Term Structure Dynamics in a Model with Stochastic Internal Habit

By Qiang Dai*

February 3, 2003

Abstract

The expectations hypothesis is violated in the U.S. nominal term structure of interest rates. This so-called "expectations puzzle" can be captured by some descriptive models of the term structure with time-varying risk premium. This paper provides an economic explanation for the success of these reduced-form specifications of pricing kernels. Specifically, we show that a consumption-based asset pricing model with stochastic internal habit, when calibrated to aggregate real macroeconomic and asset market data, generates market prices of risk with all of the requisite properties to resolve the expectations puzzle.

Keywords: stochastic internal habit, habit persistence, habit shock, time-varying risk premium, expectations puzzle, equity premium puzzle, riskfree rate puzzle

*The author is an assistant profession of finance of the Stern School of Business at New York University. Email: qdai@stern.nyu.edu. Phone: (212) 998-0358. I would like to thank Ken Singleton for extensive comments and suggestions on the paper, and many others, especially Yakov Amihud, John Campbell, John Cochrane, Darrell Duffie, Jun Liu, Anthony Lynch, and Suresh Sundaresan for their comments to earlier drafts of the paper.

1 Introduction

This paper examines the implications of a consumption-based asset pricing model with stochastic habit formation for the dynamic properties of the term structure of interest rates. The habit stock in this model is "internal" in the sense that its future evolution is affected by the current choice of aggregate consumption. Additionally, the habit shock to the preference is stochastic in the sense that it is not perfectly predictable given knowledge of current and past consumption decisions. We argue that these features together go a long ways toward resolving some the pricing anomalies in the literature on the terms structure of interest rates.

There is now a large literature demonstrating that the introduction of habit formation into preference-based models generates equity risk premiums that are more consistent with history than standard time-separable specifications of preferences (see, e.g., Sundaresan (1989), Constantinides (1990), and Campbell and Cochrane (1999)). However, these models were parameterized to yield constant interest rates in order to focus on the equity premium puzzle. As such, they are, by construction, silent on the implications of habit formation for the pricing of bonds. Stochastic interest rates are recovered in these models, simply by relaxing the endowment specification. However, these extended models are nested special cases of the flexible specification of non-time-separable preferences in Dunn and Singleton (1986), Ferson and Constantinides (1991), and Heaton (1995). Based on Euler-equation based tests that are agnostic about the distribution of consumption growth and asset returns, these authors present substantial empirical evidence against such models. Moreover, the evidence against models with deterministic habit is particularly strong when the tests involve bond yield data (see, e.g., Dunn and Singleton (1986) and Hansen and Jagannathan (1991)). One possible interpretation of this finding is that the autocorrelation properties of marginal rates of substitution (MRS) with deterministic habit are not consistent with the autocorrelation properties of bond returns (see Singleton (1993)).[1]

In this paper we extend the internal habit model of Constantinides to allow for a stochastic habit shock and examine the dynamic properties of the implied term structures. The stochastic *and* internal nature of the habit shock are crucial in taking our model outside the framework of Dunn and Singleton (1986), Ferson and Constantinides (1991), and Heaton (1995), and set the stage for a novel source of persistence in the MRS. Specifically, if the expected growth rate of the habit stock is a convex function of the surplus consumption, the MRS depends not only on the growth rate of the surplus consumption, but also the current and future level of the per capita surplus consumption. This level dependence of the MRS implies that there is a long-run value toward which the MRS is mean-reverting. We argue that these features potentially allow our model to address the tension between autocorrelation properties of bond returns and those of the MRS with deterministic habit. In particular, we demonstrate that, when calibrated to aggregate real macroeconomic and asset market data, our model is capable of resolving the so-called "expectations puzzle" (Campbell and Shiller (1991)) suggesting that that expected bond returns are predictable by term spreads

[1] Holding period returns on stocks show very little serial correlation, so the autocorrelation of the MRS may not be as central an issue in previous studies of equity returns. Though of course the same MRS should price both stocks and bonds.

(Fama (1984) and Fama and Bliss (1987)). This gives an economic explanation for the success of some reduced-form specifications of pricing kernels in explaining the expectations puzzle (see, e.g., Dai and Singleton (2002a) and Duffee (2002)).

The remainder of this paper is organized as follows. In Section 2, we start with a discrete-time formulation of stochastic internal habit and discuss how various ingredients of our habit specification contribute to the autocorrelation properties of the MRS. We then switch to the continuous-time formulation and solve the rational expectations equilibrium of the resulting model. Asset pricing implications are derived with a particular emphasis on the model-implied dynamics of the real term structure. In Section 3, we calibrate the preference parameters in the model to match the long-run average of aggregate consumption growth rate, the aggregate equity premium, and the real riskfree rate, and show that the model has enough degrees of freedom left to match the slope coefficients in the Campbell-Shiller "long-rate" regressions. Thus, the equity premium puzzle (Mehra and Prescott (1985)), the riskfree rate puzzle (Weil (1989)), and the expectations puzzle (Campbell and Shiller (1991)) can be resolved simultaneously in the model. Section 4 concludes.

2 A Model of Stochastic Internal Habit

To motivate our model and help develop some economic intuitions, we begin with a discrete-time formulation of our habit specification. The continuous-time formulation follows directly by taking the time interval to 0. Subsequent development of the model, its solution, and its asset pricing implications will be based on the continuous-time formulation.

2.1 Discrete-time Formulation of Stochastic Habit

Consider first the discrete-time formulation of Constantinides' habit stock. Let x_t be the per capita habit stock.[2] Then Constantinides' specification of the habit stock is given by

$$x_{t+1} = B(L)\,(c_t - x_t), \tag{1}$$

where c_t is the per capita consumption, L is the lag operator, and $B(L) = \sum_{j=0}^{\infty} b\,(1-\kappa)^j L^j$ is a *geometric* distributed lag with $b > 0$ and $\kappa > 0$. Equation (1) may be rewritten as $x_{t+1} = A(L)c_t$, where $A(L) = \sum_{j=0}^{\infty} b\,(1-b-\kappa)^j L^j$. Thus the habit stock is a weighted average of past consumption and is locally deterministic.

In this paper, we consider a model in which the per capita habit stock is defined by

$$x_{t+1} = B(L)\,(1 + y_{t+1})\,z_t, \tag{2}$$

[2]The purpose of working with per capita rather than aggregate variables is to take out the time trend, so that the per capita variables are stationary. We assume that the time trend grows at the rate of G_t, which represents the long-run growth rate of the economy. G_t can potentially be a persistent stochastic process itself, although it will be taken as a constant in our calibration exercise. The mapping between the parameters of the aggregate economy and those of the per capita economy will be made explicit in the calibration exercise.

where $z_t \equiv c_t - x_t$ is the per capita *surplus consumption*, and y_t is stochastic process of the following form:

$$y_{t+1} = \zeta(z_t - \bar{z}) + \beta v_{t+1}, \tag{3}$$

where v_{t+1} is a conditionally zero-mean normal variate and may be contemporaneously correlated with the consumption shock, $\bar{z}$ is interpreted (eventually) as the long-run mean of the per capita surplus consumption, and $\zeta > 0$ and $\beta > 0$. Clearly, if $\zeta = \beta = 0$, then $y \equiv 0$, and (2) reduces to (1). When $\beta \neq 0$, y_{t+1} is stochastic. In this case, the habit stock may still be viewed as a distributed lag of past consumption, except that the weights are now perturbed by y_{t+1}. The dating of the shock is such that the habit stock is locally stochastic. Depending on its sign, the effect of the shock is to weigh the lagged z_t more or less in forming the habit stock at $t+1$.

A positive ζ is a subtle but crucial feature of the model. Mechanically, what the conditional mean of y_{t+1} does is to weigh the lagged z_t more when $z_t > \bar{z}$ and less when $z_t < \bar{z}$ in forming the habit stock at $t+1$. This means that the expected habit stock at $t+1$ is a convex function of the lagged z_t. Thus, the effect associated with a positive ζ can be referred to as the *convexity effect*.[3] Economically, the convexity effect imparts a notion of "long-run mean" or "persistence" of z_t into the model (without persistence there is no long-run mean). More importantly, the convexity effect also imparts a notion of "long-run mean" or "persistence" of the MRS into the model. To elaborate the last point, let us derive the MRS, assuming that the period utility is a power function of $z = c - x$ with a constant and positive curvature $1-\gamma$ and that the subjective discount rate is a positive constant ρ.

2.2 Stochastic Internal Habit and MRS

Following Dunn and Singleton (1986), we can show that the MRS is given by

$$\mathrm{MRS}_{t,t+1} = e^{-\rho}\,\frac{\mathrm{MUC}_{t+1}}{\mathrm{MUC}_t}, \quad \mathrm{MUC}_t = E_t\left[\sum_{j=0}^{\infty} e^{-\rho j}\, a_{t,j}\, z_{t+j}^{\gamma-1}\right], \tag{4}$$

where $a_{t,0} = 1$, and for $\forall j \geq 1$,

$$a_{t,j} = \left[\prod_{i=0}^{j-1} -(1-\kappa) \prod_{i=1}^{j-1}\right] \left[1 - \kappa - b(1 + y_{t+i+1} + \zeta z_{t+i})\right], \tag{5}$$

which is in the information set at $t+j$. When κ, β, and ζ and the length of the time interval are sufficiently small, $\{a_{t,j} : j \geq 1\}$ can be approximated by

$$a_{t,j} \approx -a^j\left(1 + \zeta(2z_t - \bar{z}) + \beta v_{t+1}\right), \; a^j = b\left(1 - \kappa - b\right)^{j-1}, \; \forall j \geq 1.$$

These coefficients are referred to as "internal" coefficients because they are identically zero in an "external" habit formation model.

[3] Obviously, a negative ζ will induce a "concavity effect". The resulting model has undesirable properties and is not considered.

Armed with these results, we can now delineate the marginal contributions of various ingredients of our habit specification to the autocorrelation properties of the MRS:

1. With a *deterministic external* habit, $\beta = \zeta = 0$ and $a_{t,j} = 0$, for $\forall j \geq 1$, the marginal utility of consumption (MUC) is given by

$$\mathrm{MUC}_t^{DE} = z_t^{\gamma-1}.$$

   In this case, the autocorrelation properties of the MRS are determined by the autocorrelation properties of z_t. If z_t were a random walk, then so would be the MRS, and both the riskfree rate and the Sharpe ratio would be constant. Campbell and Cochrane (1999) assume that the endowment process c_t is a random walk, and the log surplus consumption ratio $s_t \equiv \log\left(\frac{z_t}{c_t}\right)$ is persistent. Consequently, z_t is persistent and the riskfree rate is stochastic in general.[4]

2. With a *deterministic internal* habit, $\beta = \zeta = 0$, $a_{t,j} = a^j$, $\forall j \geq 1$, the MUC is given by

$$\mathrm{MUC}_t^{DI} = \mathrm{MUC}_t^{DE} - \sum_{j=1}^{\infty} a^j \, e^{-\rho j} \, E_t\left[z_{t+j}^{\gamma-1}\right].$$

   The internal nature of the habit stock induces the extra terms that are absent in an external habit formation model.[5] These extra terms represent an additional source of persistence in the MRS.[6] Treating a^j as free parameters, the internal habit formation models of Dunn and Singleton (1986), Ferson and Constantinides (1991), and Heaton (1995) nest both the external habit formation models of Campbell and Cochrane (1999) and the internal habit formation model of Constantinides (1990).

3. With a *stochastic internal* habit ($\beta > 0$), but *no convexity effect* ($\zeta = 0$), the MUC is given by

$$\mathrm{MUC}_t^{SI} \approx \mathrm{MUC}_t^{DI} - \beta \sum_{j=1}^{\infty} a^j \, e^{-\rho j} \, E_t\left[v_{t+1} \, z_{t+j}^{\gamma-1}\right].$$

   This is an intermediate case that helps illustrate the role of the convexity effect. When the convexity effect is absent, the habit shock induces an extra term through its correlation with production shocks. This extra term does not fundamentally change the autocorrelation properties of the MRS: if z_t were a random walk, and its shock has a

[4] In focusing on equity returns, Campbell and Cochrane (1999) impose a restriction on the conditional volatility of s_t or equivalently that of z_t so that the precautionary savings motive renders the riskfree rate constant.

[5] This result is first derived by Dunn and Singleton (1986) without imposing the geometric restriction on the distributed lag $B(L)$.

[6] In the special case that z_t is a random walk, MUC_t^{DI} is proportional to $z_t^{\gamma-1}$, as in an external habit formation model. This is precisely the Constantinides result – the point of departure for his analysis of equity returns.

constant correlation with the habit shock, then MUC_t^{SI} is again proportional to $z_t^{\gamma-1}$, and the riskfree rate is again constant. Of course, stochastic interest rates can still be induced by a persistent process for z_t.

4. Finally, when the stochastic internal habit is formed with a convexity effect, i.e., both $\beta \neq 0$ and $\zeta \neq 0$, the MUC becomes

$$\mathrm{MUC}_t \approx \mathrm{MUC}_t^{SI} - \zeta\,(2z_t - \bar{z}) \sum_{j=1}^{\infty} a^j\, e^{-\rho j}\, E_t\left[z_{t+j}^{\gamma-1}\right].$$

The last term in the above equation introduces a novel source of autocorrelation in the MUC and hence in the MRS: even if z_t were a random walk, the last term would contain a term proportional to z_t^{γ}. Consequently, MUC_t would no longer be proportional to $z_t^{\gamma-1}$. In other words, the convexity effect imparts a notion of "long-run mean" or "persistence" in the MUC, because it is no longer indifferent to the level of z_t.[7] The resulting *level dependence* of the MRS is responsible for inducing a state-dependent riskfree rate and state-dependent Sharpe ratios in a Markovian equilibrium of our model.

The preceding discussion highlights the importance of both the stochastic and internal nature of the habit shock v_{t+1}. Without either feature, the model would be nested in those of Dunn and Singleton (1986), Ferson and Constantinides (1991), and Heaton (1995) and therefore would be econometrically rejected by the data. The discussion (point 3) also shows, however, that these two features are not sufficient to generate a level-dependence in the MRS. The convexity effect, which induces the level-dependence in the MRS, is indeed a crucial part of our model specification.

Up to this point, we have not taken any stand on the dynamic specification of the aggregate endowment process or equivalently the dynamic specification of z_t. Using observed aggregate consumption data to characterize the aggregate endowment process has two drawbacks. First, the consumption data is noisy. Consequently, it is challenging to econometrically pick up possible persistent and heteroskedastic features of the data generating process. Second, perhaps more importantly, if we take a model of internal habit seriously, we should not rule out the possibility that the conditional moments of the endowment process may depend on the habit level, just as the conditional moments of the habit process depend on the consumption level. Since the habit process is not observed, the aggregate consumption data, even if it is pristine, may not contain enough information to identify the *joint dynamics* of endowment and habit. These considerations lead us to consider an alternative (conceptually equivalent) formulation of the model, namely, the Merton-style continuous-time formulation of a dynamic asset pricing model that makes explicit consumption and portfolio decisions (see, e.g., Merton (1969), Breeden (1979)). This approach has two advantages. First, instead of being explicit about the joint dynamics of endowment and habit, we only need to impose

[7] This form of MUC can be generated in an external habit model with period utility of the form $\frac{z^{\gamma}}{\gamma}\left(1 + \theta \times \frac{z}{\bar{z}}\right)$, where $\theta > 0$.

the assumption that the investment opportunity set be Markovian, which is tantamount to assuming that the joint dynamics of endowment and habit be Markovian. The functional form of the investment opportunity set is then determined endogenously as part of a rational expectations equilibrium. Second, the Markovian assumption allows us to solve the model analytically using the standard dynamic programming method. The rest of the paper is based on this formulation, starting with the continuous-time formulation of the habit process in the next subsection.

2.3 Continuous-time Formulation of Stochastic Habit

Taking the continuous-time limit of equations (2) and (3), we obtain the fundamental assumption of this paper:

Assumption 1 : The habit stock x_t evolves in continuous-time according to

$$dx_t = \left(bz_t - \kappa x + \frac{\epsilon\beta^2}{2} z_t^2 \right) dt + \beta\, z_t\, dB_t^x, \tag{6}$$

where $b > 0$, $\kappa > 0$, $\epsilon > 0$, $\beta > 0$, and B_t^x is a standard Brownian motion.[8]

The quadratic term in the conditional mean of the habit stock captures the "convexity" effect induced by the conditional mean of y_{t+1} in the discrete-time formulation. The expected increase in x_t induced by a positive shock to z_t is larger than the expected decrease in x_t induced by a negative shock to z_t with the same magnitude.

To complete our model we follow Constantinides (1990) and assume:

Assumption 2 There exists a representative consumer with a power period utility $\frac{(c_t - x_t)^\gamma}{\gamma}$ with $\gamma < 0$ and a constant subjective discount rate $\rho > 0$;

Assumption 3 The consumer can insure against future consumption risk by trading a risky asset with instantaneous expected return μ_t and instantaneous volatility σ_t, and a riskfree asset with instantaneous return r_t; and

Assumption 4 Markets clear in the sense that in equilibrium, the demand of the riskfree asset is equal to the supply, which is assumed to be 0.

Under Assumptions 1–4, the representative consumer solves

$$\max_{\alpha_t, c_t : t \geq 0} E_0 \left[\int_0^\infty e^{-\rho t} \frac{(c_t - x_t)^\gamma}{\gamma}\, dt \right], \tag{7}$$

subject to the budget constraint

$$dw_t \;=\; [\alpha_t(\mu_t - r_t)w_t + r_t w_t - c_t]\, dt + \alpha_t \sigma_t w_t\, dB_t^w, \tag{8}$$

[8] The parameters in Equation (6) are reparameterized version of those that appear in equations (1) and (2). The restriction $\epsilon > 0$ is the same as the restriction $\zeta > 0$. The Brownian motion dB_t^x is the standardized version of v_{t+1} in the continuous-time limit.

and the habit specification (2), where w_t is the per capita aggregate wealth and B_t^w is a standard Brownian motion with a constant correlation δ with B_t^x. Throughout the paper, dB^w is referred to as the *production shock*, whereas dB^x is referred to as the *habit shock*.

2.4 The Markovian Solution

We will solve the model under the following assumption on the technological specification of the economy:

Assumption 5 The investment opportunity set is Markovian. That is, r_t, μ_t, and σ_t are at most functions of w_t and x_t.

Assumption 5 is a relatively weak restriction, since there is no need to specify a priori the functional form of the investment opportunity set (which will be determined endogenously).[9]

Under Assumptions 1–5, the representative consumer's utility maximization problem can be equivalently characterized in terms of the following Hamilton-Jacobi-Bellman equation:

$$0 = \max_{\alpha,z} \left[u(z,t) + V_t + \mathcal{A}^{\alpha,z} V \right], \tag{9}$$

where $u(z,t) = e^{-\rho t} \frac{z^\gamma}{\gamma}$, V is the value function given by

$$V(w_t, x_t; t) = \max_{\alpha_s, z_s : s \geq t} E_t \left[\int_0^\infty u(z_s, s)\, ds \right],$$

and $\mathcal{A}^{\alpha,z}$ is the infinitesimal generator given by[10]

$$\begin{aligned} \mathcal{A}^{\alpha,z} &= \left\{ \left[\alpha(\mu(w,x) - r(w,x)) + r(w,x) \right] w - z - x \right\} \partial_w + \frac{1}{2}\alpha^2 \sigma(w,x)^2 w^2 \partial_{ww} \\ &+ \left(bz - \kappa x - \frac{\epsilon \beta^2}{2} \right) \partial_x + \frac{\beta^2 z^2}{2} \partial_{xx} + w\alpha\sigma(w,x)\delta\beta z \partial_{xw}. \end{aligned}$$

The first-order conditions, $\frac{\partial \mathcal{A}^{\alpha,z}}{\partial \alpha} V = 0$ and $u_z + \frac{\partial \mathcal{A}^{\alpha,z}}{\partial z} V = 0$, imply that the optimal portfolio policy is given by

$$\alpha^* = A(w,x)^{-1} \sigma(w,x)^{-1} \left[\Lambda(x) - H(w,x)\delta\beta z^* \right], \tag{10}$$

[9] When $\beta = \epsilon = 0$, the model reduces to that of Constantinides (1990). The rational expectations equilibrium of the model is characterized by a constant riskfree rate and a constant Sharpe ratio of the risky asset. When $\beta \neq 0$ and $\epsilon \neq 0$, however, both the riskfree rate and the Sharpe ratio will become state-dependent – otherwise markets can not clear in every state of the world. In general, this presents a challenge in solving the model. Fortunately, under the particular formulation we have adopted, the equilibrium investment opportunity set takes a simple form and can be easily guessed (see equations (12) –(14)).

[10] For notational simplicity, we have made a change of variable from c to $z = c - x$ in the optimization problem.

where $A(w,x) = -\frac{wV_{ww}}{V_w}$ is the relative risk aversion with respect to wealth and $H(w,x) = -\frac{V_{wx}}{V_w}$; and the optimal consumption policy is given by $c^* = z^* + x$, where z^* solves:

$$u_z = V_w - \left(b + \epsilon\beta^2 z\right) V_x - \beta^2 z V_{xx} - w\alpha^* \sigma(w,x)\delta\beta z V_{xw}. \tag{11}$$

To obtain explicit expressions for the optimal policies, we conjecture that the investment opportunity set takes the following form:

$$r(w,x) = \frac{1-\kappa f'(x)}{f(x)/x} \equiv r(x), \tag{12}$$

$$\mu(w,x) = r(x) + \Lambda(x)\sigma(w,x), \tag{13}$$

$$\sigma(w,x) = \frac{w - f(x)}{w} g(x), \tag{14}$$

where $f(x) = \eta\frac{1-e^{-\epsilon x}}{\epsilon}$ with $\eta > 0$, and the functions $g(x)$ and $\Lambda(x)$ are to be determined.[11] It is easy to verify that, under this conjecture, the optimization problem is solved with the value function given by

$$V(w,x,t) = e^{-\rho t} J(w,x), \text{ where } J(w,x) = \phi\frac{(w-f(x))^\gamma}{\gamma}, \phi > 0, \tag{15}$$

and the optimal portfolio and consumption policies given by

$$\alpha^* = \frac{\Lambda(x) + (1-\gamma)\delta\beta h(x) f'(x)}{(1-\gamma)g(x)}, \quad z^* = h(x)[w - f(x)], \tag{16}$$

where

$$\frac{h(x)^{\gamma-1}}{\phi} = 1 + bf'(x) + (1-\gamma)\beta^2 h(x) f'(x)^2 - (1-\gamma)\delta\beta g(x) f'(x). \tag{17}$$

The unknown functions $g(x)$ and $\Lambda(x)$ are obtained by using the Bellman optimality principle and the market clearing condition, respectively. First, substituting the optimal policies (16) into (9) yields

$$\begin{aligned} 0 = {} & \frac{1-\gamma}{2} g(x)^2 - (1-\gamma)\delta\beta h(x) f'(x) g(x) \\ & + \left[r(x) - \frac{\rho}{\gamma} + \frac{1-\gamma}{\gamma}\frac{h(x)^\gamma}{\phi} + \frac{1-\gamma}{2}\beta^2 h(x)^2 f'(x)^2 \right]. \end{aligned} \tag{18}$$

The functions $h(x)$ and $g(x)$ are therefore obtained by solving equations (17) and (18) jointly, state by state. Once $h(x)$ and $g(x)$ are known, the function $\Lambda(x)$ is obtained by imposing the market clearing condition: $\alpha^* = 1$ implies that

$$\Lambda(x) = (1-\gamma)\left[g(x) - \delta\beta h(x) f'(x)\right]. \tag{19}$$

[11] The functional form of $f(x)$ is dictated by the habit specification (6). If the coefficients in the diffusion and in the quadratic term in the drift are state-dependent, the functional form of $f(x)$ will be different. To preserve key properties of our model, any alternative habit specification must be such that the function $f(x)$ satisfies the properties: $f(0) = 0$, $f'(x) > 0$, and $f''(x) < 0$. Under our specification, the sign restriction $\epsilon > 0$ ensures that $f(x)$ is monotonically increasing and concave.

We can easily verify that all necessary and sufficient conditions (including the transversality condition) are satisfied and all markets clear. Thus the initial conjecture of the equilibrium investment opportunity set, (12)–(14), is validated.

2.5 Pricing Contingent Claims

The model contains a contingent claims pricing model, which is fully specified in terms of the pricing kernel and the state dynamics.

The Pricing Kernel

The pricing kernel m_t is proportional to the marginal utility of wealth, i.e., $m_t \propto V_w = e^{-\rho t}\phi(w - f(x))^{\gamma-1}$. Ito's lemma implies

$$\frac{dm_t}{m_t} = -r(x_t)\,dt - \Lambda^w(x_t)\,dB_t^w - \Lambda^x(x_t)\,dB_t^x, \tag{20}$$

where $\Lambda^w(x) = -\frac{wV_{ww}}{V_w}\alpha^*\sigma = (1-\gamma)g(x)$ and $\Lambda^x(x) = -\frac{V_{wx}}{V_w}\beta z^* = -(1-\gamma)\beta h(x)f'(x)$.

The State Dynamics

Since $w = f(x) + \frac{z^*}{h(x)}$ is a deterministic function of x and z^*, we can alternatively treat (x, z^*) as the state vector. The equilibrium solution implies that the joint dynamics of (x, z^*) are given by

$$dx = \mu^x(x, z^*)\,dt + \beta z^*\,dB^x, \tag{21}$$

$$\frac{dz^*}{z^*} = \mu^z(x, z^*)\,dt + g(x)\,dB^w + \beta\left[\frac{zh'(x)}{h(x)} - h(x)f'(x)\right]dB^x, \tag{22}$$

where $\mu^x(x, z) = bz - \kappa x + \frac{1}{2}\epsilon\beta^2 z^2$, and

$$\begin{aligned}\mu^z(x, z) &= \left[\frac{r(x)-\rho}{1-\gamma} + \frac{2-\gamma}{2(1-\gamma)^2}\Lambda(x)^2 + \frac{2-\gamma}{2}(1-\delta^2)\beta^2 h(x)^2 f'(x)^2\right] \\ &+ \left[\frac{h''(x)}{h(x)}\frac{\beta^2 z^2}{2} + \frac{h'(x)}{h(x)}\beta z\frac{\delta\Lambda(x) + \sqrt{1-\delta^2}\Lambda_\perp(x)}{1-\gamma}\right] + \frac{h'(x)}{h(x)}\mu^x(x, z).\end{aligned}$$

The real functions $h(x)$, $g(x)$, $r(x)$, $f(x)$, $\Lambda^w(x)$, and $\Lambda^x(x)$ that appear in the pricing kernel and the state dynamics are all part of the equilibrium solution and known. The resulting contingent claims pricing model can be used to price any security whose payoffs load on the two state variables, including equity and debt claims. In the rest of this section, we use this model to determine the term structure of real interest rates and examine some of its key properties.

Pricing Real Zero-coupon Bonds

Let $D_{t,T}$ be the time-t price of a real zero-coupon bond with maturity date T, i.e.,

$$D_{t,T} = E\left[\frac{m_T}{m_t}\middle| x_t, z_t^*\right]. \tag{23}$$

In general, $D_{t,T}$ is a function of both x_t and z_t^*. Ito's lemma implies that

$$\frac{dD_{t,T}}{D_{t,T}} = \mu_{t,T}\,dt + \sigma_{t,T}^{w}\,dB_t^{w} + \sigma_{t,T}^{x}\,dB_t^{x} \tag{24}$$

where $\mu_{t,T}$ is the instantaneous expected return on the bond, $\sigma_{t,T}^{w} = g(x)\frac{\partial \log D_{t,T}}{\partial \log z^*}$, and $\sigma_{t,T}^{x} = \beta\left[\frac{z^* h'(x)}{h(x)} - h(x)f'(x)\right]\frac{\partial \log D_{t,T}}{\partial \log z^*} + \beta z^* \frac{\partial \log D_{t,T}}{\partial x}$. Since $m_t\,D_{t,T}$ must be a Martingale, the expected excess return is given by

$$\mu_{t,T} - r_t = \sigma_{t,T}^{w}[\Lambda^w(x_t) + \delta\Lambda^x(x_t)] + \sigma_{t,T}^{x}[\Lambda^x(x_t) + \delta\Lambda^w(x_t)]. \tag{25}$$

Equation (25) shows that, in general, the expected bond returns are predictable.

Substituting equation (25) in equation (24), we obtain

$$\frac{dD_{t,T}}{D_{t,T}} = r(x_t)\,dt + \sigma_{t,T}^{w}\,dB^{w,\mathbb{Q}} + \sigma_{t,T}^{x}\,dB^{x,\mathbb{Q}}, \tag{26}$$

where,

$$B_t^{w,\mathbb{Q}} = B_t^{w} + \int_0^t \lambda^w(x_s)\,ds, \qquad \lambda^w(x_s) = \Lambda^w(x_s) + \delta\Lambda^x(x_s),$$
$$B_t^{x,\mathbb{Q}} = B_t^{x} + \int_0^t \lambda^x(x_s)\,ds, \qquad \lambda^x(x_s) = \Lambda^x(x_s) + \delta\Lambda^w(x_s).$$

Let $\mathbb{P}$ be the physical probability measure under which B^w and B^x are standard Brownian motions with correlation δ, and define an equivalent probability measure $\mathbb{Q}$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{-\frac{1}{2}\int_0^T\left[(\lambda^w(x_s))^2 + (\lambda^x(x_s))^2 + 2\delta\lambda^w(x_s)\lambda^x(x_s)\right]ds - \int_0^T \lambda^w(x_s)\,dB_s^w - \int_t^T \lambda^x(x_s)\,dB_s^x}. \tag{27}$$

Then, $B_t^{w,\mathbb{Q}}$ and $B_t^{x,\mathbb{Q}}$ are standard Brownian motions with correlation δ under $\mathbb{Q}$. Thus, $\mathbb{Q}$ is precisely the risk-neutral measure, and the zero-coupon bond prices can be obtained from the following risk-neutral pricing formula:

$$D_{t,T} = E^{\mathbb{Q}}\left[e^{-\int_t^T r(x_s)\,ds}\middle| x_t, z_t^*\right]. \tag{28}$$

Equation (28), together with equations (21), (22), and (27), form a two-factor nonlinear model of the real term structure. Although bond prices are not given by an explicit analytic expression, they can be readily computed using numerical methods such as a Monte Carlo simulation or a finite-difference method.

2.6 Model-implied Properties of the Real Term Structure

The rest of the paper focuses on the term structure implications. We begin in this subsection with a discussion on some general properties of the real term structure implied by the

model, followed in the next two subsections by a discussion on how the model resolves the expectations puzzle.

First, the real term structure is affected not only by the growth rate of per capita consumption c^*, but also by the growth rate of either the habit stock x or the surplus consumption $z^* = c^* - x$. In general, both the production/consumption shock and the habit shock are priced, with market prices of risk given by $\lambda^w(x)$ (which is equal to the Sharpe ratio $\Lambda(x)$ of the risky asset) and $\lambda^x(x)$, respectively.

Second, the instantaneous riskfree rate is determined entirely by the habit stock x, and does not depend on the surplus consumption z^*. Thus, the production shock (which affects directly z^* not x^*) has a delayed effect on the term structure. That is, letting $R_{t,T} \equiv -\frac{\log D_{t,T}}{T-t}$ be the yield on a zero-coupon bond with maturity T, the loading, $\frac{\partial R_{t,T}}{\partial \log z^*}$, of the zero yield on $\log z^*$ is expected to have a humped shape: starting at zero initially, reaching a peak or valley at some intermediate maturity range, and converging to 0 again at long maturities (due to mean reversion). Thus, the production shock may potentially induce a hump in the term structure of volatility. This is consistent with the finding by Fleming and Remolona (1999) that macro-economic announcements may induce a humped response function in the term structure.[12] A humped term structure of volatility may also induce a humped term structure of expected returns: in the limit $\beta \to 0$,

$$\mu_{t,T} - r_t = \sigma^w_{t,T}\Lambda(x_t) = \left[-\frac{\Lambda(x_t)}{1-\gamma}\frac{\partial R_{t,T}}{\partial \log z^*}(T-t)\right] \times \Lambda(x_t), \tag{29}$$

where we have used the fact that $\Lambda^w(x) = \Lambda(x) = (1-\gamma)g(x)$ is the Sharpe ratio of the risky investment opportunity.[13]

Third, the stochastic shock to the habit stock induces a positive correlation between the expected bond returns and the term spreads, which implies that the "expectations hypothesis" does not hold in the model. Indeed, the degree in which the expectations hypothesis is violated is controlled by the volatility of the habit shock. This fact allows us to calibrate the parameter β. To elaborate this implication of the model, let us digress for a moment to introduce formally the notion of "expectations hypothesis" and that of "expectations puzzle".

2.7 The Expectations Hypothesis and the Puzzle

For any fixed integer $n \geq 1$, let

$$R^n_t \equiv -\frac{1}{n}\log D_{t,t+n}, \quad e^n_t \equiv E_t\left[\log\frac{D_{t+1,t+n}}{D_{t,t+n}} - R^1_t\right],$$

[12] Empirical evidence for such a humped term structure of volatility is discussed by Litterman, Scheinkman, and Weiss (1988), Dai and Singleton (2000), Leippold and Wu (2001), and Piazzesi (2001).

[13] Note that, because $\epsilon > 0$, our model does not converge to the Constantinides model as $\beta \to 0$ (in the Constantinides model, the riskfree rate is constant). If we impose the constraint that $\epsilon = 0$, then the model converges to the Constantinides result when $\beta \to 0$. This shows that our model is a more radical departure from Constantinides model than what might appear at first glance.

be the time-t yield of a n-period zero-coupon bond and the expected one-period holding period excess return on the bond, respectively. Then it is definitionally true that (see Dai and Singleton (2002a))

$$E_t\left[R_{t+1}^{n-1} - R_t^n\right] + \frac{e_t^n}{n-1} = \frac{1}{n-1}\left(R_t^n - R_t^1\right),\ n > 1. \tag{30}$$

The *expectations hypothesis* is basically the assumption that e_t^n is constant for all n. Under this hypothesis, the following regression (which is often referred to as the Campbell-Shiller "long-rate" regression)

$$R_{t+1}^{n-1} - R_t^n = \delta_n + \beta_n \frac{R_t^n - R_t^1}{n-1} + \epsilon_{t+1},\ n > 1, \tag{31}$$

should produce a slope coefficient of 1. When the regression is run on nominal Treasury yields with sample length $\mathcal{T}$, however, the sample slope coefficients $\beta_n^{\mathcal{T}}$ are not only different from 1, but are negative. Furthermore, the magnitude of $\beta_n^{\mathcal{T}}$ increases with n. These and some other related stylized facts are collectively referred to as the *expectations puzzle* (see Campbell and Shiller (1991)). Recently, Dai and Singleton (2002a) and Duffee (2002) show that the expectations puzzle can be explained by a reduced-form dynamic term structure model, provided that the market prices of risk take some flexible form so that the expected excess bond returns are positively correlated with the term spreads. These studies, however, do not provide an economic explanation for the requisite behavior of the market prices of risk, a gap that this paper tries to fill. It is important to point out that the observed violation of the expectations puzzle is an empirical statement on nominal yields, rather than real yields, and there are both theoretical (see, e.g., McCallum (1994)), and empirical (see, e.g., Kugler (1997)) reasons to believe that monetary policies may play an important role in the violation of the expectations hypothesis in nominal yields. Nevertheless, it is useful to know to what extent the puzzle may be explained by real effects alone.

The ability of an affine term structure model to capture the violation of the expectations hypothesis can be illustrated by the following one-factor extended Gaussian model:[14]

$$D_{t,T} = E^{\mathbb{Q}}\left[e^{-\int_t^T r(x_s)\,ds}\middle| x_t\right], \tag{32}$$

$$r(x) = \delta_0 + \delta_1 x, \quad dx = \kappa(\bar{x} - x)dt + \Sigma dB^x, \tag{33}$$

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = e^{-\frac{1}{2}\int_0^T \lambda(x_s)^2\,ds - \int_t^T \lambda(x_s)\,dB_s^x}, \quad \lambda(x) = \Sigma^{-1}\left(\lambda_0 + \lambda_1 x\right), \tag{34}$$

where $\delta_1 > 0$, $\kappa > 0$, and $\lambda_1 > 0$. Equations (33) and (34) imply that the risk-neutral dynamics of the state variable x is given by

$$dx = \kappa^{\mathbb{Q}}(\bar{x}^{\mathbb{Q}} - x)\,dt + \Sigma\,dB^{x,\mathbb{Q}}, \quad \text{where} \quad \kappa^{\mathbb{Q}} = \kappa + \lambda_1 \text{ and } \bar{x}^{\mathbb{Q}} = \frac{\kappa\bar{x} - \lambda_0}{\kappa^{\mathbb{Q}}}. \tag{35}$$

[14]This model and its multi-factor extensions are used by Dai and Singleton (2002a) to explain the expectations puzzle.

Substituting (35) into the bond pricing equation (32), we obtain

$$D_{t,T} = e^{-A(T-t)-B(T-t)r_t},$$

$$\text{where} \qquad A(\tau) = \left(\delta_0 + \delta_1 \bar{x}^{\mathbb{Q}} - \frac{\delta_1^2 \Sigma^2}{2\kappa^{\mathbb{Q}}}\right)(B(\tau) - \tau) - \frac{\delta_1^2 \Sigma^2}{4\kappa^Q} B(\tau)^2 \text{ and } B(\tau) = \frac{1 - e^{-\kappa^{\mathbb{Q}}\tau}}{\kappa^{\mathbb{Q}}}.$$

From the loadings $A(\tau)$ and $B(\tau)$, we can compute

1. the n-period zero coupon bond yield: $R_t^n = \frac{A(n)}{n} + \frac{B(n)}{n} r_t$;
2. the one-period forward rate with settlement date n periods hence:

$$f_t^n = [A(n+1) - A(n)] + [B(n+1) - B(n)]\, r_t;$$

3. The conditional expectation of one-period spot rate prevailing at $t+n$:

$$E_t[R_{t+n}^1] = A(1) + B(1)\left[\left(1 - e^{-\kappa n}\right)\bar{r} + e^{-\kappa n} r_t\right].$$

It follows that the linear projection coefficient of the forward premium, $f_t^n - E_t[R_{t+n}^1]$, onto the forward spread, $f_t^n - R_t^1$, is given by

$$\alpha_n = \frac{e^{-\kappa n} - [B(n+1) - B(n)]/B(1)}{1 - [B(n+1) - B(n)]/B(1)} \approx \frac{e^{-\kappa n} - e^{-\kappa^{\mathbb{Q}} n}}{1 - e^{-\kappa^Q n}} \approx \frac{\kappa^{\mathbb{Q}} - \kappa}{\kappa^{\mathbb{Q}}} = \frac{\lambda_1}{\kappa + \lambda_1},$$

where the approximation holds if κ^Q, κ, and the observation internal are sufficiently small. If $\kappa > 0$ and $\lambda_1 > 0$, then $0 < \alpha_n < 1$. If $\beta = 0$, then $\lambda_1 = 0$ and $\alpha_n = 0$. Thus, the forward premium is predictable by the forward spread if and only if the habit stock is stochastic ($\beta \neq 0$).

Backus, Foresi, Mozumdar, and Wu (2001) show that the predictability of forward premium by forward spread and the expectations puzzle are two sides of the same coin. We can confirm this by computing the linear projection coefficient β_n of the yield change, $R_{t+1}^{n-1} - R_t^n$, onto the scaled term spread, $\frac{1}{n-1}(R_t^n - R_t^1)$, i.e.,

$$\beta_n = \frac{\operatorname{cov}\left(R_{t+1}^{n-1} - R_t^n, \frac{R_t^n - R_t^1}{n-1}\right)}{\operatorname{var}\left(\frac{R_t^n - R_t^1}{n-1}\right)} = (n-1) \times \frac{\operatorname{cov}\left(R_{t+1}^{n-1} - R_t^n, R_t^n - R_t^1\right)}{\operatorname{var}\left(R_t^n - R_t^1\right)},$$

which are the population counterparts of $\beta_n^{\mathcal{T}}$. It is easy to show that if $0 < \alpha_n < 1$ and roughly constant, then β_n becomes negative and increases in magnitude with n at sufficiently large n – capturing the essence of the expectations puzzle. Furthermore, the speed at which $|\beta_n|$ increases with n is controlled by the value of λ_1 or equivalently the volatility of the habit shock β. These results will be illustrated graphically in the next section.

2.8 Stochastic Habit and the Expectations Puzzle

We will now close the loop by showing that the extended one-factor Gaussian model (32)–(34) emerges as an approximation to the two-factor term structure dynamics implied by our model of stochastic internal habit. As a by-product, we will derive the reduced-form parameters $(\delta_0, \delta_1, \kappa, \bar{x}, \Sigma, \lambda_0, \lambda_1)$ in terms of the structural parameters $(\rho, \gamma, b, \kappa, \beta, \delta, \bar{z}, \epsilon)$. The mean-reversion coefficient κ that appears in both parameter vectors is indeed the same: mean reversion of the interest rate comes directly from habit persistence in this model.

An one-factor approximation of the structural model is obtained by shutting down the effect of the production shock on the term structure. This is done by setting z^* to its long-run mean $\bar{z}$ in the dynamics of the habit stock x (21), and by integrating out the production shock in (27). The resulting model is given by

$$\begin{aligned} D_{t,T} &= E^{\mathbb{Q}}\left[e^{-\int_t^T r(x_s)\,ds}\middle| x_t\right], \quad r(x) = \frac{1-\kappa f'(x)}{f(x)/x}, \quad f(x) = \eta\frac{1-e^{-\epsilon x}}{\epsilon}, \\ dx &= \left(b\bar{z} - \kappa x + \frac{\epsilon\beta^2}{2}\bar{z}^2\right)dt + \beta\bar{z}dB^x, \\ \frac{d\mathbb{Q}}{d\mathbb{P}} &= e^{-\frac{1}{2}\int_0^T \lambda^x(x_s)^2\,ds - \int_t^T \lambda^x(x_s)\,dB_s^x}, \quad \lambda^x(x) = \delta\Lambda^w(x). \end{aligned}$$

The one-factor extended Gaussian model (32)–(34) is then obtained by linearizing the riskfree rate and the market price of risk around the long-run mean of x.[15] This procedure provides a mapping between the reduced-form parameters and the structural parameters:

$$\delta_0 = \bar{r} - \delta_1\bar{x}, \quad \delta_1 = \bar{r}\left[\frac{-\kappa f''(\bar{x})}{1-\kappa f'(\bar{x})} + \frac{f(\bar{x}) - \bar{x}f'(\bar{x})}{\bar{x}f(\bar{x})}\right], \quad \bar{r} = \frac{1-\kappa f'(\bar{x})}{f(\bar{x})/\bar{x}}, \tag{36}$$

$$\bar{x} = \frac{b}{\kappa}\bar{z} + \frac{\epsilon\beta^2}{2\kappa}\bar{z}^2, \quad \Sigma = \beta\bar{z}, \quad \lambda_0 = \delta\beta\bar{z}\Lambda(\bar{x}) - \lambda_1\bar{x}, \quad \lambda_1 = \delta\beta\bar{z}\Lambda'(\bar{x}). \tag{37}$$

The sign restrictions $\kappa > 0$, $\delta_1 > 0$, and $\lambda_1 > 0$ in the reduced-form model can now be explained in terms of the theoretical properties of the structural model. First, $\kappa > 0$ because habit is persistent. Second, $\delta_1 > 0$ because $r(x)$ is monotonically increasing in x. Finally, $\lambda_1 > 0$ because the habit shock B^x is negatively correlated with the production shock B^w, namely $\delta < 0$, and the Sharpe ratio $\Lambda(x)$ is monotonically decreasing in x under very general conditions.[16]

[15] The one-factor Quadratic-Gaussian (QG) model is obtained by a quadratic approximation of $r(x)$ and a linear approximation of $\lambda(x)$. Dai and Singleton (2002b) report that when calibrated to the observed forward rates, this model has virtually the same performance as the one-factor extended Gaussian model in explaining the expectations puzzle. The multi-factor versions of the QG model are extensively studied by, e.g., Lu (1999), Leippold and Wu (2001), and Ahn, Dittmar, and Gallant (2002).

[16] The structural model does not always have the properties $r'(x) > 0$ and $\Lambda'(x) < 0$. The requirement that the model possesses these properties amounts to an extra assumption that characterizes implicitly a region in the parameter space in which the model is consistent with the expectations puzzle (at least in terms of the direction in which the expectations hypothesis is violated). As we shall see in the next section, this is not a strong restriction.

3 Calibration

The preceding sections have laid out the theoretical model and its term structure implications, and have established a close linkage between the expectations puzzle and the volatility of the habit shock. However, we have left open the question of whether there exist reasonable values of the structural parameters under which the model can match the Campbell-Shiller regression coefficients computed from the observed data. In this section, we answer this question by calibrating the model to observed macro and asset return moments, and compute Campbell-Shiller regression coefficients explicitly under different volatility levels of the habit shock. To facilitate comparison with extant models, we will calibrate our model to both sample moments used by Campbell and Cochrane (1999) and those used by Constantinides (1990). In particular, we will force our model to fit the average level of the riskfree rate and the Sharpe ratio. Thus, the riskfree rate puzzle and the equity premium puzzle are absent in the model, no matter which set of sample moments are used.

We begin by calibrating the model to sample moments used by Campbell and Cochrane (1999): the historical mean of consumption growth rate is $G = 1.89\%$, volatility of consumption growth rate is $\sigma_c = 1.5\%$, historical mean of riskfree rate is $R = 0.94\%$, and the the historical average of the equity Sharpe ratio is $\bar{\Lambda} = 0.4341$. To calibrate the correlation between the habit shock and the production shock, we use the historical average Sharpe ratio of real bonds is $\bar{\Lambda}_P = 0.374$ from Campbell and Viceira (2001).

In linking the model to the sample moments, we need to make three assumptions. First, we assume that the population growth rate (perhaps augmented by technology growth) is the same as the aggregate consumption growth rate, and is equal to G. This means that the subjective discount rate in the per capita economy is equal to that in the aggregate economy plus G, the mean reversion coefficients of the per capita variables are equal to those of the aggregate variables plus G, the annual returns and interest rates in the per capita economy are equal to equal to those in the aggregate economy minus G, and the contemporaneous ratios, volatilities, correlations in the per capita economy are the same as those in the aggregate economy. This implies, for an example, that the long-run value of the riskfree rate $r(x)$, defined by $\bar{r} = r(\bar{x})$, is equal to $R - G = -0.95\%$. Second, since the habit process is not observed, the parameters b, κ, and β can not be calibrated directly. What we will do is to fix them at some plausible values, and calibrate the preference parameters from moments of asset returns and consumption growth. Finally, it is easy to check that, if we don't care about the absolute scale of the economy, then ϵ is not identified. Thus, without loss of generality, we will set $\epsilon = 1$. Consequently, $f(x) = \eta\,(1 - e^{-x})$.

The rest of the subsection gives a detailed illustration of the calibration procedure and the performance of the resulting model in explaining the expectations puzzle. Initially, we fix $\gamma = -2$, $\kappa = 0.0232$, and $\beta = 0.5$, which is equivalent to fixing $b = 0.201$, $\kappa = 0.0232$, and $\beta = 0.5$. The benefit of fixing γ rather than b is that it allows us to calibrate the rest of the parameters one by one using an associated moment in a sequential manner, rather than calibrate several parameters jointly. We will then tabulate and discuss calibration results for other parameter configurations.

The Subjective Discount Rate

The subjective discount rate ρ can be obtained by imposing the restriction (for stationarity) that the expected growth rate of z^* is equal to 0 at $z^* = \bar{z}$ and $x = \bar{x}$. Ignoring terms associated with β, we have

$$\rho \approx \bar{r} + \frac{2-\gamma}{2(1-\gamma)}\bar{\Lambda}^2 = 0.116.$$

Correlation between Production and Habit Shocks

It is straight-forward to show that under the one-factor Gaussian approximation, the average Sharpe ratio of the real bonds is equal to $-\delta\Lambda(x)$. Thus,

$$\delta = -\frac{\bar{\Lambda}_P}{\bar{\Lambda}} = -0.8615 < 0.$$

Habit Parameter

Since $c^* = z^* + x$, and the volatility of z^* is approximately $g(x) = \frac{\Lambda(x)}{1-\gamma}$, we have

$$\sigma_c \approx \frac{\bar{z}}{\bar{z}+\bar{x}}\frac{\Lambda(\bar{x})}{1-\gamma} \approx \frac{\kappa}{b+\kappa} \times \frac{\bar{\Lambda}}{1-\gamma} = 1.5\%.$$

It follows that $b = 0.201$.

Long-run Means of Habit Stock and Surplus Consumption

First, we assume that the riskfree rate is equal to 0% at zero habit, i.e., $R_0 = r(0) + G = 0\%$, or $r(0) = -1.89\%$. Thus,

$$\eta = \frac{1}{r(0)+\kappa} = \frac{1}{-1.89\% + 0.0232} = 230.947.$$

Next, equating the long-run mean of the riskfree rate with the historical mean of the riskfree rate yields

$$\bar{r} = r(\bar{x}) = \frac{1-\kappa f'(\bar{x})}{f(\bar{x})/\bar{x}} = -0.95\%, \tag{38}$$

which implies $\bar{x} = 0.745$.

It follows that $f(\bar{x}) = 121.310$, $f'(\bar{x}) = 109.637$, and $f''(\bar{x}) = -109.637$, which will be used to compute the reduced-form parameters δ_0 and δ_1. Furthermore, assuming that the expected growth rate of the habit stock is zero at the long-run means, we obtain

$$\bar{z} \approx \frac{\kappa}{b}\bar{x} = 0.0862.$$

State-dependence: $r'(\bar{x})$ and $\Lambda'(\bar{x})$

Differentiating (12) with respect to x, we have

$$r'(\bar{x}) = \frac{1-\kappa(f'(\bar{x})+\bar{x}f''(\bar{x}))}{f(\bar{x})} - \frac{f'(\bar{x})}{f(\bar{x})}\bar{r} = 0.0115 > 0.$$

Equation (19) implies that $\Lambda'(x) \approx (1-\gamma)g'(x)$. Differentiating equation (18) yields

$$\Lambda'(\bar{x}) \approx -(1-\gamma)\frac{r'(\bar{x}) - bf''(\bar{x})h(\bar{x})}{\Lambda(\bar{x})} = -0.432 < 0,$$

where we have used equations (19) and (18), respectively, to obtain

$$g(\bar{x}) \approx \frac{\Lambda(\bar{x})}{1-\gamma} = 0.145, \quad h(\bar{x}) \approx -\frac{\frac{\gamma}{2}g(\bar{x})^2 + \frac{\gamma}{1-\gamma}\left[r(\bar{x}) - \frac{\rho}{\gamma}\right]}{1 + bf'(\bar{x})} = 0.232\%.$$

Reduced-form Parameters

Now that we have all the information we need to compute the reduced-form parameters, we can use equations (36) and (37) to obtain

$$\delta_0 = -0.0181,\ \delta_1 = 0.0115,\ \kappa = 0.0232,\ \bar{x} = 0.745,\ \Sigma = 0.043,\ \lambda_0 = -0.028,\ \lambda_1 = 0.016.$$

Implications for the Expectations Puzzle

Based on the reduced-form parameters, we can compute the model-implied expected excess return e_t^n, and then compute the linear projection coefficient ϕ_n of e_t^n on the term spread $R_t^n - R_t^1$:

$$\phi_n = \frac{\text{cov}\left(e_t^n, R_t^n - R_t^1\right)}{\text{var}\left(R_t^n - R_t^1\right)}.$$

If the model is correct, then for each n, equation (30) implies that $1 - \phi_n$ is equal to β_n, i.e., ϕ_n represents the population value of the negative bias from 1 in the Campbell-Shiller "long-rate" regression. With sample length $\mathcal{T}$, the sample counterpart $1 - \phi_n^{\mathcal{T}}$ should be close the Campbell-Shiller regression coefficients $\beta_n^{\mathcal{T}}$. Figure 1 plots $1 - \phi_n^{\mathcal{T}}$ implied by the model (solid line) against $\beta_n^{\mathcal{T}}$. Since the model has only one factor, it misses some details near the short end (less than 2 years) and then overcompensates slightly in the long-end. Overall, however, the model does a fine job of capturing the broad trend and magnitude of the regression coefficients across the entire maturity spectrum.

Table 1 reports the parameter values calibrated above in column I, and 7 additional parameter configurations calibrated in the same way under different fixed values of γ, κ, and β. In columns II to VII, the values of γ, κ, or β are relatively small variations of those used in column I. In column VIII, the risk aversion coefficient is much larger ($1 - \gamma = 9$) and the volatility of the habit shock is much smaller ($\beta = 0.1$) compared to those in column I, but the two models are almost indistinguishable in their ability to fit the Campbell-Shiller regression coefficients. The model-implied Campbell-Shiller regression coefficients for all eight parameter configurations are reported in the top half of Table 2 and plotted in Figure 2, together with the actual Campbell-Shiller regression coefficients (the last column in Table 2 and the "cross" line indicated by "C-S").

The last row of Table 1, which is the ratio of κ and λ_1, predicts how well each model captures the expectations puzzle. When $\frac{\kappa}{\lambda_1}$ is around 1.45, as in columns I and VIII, the

model replicates nicely the biases in Campbell-Shiller regression coefficients. This value of κ/λ_1 represents some sort of a "magic number" for the U.S. nominal term structure, it implies that

$$\alpha_n \approx \frac{\lambda_1}{\kappa + \lambda_1} = \frac{1}{1.45 + 1} = 0.41,$$

which is roughly the average level of predictability of the forward premiums documented by Fama and Bliss (1987) and Backus, Foresi, Mozumdar, and Wu (2001). When $\frac{\kappa}{\lambda_1} > 1.45$, as in columns II, V, and VI, the model does not generate enough biases, because the predictability (or state-dependence) of market prices of risk is not strong enough against mean reversion of the interest rate. In contrast, when $\frac{\kappa}{\lambda_1} < 1.45$, as in columns III, IV, and VII, the model generates too much biases.

Table 1 also indicates that, keeping everything else the same (adjusting calibrated parameters if necessary so that the model remains free of the equity premium puzzle and the riskfree rate puzzle), a larger risk aversion coefficient, or a smaller mean reversion coefficient, or a larger volatility of the habit shock helps generates larger biases from the expectations hypothesis and therefore can be helpful in explaining the expectations puzzle. The effects of varying κ and β are easy to understand. It is a bit more subtle why the risk aversion coefficient also affects the model's ability to explain the expectations puzzle. The reason is that the model is required to explain *both* the equity premium puzzle and the riskfree rate puzzle. A larger risk aversion parameter implies that the equity premium puzzle can be explained at a higher level of surplus consumption ratio. Since the riskfree rate puzzle pins down the level of the per capita habit stock, a larger risk aversion parameter therefore implies a higher level of per capita surplus consumption, which in turn implies that the expected bond returns or forward premiums are more predictable (λ_1 is larger – c.f. columns II and III).[17]

The above results indicate that the model has a great deal of flexibility in explaining simultaneously the expectations puzzle, the equity premium puzzle, and the riskfree rate puzzle. Another useful illustration of this point is to calibrate the model to sample moments used by Constantinides (1990). Table 3 reports the result of this exercise. In each column, the values of $1 - \gamma$, $a = b + \kappa$, and b are fixed at one of six set of values used by Constantinides (1990), and two new parameters are added to account for stochastic nature of the habit stock: the correlation coefficient is fixed at $\delta = -0.8615$ as before, and the volatility of the habit shock is fixed at either of the following two values: 0.1 or 0.5. Note, however, that we do not take as given the subjective discount rate ρ used by Constantinides (1990), because it would imply that the long-run growth rates of consumption and habit are different. Instead, we calibrate ρ so that both growth rates are zero in the per capita or detrended economy. Columns I and VII show that when a or b is too small (the ratio of a and b is fixed), the model is incapable of explaining the expectations puzzle: the riskfree rate is decreasing in x, the Sharpe ratio is increasing in x, consequently λ_1 is negative. Since $\lambda_1 \neq 0$,

[17] The bottom half of Table 2 reports and Figure 3 plots the "risk-premium adjusted Campbell-Shiller regression coefficients", defined as the linear projection coefficients of $R^{n-1}_{t+1} - R^n_t + \frac{e^n_t}{n-1}$ on $\frac{R^n_t - R^1_t}{n-1}$. If the model captures the time-varying risk premium correctly, then equation (30) implies that these projection coefficients should be close to 1 for all n (see Dai and Singleton (2002a)). This is just a different way of viewing the same results reported above.

the expectations hypothesis is violated, but the biases are in the wrong direction. For all other parameter configurations, the expectations hypothesis is violated in the right direction. Quantitatively, however, only column VIII comes close to matching the observed Campbell-Shiller regression coefficients, which is seen clearly by the fact that κ/λ_1 is the closest to 1.45 in column VIII. This is also illustrated in Figures 4 and 5, which plot, respectively, the model-implied Campbell-Shiller regression coefficients and risk-premium adjusted regression coefficients against the observed Campbell-Shiller regression coefficients. Only the curve labeled VIII comes close to matching the data. (The cases I and XII are excluded in the plot because they are in the wrong direction.) Clearly the model is capable of fitting the Campbell-Shiller regression coefficients closely by adjusting β upward by an appropriate amount.

4 Conclusion

This paper presents an economic explanation for the violation of the expectations hypothesis based on a consumption based asset pricing model with stochastic internal habit. In our model, habit persistence gives rise to mean reversion in real interest rates, and habit shocks give rise to a time-varying risk premium or predictability of expected bond returns. We show how the competition between mean reversion and time-varying risk premium determines the degree in which the expectations hypothesis is violated.

The model explains the expectations puzzle without affecting its ability to explain both the equity premium puzzle and the riskfree rate puzzle. Even when the model is required to explain all three puzzles simultaneously, it is still left with ample degrees of freedom to potentially explain additional features of asset returns and macro-economic aggregates.

Although our model allows the long-run growth rate of the economy to be persistent or mean-reverting, it is calibrated under the assumption that the long-run growth rate economy be constant. An autonomous mean-reverting long-run growth rate amounts to adding an additional factor to the real riskfree rate, but does not affect the market prices of consumption and habit risk. To the extent that the long-run risk is correlated with the consumption or the habit risk, it is priced in the equilibrium and represents an additional source of variation in the expected bond returns. Thus, our model is complimentary to (and in some sense nests) the model of Wachter (2001), who extends Campbell and Cochrane (1999) by introducing a mean-reverting long-run growth rate to the economy, and Bansal and Yaron (2000), who consider the joint effect of the long-run risk and a non-state-separable preference.

We end the paper by raising some outstanding theoretical and empirical questions that are beyond the scope of this paper. First, McCallum (1994) shows that if a monetary policy aimed at interesting rate smoothing induces mean reversion in the nominal interest rates and the risk premium associated with monetary shocks is an AR(1) process, then the expectations hypothesis is violated in the nominal term structure. Furthermore, his model predicts that the degree in which the expectations hypothesis is violated can vary across countries with different type of monetary policies and within a country if monetary regimes changes over time. Such variations are indeed found by Kugler (1997) and Hsu and Kugler (1997). Our

model captures very much the same intuition that the competition between mean reversion and time-varying risk premium lies at the heart of the expectations puzzle. However, in our model, mean reversion of the real interest rates arises from habit persistence rather than monetary policy, and the time-varying risk premium arises from real shocks (habit shocks) rather than nominal shocks. To the extent that the degree of habit persistence and the volatility of the habit shock may vary across different countries and over time, the real effect captured by our model may also help explain cross-country and/or temporal variation in the violation of the expectations hypothesis. Obviously, the nominal and real effects are complimentary to each other and both can be accommodated in a richer model. An interesting empirical question is whether and how the two effects can be disentangled.

Second, for the purpose of explaining the expectations puzzle, our relatively simple continuous-time specification of the habit process is sufficient, and the Markovian structure imposed on the investment opportunity set facilitates the development of the economic intuition. However, in order to address more general asset pricing questions such as the econometric rejection of parametric models with deterministic habit by Euler-equation based tests and the difficulty of such models to explain the Hansen-Jagannathan volatility bounds for the intertemporal marginal rate of substitution, it may very well be necessary to consider more general specifications of the habit and endowment processes. Our discrete-time formulation of a stochastic habit stock (see equations (2) and (3)) are amendable for such extensions. For examples, (i) the distributed lag $B(L)$ may contain arbitrary weights, relaxing the assumption of uniformly positive and geometrically declining weights; this allows the model to accommodate both habit formation and local substitution along the line of Dunn and Singleton (1986), Gallant and Tauchen (1989), Ferson and Constantinides (1991), Eichenbaum and Hansen (1990), Hansen and Jagannathan (1991), and Heaton (1995); (ii) the conditional distribution of y_{t+1} may depend on the entire history of z_t, rather than the current z_t alone; (iii) the habit shock v_{t+1} may have an autocorrelation structure of its own rather than being an i.i.d. shock; and (iv) the Markovian assumption may be relaxed by considering more flexible endowment specifications.

Finally, the model begs a rather intriguing question: where do the habit shocks come from? Do they arise from unobservable taste shocks? Can they be linked to or interpreted as observed macro-economic shocks?

Table 1: Calibrated Parameters

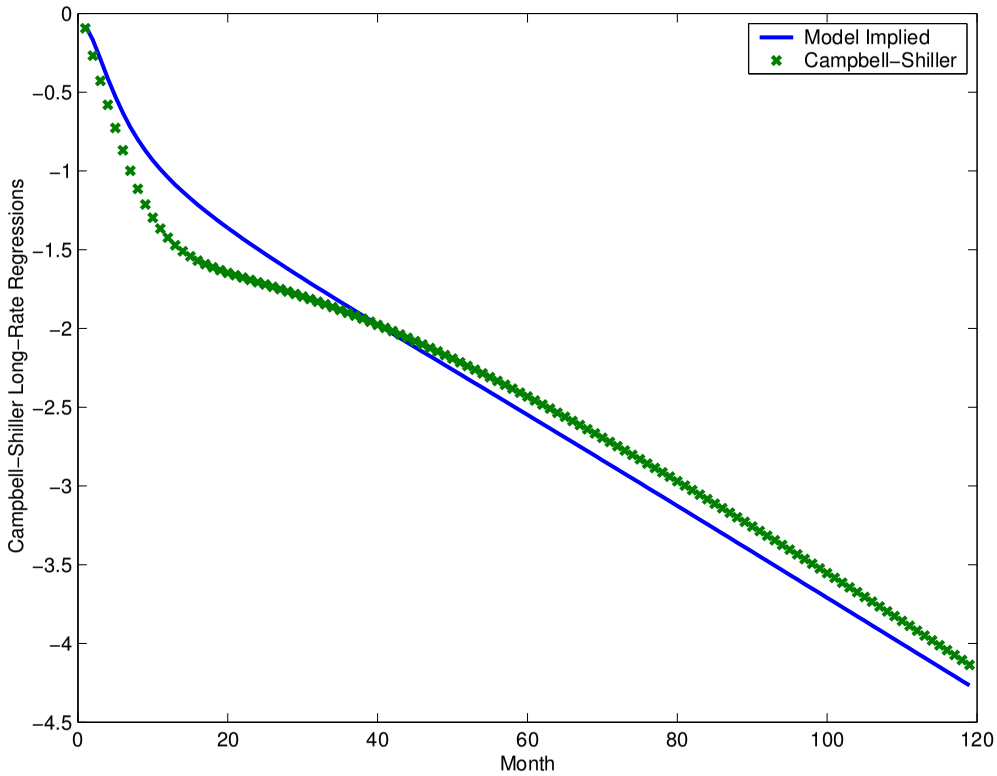
Param.	I	II	III	IV	V	VI	VII	VIII
$1-\gamma$	3.0000	2.0000	4.0000	3.0000	3.0000	3.0000	3.0000	9.0000
κ	0.0232	0.0232	0.0232	0.0200	0.0300	0.0232	0.0232	0.0208
β	0.5000	0.5000	0.5000	0.5000	0.5000	0.1000	1.0000	0.1000
δ	-0.8616	-0.8616	-0.8616	-0.8616	-0.8616	-0.8616	-0.8616	-0.8616
b	0.2009	0.3129	0.1448	0.1729	0.2594	0.2009	0.2009	0.0462
ρ	0.1161	0.1318	0.1083	0.1161	0.1161	0.1161	0.1161	0.0952
$\bar{x}$	0.7450	0.7450	0.7450	1.0540	0.4746	0.7450	0.7450	0.9477
$\bar{z}$	0.0862	0.0553	0.1195	0.1219	0.0549	0.0862	0.0862	0.4277
$r'(\bar{x})$	0.0115	0.0115	0.0115	0.0073	0.0191	0.0115	0.0115	0.0085
$\Lambda'(\bar{x})$	-0.4318	-0.4321	-0.4318	-0.4126	-0.4765	-0.4318	-0.4318	-0.3900
δ_0	-0.0181	-0.0181	-0.0181	-0.0172	-0.0186	-0.0181	-0.0181	-0.0175
δ_1	0.0115	0.0115	0.0115	0.0073	0.0191	0.0115	0.0115	0.0085
κ	0.0232	0.0232	0.0232	0.0200	0.0300	0.0232	0.0232	0.0208
$\bar{x}$	0.7450	0.7450	0.7450	1.0540	0.4746	0.7450	0.7450	0.9477
Σ	0.0431	0.0277	0.0597	0.0609	0.0274	0.0086	0.0862	0.0428
λ_0	-0.0280	-0.0180	-0.0389	-0.0456	-0.0156	-0.0056	-0.0561	-0.0296
λ_1	0.0160	0.0103	0.0222	0.0217	0.0113	0.0032	0.0321	0.0144
κ/λ_1	1.4500	2.2560	1.0450	0.9232	2.6630	7.2480	0.7248	1.4500

In each parameter configuration, $1-\gamma$, κ, and β are fixed. The rest of the structural parameters are calibrated to the following moments: the long-run growth rate $G = 1.89\%$, the riskfree rate $R = 0.94\%$, volatility of consumption growth rate $\sigma_c = 1.5\%$, the equity Sharpe ratio $\bar{\Lambda} = 0.4341$, the real bond Sharpe ratio $\bar{\Lambda}_P = 0.374$. We also impose the following restrictions: the long-run mean of the growth rate for per capita surplus consumption and that of habit stock are 0 and the riskfree rate at zero habit level is 0%.

Table 2: Model-implied Campbell-Shiller Regression Coefficients

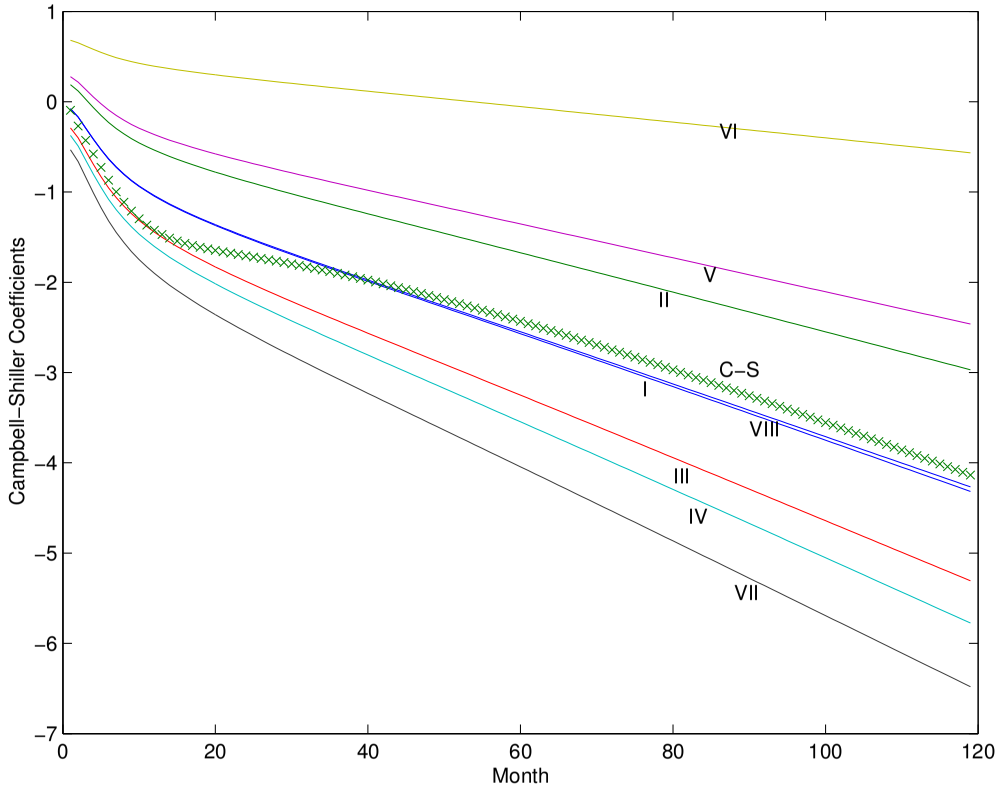
Year	I	II	III	IV	V	VI	VII	VIII	C-S
Original Campbell-Shiller Regressions									
1	-0.9885	-0.4979	-1.3822	-1.5378	-0.3300	0.4095	-1.8279	-0.9945	-1.3664
2	-1.4628	-0.8552	-1.9504	-2.1463	-0.6439	0.2686	-2.5023	-1.4722	-1.6921
3	-1.8299	-1.1318	-2.3901	-2.6186	-0.8851	0.1595	-3.0241	-1.8428	-1.8821
4	-2.1759	-1.3926	-2.8045	-3.0648	-1.1114	0.0567	-3.5156	-2.1926	-2.1239
5	-2.5195	-1.6515	-3.2159	-3.5085	-1.3353	-0.0455	-4.0035	-2.5405	-2.4075
6	-2.8653	-1.9121	-3.6298	-3.9557	-1.5599	-0.1483	-4.4944	-2.8910	-2.7207
7	-3.2137	-2.1748	-4.0468	-4.4070	-1.7855	-0.2520	-4.9886	-3.2447	-3.0550
8	-3.5639	-2.4389	-4.4659	-4.8614	-2.0114	-0.3562	-5.4852	-3.6006	-3.4045
9	-3.9151	-2.7038	-4.8861	-5.3177	-2.2372	-0.4608	-5.9829	-3.9581	-3.7656
10	-4.2666	-2.9690	-5.3065	-5.7752	-2.4624	-0.5656	-6.4807	-4.3164	-4.1364
Risk-premium Adjusted Campbell-Shiller Regressions									
1	0.6220	0.1314	1.0158	1.1713	-0.0364	-0.7759	1.4614	0.6281	-1.3664
2	0.7707	0.1632	1.2583	1.4542	-0.0482	-0.9607	1.8102	0.7801	-1.6921
3	0.9478	0.2498	1.5081	1.7366	0.0030	-1.0416	2.1420	0.9607	-1.8821
4	1.0520	0.2686	1.6806	1.9409	-0.0125	-1.1806	2.3917	1.0687	-2.1239
5	1.1120	0.2440	1.8083	2.1010	-0.0722	-1.3620	2.5960	1.1329	-2.4075
6	1.1445	0.1914	1.9091	2.2350	-0.1608	-1.5724	2.7736	1.1703	-2.7207
7	1.1586	0.1198	1.9918	2.3520	-0.2696	-1.8031	2.9336	1.1896	-3.0550
8	1.1594	0.0344	2.0614	2.4569	-0.3931	-2.0483	3.0807	1.1961	-3.4045
9	1.1496	-0.0617	2.1205	2.5522	-0.5283	-2.3047	3.2174	1.1925	-3.7656
10	1.1303	-0.1673	2.1702	2.6389	-0.6739	-2.5708	3.3444	1.1800	-4.1364

Figure 1: Model-implied and Actual Campbell-Shiller Regression Coefficients



The solid line plots the model-implied Campbell-Shiller regression coefficients, i.e., 1 - $\phi_n^{\mathcal{T}}$, where $\phi_n^{\mathcal{T}}$ is the linear projection coefficient from model-implied expected excess return e_t^n on the term spread $R_t^n - R_t^1$. The model parameters are reported in column "I" in Table 1. The yield data comes from the smoothed Fama-Bliss data set covering January, 1970 to December, 1995 ($\mathcal{T} = 320$ months).

Figure 2: Original Campbell-Shiller Regression Coefficients



Each solid line plots the model-implied Campbell-Shiller regression coefficients, i.e., 1 - $\phi_n^{\mathcal{T}}$, where $\phi_n^{\mathcal{T}}$ is the linear projection coefficient from model-implied expected excess return e_t^n on the term spread $R_t^n - R_t^1$. The label attached to each line indicates the column in Table 1 for the associated model parameters. The line indicated by "C-S" are the actual Campbell-Shiller regression coefficients.

Figure 3: Risk-premium Adjusted Campbell-Shiller Regression Coefficients

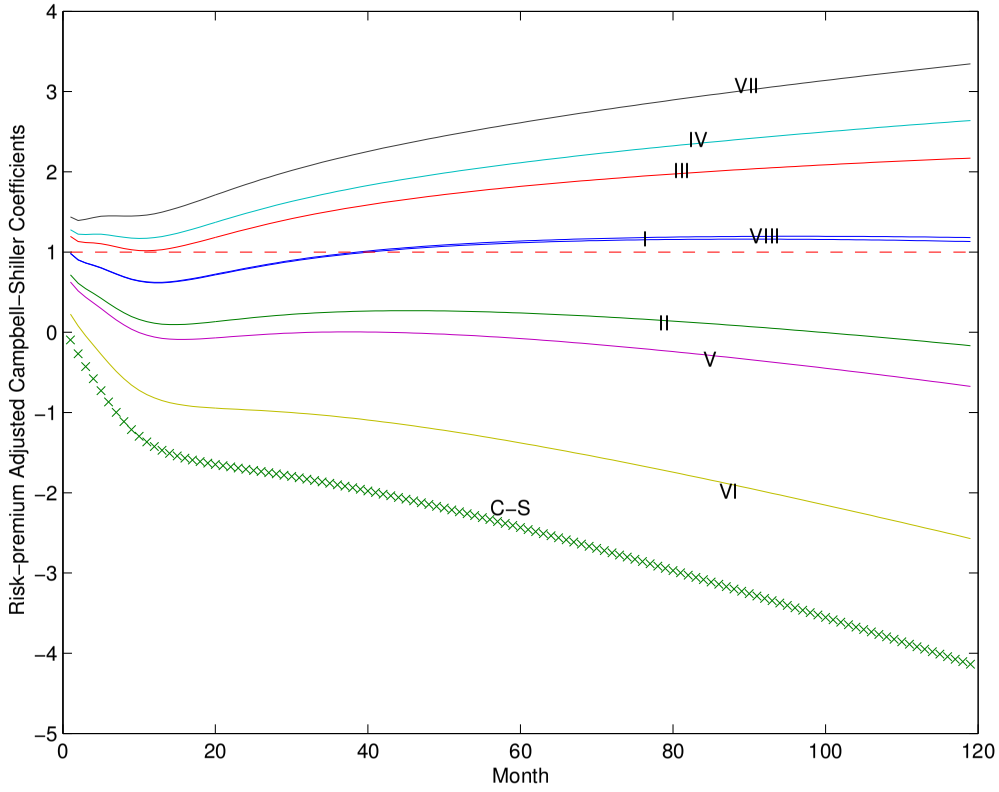
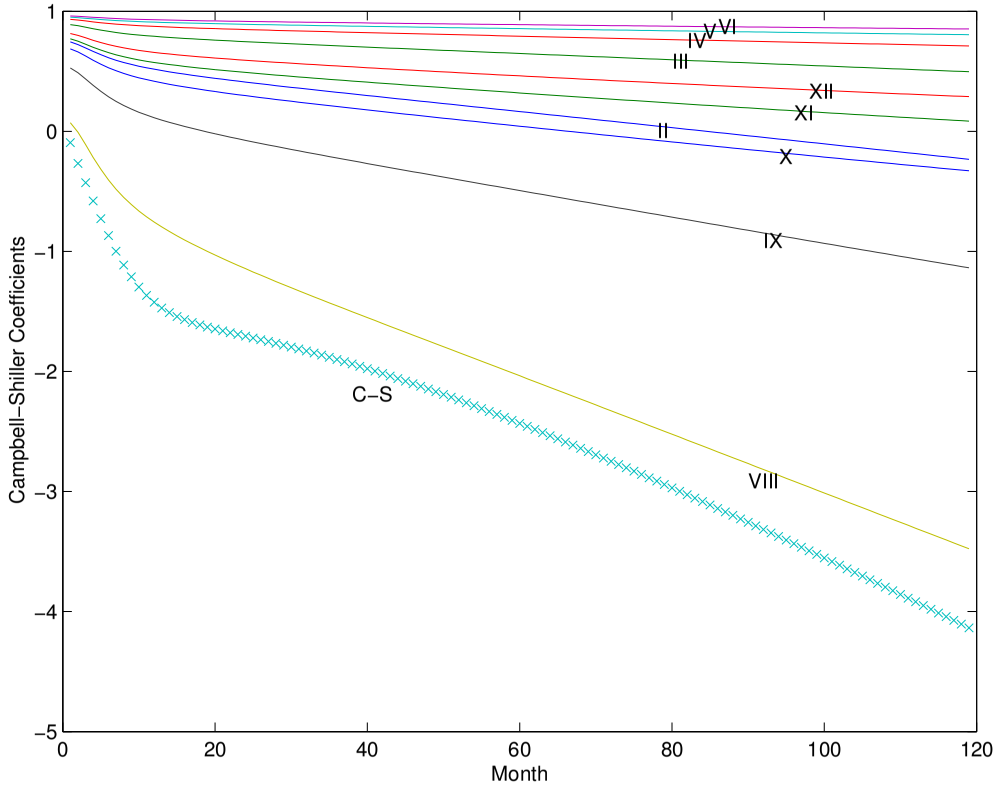


Table 3: Constantinides Model Recalibrated

Param.	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
$1-\gamma$	2.2000	2.2000	2.2000	2.2000	2.2000	2.2000	2.2000	2.2000	2.2000	2.2000	2.2000	2.2000
κ	0.0070	0.0280	0.0500	0.0720	0.0950	0.1080	0.0070	0.0280	0.0500	0.0720	0.0950	0.1080
β	0.1000	0.1000	0.1000	0.1000	0.1000	0.1000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
δ	-0.8615	-0.8615	-0.8615	-0.8615	-0.8615	-0.8615	-0.8615	-0.8615	-0.8615	-0.8615	-0.8615	-0.8615
b	0.0930	0.1720	0.2500	0.3280	0.4050	0.4920	0.0930	0.1720	0.2500	0.3280	0.4050	0.4920
ρ	0.0879	0.0879	0.0879	0.0879	0.0879	0.0879	0.0879	0.0879	0.0879	0.0879	0.0879	0.0879
$\bar{x}$	-0.7170	0.5553	0.2471	0.1597	0.1167	0.1013	-0.7170	0.5553	0.2471	0.1597	0.1167	0.1013
$\bar{z}$	-0.0540	0.0904	0.0494	0.0351	0.0274	0.0222	-0.0540	0.0904	0.0494	0.0351	0.0274	0.0222
$r'(\bar{x})$	0.0181	0.0172	0.0401	0.0624	0.0855	0.0985	0.0181	0.0172	0.0401	0.0624	0.0855	0.0985
$\Lambda'(\bar{x})$	-0.4424	-0.3893	-0.5122	-0.6402	-0.7757	-0.8570	-0.4424	-0.3893	-0.5122	-0.6402	-0.7757	-0.8570
δ_0	-0.0083	-0.0178	-0.0182	-0.0183	-0.0183	-0.0183	-0.0083	-0.0178	-0.0182	-0.0183	-0.0183	-0.0183
δ_1	0.0000	0.0172	0.0401	0.0624	0.0855	0.0985	0.0000	0.0172	0.0401	0.0624	0.0855	0.0985
κ	0.0070	0.0280	0.0500	0.0720	0.0950	0.1080	0.0070	0.0280	0.0500	0.0720	0.0950	0.1080
$\bar{x}$	-0.7170	0.5553	0.2471	0.1597	0.1167	0.1013	-0.7170	0.5553	0.2471	0.1597	0.1167	0.1013
Σ	-0.0054	0.0090	0.0049	0.0035	0.0027	0.0022	-0.0270	0.0452	0.0247	0.0175	0.0137	0.0111
λ_0	0.0002	-0.0045	-0.0021	-0.0014	-0.0011	-0.0009	0.0011	-0.0226	-0.0104	-0.0070	-0.0054	-0.0043
λ_1	-0.0021	0.0030	0.0022	0.0019	0.0018	0.0016	-0.0103	0.0152	0.0109	0.0097	0.0091	0.0082
κ/λ_1	-3.4040	9.2360	22.9300	37.2300	51.9200	65.7700	-0.6807	1.8470	4.5860	7.4460	10.3800	13.1500

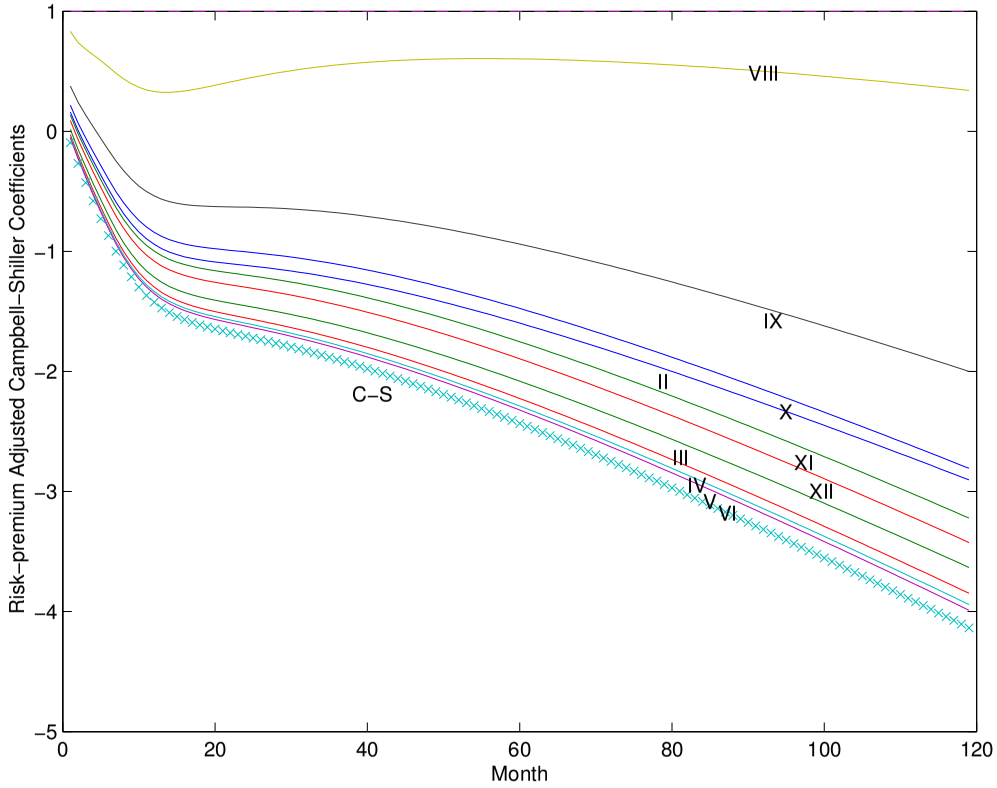
The model is calibrated to the following moments used by Constantinides (1990): the mean consumption growth rate $G = 1.83\%$; the mean Sharpe ratio $\bar{\Lambda} = \frac{0.06}{16.5\%} = 0.3636$, the mean riskfree rate $R = 1\%$. The parameters γ, $a = \kappa + b$, and b are also taken from Constantinides (1990). However, ρ is calibrated here, rather than taken from Constantinides (1990). The value of $\rho = 0.037$ used in Constantinides (1990) would imply that the long-run growth rates of the consumption and the habit stock are different. Two parameters associated habit shocks are new: $\delta = -0.8615$ and $\beta = 0.1$ or 0.5.

Figure 4: Original Campbell-Shiller Regression Coefficients



Each solid line plots the model-implied Campbell-Shiller regression coefficients, i.e., 1 - ϕ_n, where ϕ_n is the linear projection coefficient from model-implied expected excess return e_t^n on the term spread $R_t^n - R_t^1$. The label attached to each line indicates the column in Table 3 for the associated model parameters. The line indicated by "C-S" are the actual Campbell-Shiller regression coefficients.

Figure 5: Risk-premium Adjusted Campbell-Shiller Regression Coefficients



References

Ahn, D.-H., Dittmar, R. F., Gallant, A. R., 2002. The extended saints model of the term structure of interest rates: Theory and evidence. Review of Financial Studies 15, 243–288.

Backus, D., Foresi, S., Mozumdar, A., Wu, L., 2001. Predictable changes in yields and forward rates. Journal of Financial Economics 59(3), 281–311.

Bansal, R., Yaron, A., 2000. Risks for the long run: A potential resolution of asset pricing puzzles. .

Breeden, D., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7, 265–296.

Campbell, J., Cochrane, J., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107(2), 205–251.

Campbell, J. Y., Shiller, R. J., 1991. Yield spreads and interest rate movements: A bird's eye view. Review of Economic Studies 58, 495–514.

Campbell, J. Y., Viceira, L. M., 2001. Who should buy long-term bonds?. American Economic Review 91(1), 99–127.

Constantinides, G., 1990. Habit formation: A resolution of the equity premium puzzle. Journal of Political Economy 98(3), 519–543.

Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. Journal of Finance LV, 1943–1978.

Dai, Q., Singleton, K. J., 2002a. Expectations puzzles, time-varying risk premia, and dynamic models of term structure. forthcoming, Journal of Financial Economics.

——— , 2002b. Term structure dynamics in theory and reality. Review of Financial Studies, forthcoming.

Duffee, G. R., 2002. Term premia and interest rate forecasts in affine models. Journal of Finance 57, 405–443.

Dunn, K. B., Singleton, K. J., 1986. Modeling the term structure of interest rates under non-separable utility and durability of goods. Journal of Financial Economics pp. 27–55.

Eichenbaum, M. S., Hansen, L. P., 1990. Estimating models with intertemporal substitution using aggregate time series data. Journal of Business and Economic Statistics 8, 53–69.

Fama, E., 1984. Term premiums in bond returns. Journal of Financial Economics 13, 529–546.

Fama, E., Bliss, R., 1987. The information in long-maturity forward rates. American Economic Review 77(4), 680–692.

Ferson, W. E., Constantinides, G. M., 1991. Habit formation and durability in aggregate consumption: Empirical tests. Journal of Financial Economics 29, 199–240.

Fleming, M. J., Remolona, E. M., 1999. The term structure of announcement effects. FRB New York Staff Report No. 76.

Gallant, R., Tauchen, G., 1989. Seminonparametric estimation of conditionally constrained heterogeneous processes: Asset pricing applications. Econometrica 57, 1091–1120.

Hansen, L. P., Jagannathan, R., 1991. Implications of security market data for models of dynamic economies. Journal of Political Economy 99(2), 225–262.

Heaton, J., 1995. An empirical investigation of asset pricing with temporically dependent preference specifications. Econometrica.

Hsu, C., Kugler, P., 1997. The revival of the expectations hypothesis of the us term structure of interest rates. Economic Letters pp. 115–120.

Kugler, P., 1997. Central bank policy reaction and expectations hypothesis of the term structure. International Journal of Financial Economics 2, 217–224.

Leippold, M., Wu, L., 2001. Design and estimation of quadratic term structure models. Working paper, Fordham University.

Litterman, R., Scheinkman, J., Weiss, L., 1988. Volatility and the yield curve. Unpublished working paper. Goldman Sachs.

Lu, B., 1999. An empirical analysis of the constantinides model of the term structure. Working paper, University of Michigan.

McCallum, B. T., 1994. Monetary policy and the term structure of interest rates. NBER working paper No. 4938.

Mehra, R., Prescott, E. C., 1985. The equity premium: A puzzle. Journal of Monetary Economics 15, 145–161.

Merton, R. C., 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. Review of Economics and Statistics 51, 247–257.

Piazzesi, M., 2001. An econometric model of the term structure with macroeconomic jump effects. Working Paper, UCLA.

Singleton, K. J., 1993. Econometric implications of consumption-based asset pricing models. in *Advances in Econometrics, Sixth World Congress*, ed. by J. J. Laffont, and C. A. Sims. Cambridge University Press.

Sundaresan, S. M., 1989. Intertemporally dependent preferences and the volatility of consumption and wealth. Review of Financial Studies 2, 73–88.

Wachter, J., 2001. Habit formation and returns on bonds and stocks. Working paper, NYU.

Weil, P., 1989. The equity premium puzzle and the riskfree rate puzzle. Journal of Monetary Economics 24, 401–421.