Habit Formation and Returns on Bonds and Stocks

Jessica A. Wachter*

March 8, 2001

Abstract

This paper proposes a habit formation model that captures the ability of the yield spread to predict excess returns on bonds as documented in empirical studies. The model, a generalization of Campbell and Cochrane (1999), also captures the predictability of stock returns by the price-dividend ratio, a high equity premium, excess volatility, positive excess returns on bonds, and an upward sloping average yield curve. The model is shown to imply a joint process for interest rates and consumption. When this process is estimated from the data, a new empirical fact emerges: Controlling for contemporaneous consumption growth, long lags of consumption predict the interest rate. Thus the success of the model is based on a more realistic process for consumption and the interest rate, rather than additional degrees of freedom in the utility function.

*New York University, Stern School of Business. Email: jwachter@stern.nyu.edu, Phone: (212)998-0799. I thank Geert Bekaert, John Cochrane, Martin Lettau, Anthony Lynch, Andre Perold, Jeremy Stein, Matt Richardson, Stephen Ross, seminar participants at the New York Federal Reserve and especially my advisor John Campbell for helpful comments. I thank Lehman Brothers for financial support.

Introduction

While it has long been known that risk premia on stocks and bonds are time-varying, an explanation for this fact remains elusive. Campbell and Cochrane (1999) propose a consumption-based model in which agents have time-varying risk aversion driven by habit formation. The model of Campbell and Cochrane successfully captures time-variation in equity risk premia (e.g. Campbell and Shiller (1988), Fama and French (1989)). At the same time, the model reconciles the high equity premium with a low riskfree rate.

Campbell and Cochrane (1999) assume that the riskfree rate is constant. Thus their model captures only one piece of the puzzle, the predictability of excess stock returns on the basis of the price-dividend ratio. Risk premia on bonds are also time-varying: Campbell and Shiller (1991) and Fama and Bliss (1987) demonstrate that high yield spreads predict high risk premia on long bonds, a violation of the so-called expectations hypothesis. In fact, Fama and French (1989) argue that the same underlying business-cycle fluctuations, as captured by the dividend-price ratio and yield spread, produce variation in the risk premia on bonds and stocks.

This paper proposes a model that captures predictability in both stock and bond returns driven by the price-dividend ratio and the yield spread. As in the data, the price-dividend ratio predicts excess returns on both bonds and stocks with a negative sign, while the yield spread predicts returns on both bonds and stocks with a positive sign. The amount of predictability is realistic in both cases, with the R^2 on the yield spread regressions lower than that of the dividend-price regressions. The model also captures the high equity premium and the low riskfree rate, an upward sloping yield curve, low volatility of interest rates, and bond premia that increase with the maturity. Bonds have a low beta on the market, and lower expected returns than stocks.

Rather than adding parameters to the utility function, the model generalizes the model of Campbell and Cochrane (1999) to better capture features of the consumption and interest rate data.The distinction is an important one; the results would be less impressive if they occurred by adding degrees of freedom. The proposed model implies a nonlinear joint process for consumption and interest rates. This nonlinear process is shown to be well-approximated by a linear process that can be estimated directly from the data. The first equation in the linear process is

$$r^f_{t+1} \approx \text{constant} + \gamma E[\Delta c_{t+1}] - b \sum_{j=0}^{\infty} \phi^j \Delta c_{t-j} + \epsilon_{t+1}, \tag{1}$$

where ϕ is an autoregressive coefficient. Thus the model nests the traditional relationship between consumption and interest rates that has been estimated by Hansen and Singleton (1983), and many others. The estimation itself reveals a new empirical fact. b turns out to be positive and significant, implying that past consumption growth effects interest rates with a negative sign. This relationship turns out to be a central ingredient in fitting stylized facts about the bond market.

Besides the empirical literature on predictability, this paper draws on the theoretical habit formation literature (e.g., Chapman, 1998, Constantinides, 1990, Dybvig, 1995, and Sundaresan, 1989). Constantinides (1990) and Sundaresan (1989) show that habit formation models can be used to explain a high equity premium with low values of risk aversion. Like these models, the model proposed here assumes the agent evaluates today's consumption relative to a reference point that increases with past consumption. Following Campbell and Cochrane (1999), this paper departs from earlier work by assuming that habit is external to the agent, namely that the agent does not take into account future habit when deciding on today's consumption.[1] Recently, Dai (2000) proposes an extension of the Constantinides model that produces time-varying risk premia on bonds. Dai's approach differs from the one here in that his focus is on the theoretical properties of the model rather than the predictions for consumption and returns.[2]

An intriguing feature of the model in this paper is the link it produces between asset returns and underlying macroeconomic variables. The dividend-price ratio captures past consumption growth, while the yield spread depends both on past consumption growth and on the long-term consumption trend. Besides its success in matching moments of the data, the model makes progress in linking return characteristics with features of the macroeconomy.

This paper is organized into three sections. The first section describes the assump-

[1] Abel (1990) also assumes external habit formation, but in his specification, agents care about the ratio of consumption to habit, rather than the difference. As a result, risk aversion is constant and risk premia do not vary through time

[2] Ferson and Constantinides (1991), Heaton (1995) and Li(2000) also empirically investigate habit formation models, but do not discuss the implications of habit formation for interest rates and returns on bonds

tions on the utility function and the endowment process. The second section describes the solution method, and characteristics of the solution. The solution method involves reducing the problem to one in which there is a single state variable using techniques borrowed from affine bond pricing literature. The third section describes the estimation and results. The final section calibrates the model and describes statistics from simulated data.

1 Aggregate Consumption and Habit

Consumption growth is assumed to follow the bivariate process from Campbell (1999). Log consumption c_t is given by:

$$\begin{aligned} \Delta c_{t+1} &= z_t + v_{t+1} && (2) \\ z_{t+1} &= (1-\psi)g + \psi z_t + u_{t+1}, && (3) \end{aligned}$$

where u_{t+1} and v_{t+1} are jointly normally distributed, $\mathrm{Var}(v_{t+1}) = \sigma_v^2$, $\mathrm{Var}(u_{t+1}) = \sigma_u^2$, and $\mathrm{Corr}(u_{t+1}, v_{t+1}) = \rho$. Related processes are studied by Cecchetti, Lam and Mark (1990) and Kandel and Stambaugh (1991). When $\sigma_u^2 = 0$, the process reduces to i.i.d. consumption growth, which is assumed by Campbell and Cochrane (1999).

Identical agents maximize expected utility relative to a reference point X_t:

$$E\sum_{t=0}^{\infty} \delta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma}. \tag{4}$$

The reference point X_t represents the agent's habit. Define the *surplus consumption ratio*

$$S_t \equiv \frac{C_t - X_t}{C_t}.$$

Following Campbell and Cochrane (1999), X_t is modeled implicitly by defining a process for $s_t = \ln S_t$. This has a double advantage in that it prevents habit from ever falling below consumption, and it provides a stationary state variable for the model. s_t is assumed to follow a first-order autoregressive process:

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t) v_{t+1}. \tag{5}$$

Note that changes in s_t are perfectly correlated with innovations in consumption growth. The conditional volatility of s_t is time-varying: this turns out to be necessary to produce low volatility in the riskfree rate.

The agent's habit is assumed to be *external*, i.e., the agent does not take into account the effect that today's consumption decisions have on X_t in the future. Formally, X_t can be considered as aggregate habit and the agent as evaluating consumption relative to aggregate habit. Because all agents are identical, individual consumption and habit and aggregate consumption and habit can be treated interchangeably.

What does the process (5) for s_t imply about the evolution of habit x_t? Appendix C shows that near the steady state $s = \bar{s}$ and $z_t = g$:

$$x_{t+1} \approx \ln\left(1 - \bar{S}\right) + \frac{g}{1-\phi} + (1-\phi)\sum_{j=0}^{\infty} \phi_t c_{t-j}. \tag{6}$$

Equation (6) shows that x_t approximates an intuitive notion of habit. Namely, x_t is a slowly decaying, weighted average of past consumption. Similarly, in Appendix C, it is shown that

$$s_t - \bar{s} \approx \left(\frac{1}{\bar{S}} - 1\right)\left(-\frac{g}{1-\phi} + \sum_{j=0}^{\infty} \phi^j \Delta c_{t-j}\right).$$

This equation sheds light on the role that s_t plays in the economy. While z_t measures future consumption growth, s_t is a measure of past consumption growth. Because (as shown below) $\bar{S} << 1$, s_t depends positively on recent past consumption growth. Thus s_t is essentially measures recessions (and booms): the higher consumption growth has been in recent years, the further agents' consumption is from habit and the better they feel.

In terms of s_t, the intertemporal marginal rate of substitution equals:

$$M_{t+1} = \delta\left(\frac{S_{t+1}}{S_t}\frac{C_{t+1}}{C_t}\right)^{-\gamma}. \tag{7}$$

The IMRS has a conditional lognormal distribution. Therefore the riskfree rate equals

$$\begin{aligned} r^f_{t+1} &= \ln 1/E_t[M_{t+1}] \\ &= -\ln\delta + \gamma z_t + \gamma(1-\phi)(\bar{s} - s_t) - \frac{\gamma^2\sigma_v^2}{2}(1+\lambda(s_t))^2. \end{aligned}$$

This riskfree rate has some familiar terms from the power utility case and others that are less familiar. First, positive expected consumption growth lead investors to borrow from the future to smooth consumption. This is reflected in the term γz_t. It is important to note that γ is simply a utility curvature parameter, not risk aversion. Like the non-time separable utility functions of Epstein and Zin (1991), the non-state-separable habit formation model drives a wedge between the willingness to substitute intertemporally and aversion to risk. The second term, proportional to $\bar{s} - s_t$ is less familiar. This term implies that as surplus consumption falls relative to its long-term mean, investors want to borrow more. This is due to the mean-reverting nature of surplus consumption: investors borrow against future periods when habit has had time to adjust and surplus consumption is higher. The last term represents precautionary savings. A higher $\lambda(s_t)$ implies that surplus consumption, and therefore marginal utility, is more volatile. Investors increase saving, and r^f falls. The sign of s_t in the riskfree rate depends on which effect wins out.

As in the model of Campbell and Cochrane (1999), three considerations determine the parameterization of $\lambda(s_t)$. First it is required that habit be predetermined at the steady state:

$$\left.\frac{dx}{dc}\right|_{s_t=\bar{s}} = 0. \tag{8}$$

Second, habit must be predetermined in a neighborhood of the steady-state, or equivalently, that habit move nonnegatively with consumption everywhere:

$$\left.\frac{d}{ds}\left(\frac{dx}{dc}\right)\right|_{s_t=\bar{s}} = 0. \tag{9}$$

Finally, the riskfree rate is required to be linear in s_t.[3] It is automatically linear in z_t because λ is a function of s_t alone. As shown in Section 3, this requirement implies a joint system of consumption and interest rates that can be estimated from the data.

The last requirement imposes the condition that $\lambda(s_t)$ be a quadratic function. Conditions (8) and (9) then determine the two constants in the square root. The resulting function equals

$$\lambda(s_t) = \begin{cases} (1/\bar{S})\sqrt{1-2(s_t-\bar{s})} - 1, & s_t \leq s_{\max} \\ 0 & s_t \geq s_{\max} \end{cases} \tag{10}$$

[3]Thus the model nests the single-factor model for interest rates considered in the working paper Campbell and Cochrane (1995).

and

$$s_{\max} \equiv \bar{s} + \frac{1}{2}\left(1 - \bar{S}^2\right) \tag{11}$$

$$\bar{S} \equiv \sigma_v \sqrt{\frac{\gamma}{1 - \phi - b/\gamma}} \tag{12}$$

The constant $s_{\max}$ is determined by $\lambda(s_{\max}) = 0$. In the continuous-time limit, s_t would never venture above $s_{\max}$. In discrete time it happens sufficiently rarely that it does not affect the behavior of the model. Substituting into the equation for the riskfree rate produces

$$r^f_{t+1} = -\ln\delta + \gamma z_t - b(s_t - \bar{s}) - \left(\frac{\gamma}{\bar{S}}\right)\frac{\sigma_v^2}{2} \tag{13}$$

$$= -\ln\delta + \gamma z_t - b(s_t - \bar{s}) - \frac{\gamma}{2}\left(1 - \phi - \frac{b}{\gamma}\right). \tag{14}$$

It is straightforward to check that (8) and (9) are satisfied.

It follows from (10) that $\lambda(s_t)$ is a decreasing function of s_t, namely high values of s_t decreases the variance. Thus the precautionary savings term is decreasing in s_t. As shown in (13), the net effect of intertemporal substitution and precautionary savings is summarized in the parameter b. When $b > 0$, higher values of s_t decrease the riskfree rate, and intertemporal substitution dominates. When $b < 0$, precautionary savings dominates.

Based on (10), it is already possible to see how this model might be successful at producing time-varying risk premia. It follows from the investor's Euler equation

$$E_t[M_{t+1}R_{t+1}] = 1. \tag{15}$$

that

$$\frac{E_t(R^e_{t+1})}{\sigma_t(R^e_{t+1})} = -\rho_t(M_{t+1}, R^e_{t+1})\frac{\sigma_t(M_{t+1})}{E_t(M_{t+1})}. \tag{16}$$

where R^e denotes the return in excess of the riskfree asset. Therefore

$$\max\frac{E_t(R^e_{t+1})}{\sigma_t(R^e_{t+1})} = \left(e^{\gamma^2\sigma_v^2(1+\lambda(s_t))^2} - 1\right)^{\frac{1}{2}}$$

$$\approx \gamma\sigma_v(1 + \lambda(s_t)). \tag{17}$$

It follows from (10) that the maximum Sharpe ratio is decreasing in s_t. Following the interpretation of s_t as a recession variable, investors demand a higher return for unit risk in recessions rather than in booms.[4] This feature allows the model to match features of the aggregate stock market. In addition, the maximum Sharpe ratio depends only on s_t, while the riskfree rate depends on both s_t and z_t.

Equation (16) shows that risk premia on bonds also vary according to s_t. However, these equations alone do not shed light on the sign of the risk premia on bonds. This depends on the underlying relationship between consumption and the interest rate, which is to be determined by the data.

2 Model Solution

The underlying assets in this economy are claims to future consumption. The notation P_{nt} is used to denote the price of the security that pays out aggregate consumption n periods from now. P_{nt} is characterized by the equations

$$P_{0t} = C_t \tag{18}$$

$$P_{nt} = E_t[M_{t+1}P_{n-1,t+1}], \tag{19}$$

where (19) follows from the representative investor's first order condition (15) and the fact that P_{nt}, by definition, pays no dividends. The total wealth of the economy is the claim to all future consumption. Namely,

$$P_t = \sum_{n=1}^{\infty} P_{nt} \tag{20}$$

In what follows, P_t is taken to be the aggregate stock market. It is also possible to consider levered claims to consumption, namely, assets that pay out C_{t+n}^{θ}. In this case, the equations are

$$P_{0t} = C_t^{\theta} \tag{21}$$

$$P_{nt} = E_t[M_{t+1}P_{n-1,t+1}] \tag{22}$$

The advantage of the approach is when $\theta = 0$, P_n denotes the price of a real bond, while when $\theta = 1$, P_n gives the value of equity.

[4]Chou, Engle and Kane (1992) and Harvey (1989) present evidence that the Sharpe ratio, as well as the risk premium, varies through time.

For $\theta = 0$, equations (21) and (22) form the basis for calculation of bond prices in discrete-time affine models (e.g. Backus, Foresi, and Telmer, 2000, Singleton, 1990, Sun, 1992). Here, their use is expanded to price both bonds and equities in a non-affine model.

2.1 Prices as Functions of the State Variables

Under habit formation $M_{t+1} = \delta \left(\frac{S_{t+1}}{S_t} \frac{C_{t+1}}{C_t} \right)^{-\gamma}$. It turns out to be easier to solve for the scaled-price measures because, unlike prices, they are stationary and are purely functions of the state variables of the economy. For the habit formation model, prices are determined by

$$\frac{P_{0t}}{C_t^\theta} = 1 \tag{23}$$

$$\frac{P_{nt}}{C_t^\theta} = E_t \left[\delta \left(\frac{S_{t+1}}{S_t} \right)^{-\gamma} \left(\frac{C_{t+1}}{C_t} \right)^{\theta-\gamma} \frac{P_{n-1,t+1}}{C_t^\theta} \right]. \tag{24}$$

Because M_{t+1} is conditionally lognormal,

$$\frac{P_{1t}}{C_t^\theta} = \delta G^{\theta-\gamma} \exp \left\{ (\theta - \gamma)(z_t - g) - \gamma(1-\phi)(\bar{s} - s_t) + [(\theta - \gamma) - \gamma\lambda(s_t)]^2 \frac{\sigma_v^2}{2} \right\}$$

P_{1t}/C_t^θ is loglinear in z_t. However, $\ln \left(P_{1t}/C_t^\theta\right)$ is nonlinear in s_t so expectation (24) cannot be calculated in closed form when $n > 1$.[5] Therefore prices must be solved for numerically.

One possible approach to the calculating prices is the use of Monte Carlo simulation to solve the expectation in (24). The expectation can be calculated directly by simulating sample paths for S and C. This method works poorly for the habit formation model, where there is a lower tail of bad events. Because the probabilities of rare events are so important, Monte Carlo estimation converges very slowly. Another possible approach is to solve (24) through numerical integration on s_t and z_t. While more accurate, this method suffers from the "curse of dimensionality". Each step in the recursion requires a two-dimensional integration and interpolation on a two-dimensional grid.

[5] When $\theta = 0$, the right hand side gives the price of a one-period bond which, by design, does not depend on s_t. However, when $n = 2$, the expression depends on s_t in a non-linear way.

Fortunately, by combining the analytical methods of affine bond pricing with numerical integration, the curse of dimensionality can be avoided and prices can be computed with no more effort than that required in the model of Campbell and Cochrane (1999). The idea is to express price-dividend ratios as the product of two functions: one of z_t and one of s_t. The z_t term can be determined in closed form because the model is affine in z_t. The s_t term is determined by a one-dimensional recursion.

Following the pattern in affine bond pricing, a form for the solution is first "guessed", and then verified by substituting back into the recursion. Based on the formula for P_{1t}/C_t^θ a reasonable guess for the form of P_{nt}/C_t^θ is the following:

$$\frac{P_{nt}}{C_t^\theta} = \exp\left\{A(n)(z_t - g) + B(n)\right\} F(s_t, n), \tag{25}$$

Substituting into (24) and solving verifies that (25) is correct and that A and B take the following forms:

$$\begin{aligned} A(n) &= (\theta - \gamma)\frac{1 - \psi^n}{1 - \psi} \\ B(n) &= n \ln \delta + n(\theta - \gamma)g + \frac{(\theta - \gamma)^2}{2}(1 - \rho^2)\sigma_u^2 \sum_{k=1}^{n-1} \left(\frac{1 - \psi^k}{1 - \psi}\right)^2. \end{aligned}$$

The function F is determined by the recursion

$$F(s_t, n) = E_t\left[\exp\left\{(\theta - \gamma)\left[\frac{\sigma_u}{\sigma_v}\rho\frac{1 - \psi^{n-1}}{1 - \psi} + 1\right] v_{t+1}\right\} \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} F(s_{t+1}, n - 1)\right], \tag{26}$$

with boundary condition

$$F(s_t, 0) = 1 \tag{27}$$

Details can be found in the Appendix.

While the formulas look complicated, they can be used to gain insight into the economic behavior of the model. One-period returns on these zero-coupon assets are given by

$$R_{n,t+1} = \left(\frac{P_{n-1,t+1}}{C_{t+1}^\theta} / \frac{P_{nt}}{C_t^\theta}\right)\left(\frac{C_{t+1}}{C_t}\right)^\theta \tag{28}$$

Therefore, by (25),

$$R_{n,t+1} = \exp\left\{A(n-1)(z_{t+1} - g) - A(n)(z_t - g) + B(n-1) - B(n)\right\}\left(\frac{C_{t+1}}{C_t}\right)^\theta \frac{F(s_{t+1}, n-1)}{F(s_t, n)}$$

It is straightforward to show that

$$E_t[r_{n,t+1} - r^f_{t+1}] = E_t[\ln F(s_{t+1}, n-1) - \ln F(s_t, n)] + \text{constant}. \qquad (29)$$

Therefore risk premia on the zero-coupon assets are time-varying, and depend only on s_t. Thus the model conveniently separates the sources of variation in prices: z_t (the long-run consumption trend) controls changes in the interest rate and in expected consumption growth while s_t (the business-cycle variable) controls changes in risk premia as well as in interest rates. It is important to note that (29) applies to premia on long-term bonds as well as on the market. Thus bond premia will vary due to the same cyclical factor s_t.

Yields on zero-coupon bonds can also be calculated up to the recursively defined function $F(s_t, n)$. Yields are defined as

$$y_{nt} = -\frac{1}{n}\ln P_{nt}.$$

From (25), it follows that

$$y_{nt} = -\ln\delta + \gamma g - \frac{(\theta-\gamma)^2}{2}(1-\rho^2)\sigma_u^2 \sum_{k=1}^{n-1}\left(\frac{1-\psi^k}{1-\psi}\right)^2 + \frac{\gamma}{n}\frac{1-\psi^n}{1-\psi}(z_t - g) - \frac{1}{n}\ln F(s_t, n). \qquad (30)$$

A special case of the model occurs when $\rho = 0$, $b = 0$, and θ is set equal to zero to price bonds. The recursion (26) collapses and $F(s_t, n) = 1$ for all n. This is because the term multiplying F in the expectation (26) no longer depends on n, when ρ equals zero. Because the one-period (riskless) bond does not depend on s_t, the term must equal a constant independent of n. From the boundary condition (18), the constant must equal 1. From (29), excess returns on bonds equal zero. This is not surprising: when $\rho = 0$ and $b = 0$, interest rates and consumption growth are conditionally uncorrelated. From (16), excess returns on bonds must equal zero because they are riskless from the point of view of investors.

The $\rho = 0$, $b = 0$ case, while relatively uninteresting in itself, does lend insight into the workings of the model. $b = 0$ and $\rho = 0$ represents a hairline case. For $b > 0$, interest rates are negatively correlated with consumption, hence bond prices are positively correlated with consumption. Therefore $b > 0$ implies bonds have positive risk premia, while $b < 0$ implies negative risk premia. Similarly, when $\rho < 0$

(> 0), interest rates are negative (positively) correlated with consumption and bonds have positive (negative) risk premia. Thus the sign of the risk premia on bonds depends on the sign of b and ρ, and the relative strength of these two effects.

2.2 Power Utility: A Comparison

While the difficulties of modeling equity prices under the assumption of power utility are well-known, the term-structure implications are less well understood. In order to clarify which new results arise from the habit utility assumption and which could be derived using power utility, this section derives formulas for prices under power utility.[6] This is especially important in Section 4.5, which considers why bonds have risk premia.

Under power utility, the intertemporal marginal rate of substitution is given by:

$$M_{t+1} = \hat{\delta}\left(\frac{C_{t+1}}{C_t}\right)^{-\hat{\gamma}}. \tag{31}$$

Thus the riskless rate must be

$$r^f_{t+1} = -\ln\hat{\delta} + \hat{\gamma} z_t - \frac{\hat{\gamma}^2}{2}\sigma_v^2. \tag{32}$$

From (16), the maximal Sharpe ratio in the economy is given by

$$\begin{aligned} \max\frac{E_t(R^e_{t+1})}{\sigma_t(R^e_{t+1})} &= \left(e^{\hat{\gamma}^2\sigma_v^2} - 1\right)^{\frac{1}{2}} \\ &\approx \hat{\gamma}\sigma_v. \end{aligned} \tag{33}$$

Equations (32) and (33) point to the failures of the power-utility model. In the postwar quarterly data-set used in this paper, the Sharpe ratio is about 0.5. However, the variance of consumption growth is very low - about 1% a year. Therefore, $\hat{\gamma}$ must be about 50 for (33) to hold. This is the well-known equity premium puzzle of Mehra and Prescott (1985). Supposing this number were plausible, (32) creates another set of problems. In the data, the riskfree rate has a low mean. But if $\hat{\gamma}$ is high, then $\hat{\delta}$ must be greater than 1 to adjust for the consumption growth term. Moreover, given

[6] Bekaert and Grenadier (1999) and Campbell (1986) consider implications of power utility in similar but not identical models.

the values of σ_u considered in this paper, high values of $\hat{\gamma}$ imply a riskfree rate that is far more volatile than that implied by the data.

A similar method to that used in the previous section can be used to calculate prices under power utility. It is shown in the Appendix that

$$\frac{P_{nt}}{C_t^\theta} = \exp\left\{\hat{A}(n)(z_t - g) + \hat{B}(n)\right\} \tag{34}$$

where

$$\hat{A}(n) = (\theta - \hat{\gamma})\frac{1-\psi^n}{1-\psi} \tag{35}$$

$$\begin{aligned}\hat{B}(n) &= n\ln\hat{\delta} + (\theta-\hat{\gamma})ng + \\ &\frac{(\theta-\hat{\gamma})^2}{2}\left[\sigma_u^2\sum_{k=1}^{n-1}\left(\frac{1-\psi^k}{1-\psi}\right)^2 + 2\rho\sigma_u\sigma_v\sum_{k=1}^{n-1}\frac{1-\psi^k}{1-\psi} + n\sigma_v^2\right].\end{aligned} \tag{36}$$

It follows from (34) that returns are lognormal. Taking the expectation and rearranging yields

$$\ln E_t[R_{n,t+1}] = \ln\hat{\delta} + \hat{\gamma}z_t + \frac{(\theta-\hat{\gamma})^2}{2}\left[2(\theta-1)\rho\sigma_u\sigma_v\frac{1-\psi^n}{1-\psi} + (\theta^2-1)\sigma_v^2\right] \tag{37}$$

Note that the coefficient on z_t is $\hat{\gamma}$, the same as the coefficient in the equation for the riskfree rate. Therefore the risk premium,

$$E_t[r_{n,t+1} - r_{t+1}^f] = \text{constant}.$$

It is not possible for the power utility model to produce predictability in excess returns for any values of the parameters.[7]

Under power utility, the term structure is the discrete-time equivalent of Vasicek (1977). The yield on zero-coupon bonds is given by

$$\begin{aligned}y_{nt} &= -\ln\hat{\delta} + \hat{\gamma}g + \frac{\hat{\gamma}}{n}\left(\frac{1-\psi^n}{1-\psi}\right)(z_t - g) \\ &\quad - \frac{\hat{\gamma}^2}{2}\left[\frac{\sigma_u^2}{n}\sum_{k=1}^{n-1}\left(\frac{1-\psi^k}{1-\psi}\right)^2 + 2\frac{\rho\sigma_u\sigma_v}{n}\sum_{k=1}^{n-1}\frac{1-\psi^k}{1-\psi} + \sigma_v^2\right].\end{aligned}$$

[7] Campbell (1999) derives this result in an approximate log-linear framework.

The term in brackets determines the shape of the expected yield curve. This term arises from correcting for Jensen's inequality. When $\rho \geq 0$, yields are decreasing in the maturity n.[8] Even when $\rho < 0$, for the parameter values of interest, the middle term is dominated by the first term. Thus the power utility model produces a downward sloping expected yield curve.

3 Estimating the Model

The dependence of the interest rate on s_t and z_t is estimated directly from the data. Section 3.1 shows that the model implies a joint process for consumption and interest rates. Section 3.2 describes the data, and Section 3.3 estimates the equations from 3.1 using instrumental variables, revealing a new empirical fact: interest rates depend significantly on long lags of past consumption. In Section 3.4, the nonlinear restrictions imposed on the coefficients are taken into account using Generalized Method of Moments estimation. The result provides confirmation of the results in Section 3.3 and estimates for all model parameters.

3.1 Deriving a Joint Process for Interest Rates and Consumption

As shown in Section 1, the riskfree rate between t and $t+1$ equals

$$\begin{aligned} r^f_{t+1} &= -\ln\delta + \gamma z_t - b(s_t - \bar{s}) - \frac{\gamma}{2}\left(1 - \phi - \frac{b}{\gamma}\right) \\ &= \bar{r}^f + \gamma(z_t - g) - b(s_t - \bar{s}). \end{aligned}$$

Here, and in what follows, $\bar{r}^f$ is taken to be the unconditional mean of the riskfree rate. Rearranging, it follows that

$$z_t = g + \frac{1}{\gamma}\left(r^f_{t+1} - \bar{r}^f\right) + \frac{b}{\gamma}\left(s_t - \bar{s}\right). \tag{38}$$

[8] The quantity $(1-\psi^k)/(1-\psi)$ is increasing in k. Therefore the average of this expression (and the average of its square) where k ranges from 1 to n is increasing in n.

Using equations (2) and (3), this equation can be used to derive a relation between consumption growth, interest rates, and habit. It follows from (2) that

$$\Delta c_{t+1} = g + \frac{1}{\gamma}\left(r^f_{t+1} - \bar{r}^f\right) + \frac{b}{\gamma}(s_t - \bar{s}) + v_{t+1}, \tag{39}$$

or equivalently

$$r^f_{t+1} - \bar{r}^f = -\gamma g + \gamma \Delta c_{t+1} - b(s_t - \bar{s}) - \gamma v_{t+1}. \tag{40}$$

Equation (38) also implies an equation for the evolution of the riskfree rate. Using the process for z_t

$$z_{t+1} - g = \psi(z_t - g) + u_{t+1}$$

from (3), it follows that

$$r^f_{t+2} - \bar{r}^f + b(s_{t+1} - \bar{s}) = \psi\left[r^f_{t+1} - \bar{r}^f + b(s_t - \bar{s})\right] + \gamma u_{t+1},$$

which implies

$$\begin{aligned} r^f_{t+2} &= \bar{r}^f + \psi\left(r^f_{t+1} - \bar{r}^f\right) - b(s_{t+1} - \bar{s}) + \psi b(s_t - \bar{s}) + \gamma u_{t+1} \\ &\approx \bar{r}^f + \psi\left(r^f_{t+1} - \bar{r}^f\right) + b(\psi - \phi)(s_t - \bar{s}) - b\lambda(\bar{s})v_{t+1} + \gamma u_{t+1}. \end{aligned} \tag{41}$$

The second line follows from (54).

Equations (39) and (40) nest equations long of interest in empirical asset pricing. When $b = 0$, these equations reduce to the equations considered by Hansen and Singleton (1983). This, and many subsequent studies have failed to find a significant relationship between consumption growth and interest rates. The habit formation model of this paper is unique in suggesting a testable modification to the traditional model. Equations (39) and (40) add surplus consumption as a right hand side variable. Using an empirical proxy, these equations can be estimated using instrumental variables, as is done in the following section.

The first step in the estimation is constructing an empirical proxy for surplus consumption s_t. Appendix C demonstrates that

$$s_t - \bar{s} \approx \lambda(\bar{s})\left(\sum_{j=0}^{\infty} \phi^j \Delta c_{t-j} - \frac{g}{1-\phi}\right),$$

near the steady state $\bar{s} = s$. Define

$$B \equiv b\lambda(\bar{s}) = b\left(\frac{1}{\bar{S}} - 1\right). \tag{42}$$

Equations (39) and (41) imply that

$$\Delta c_{t+1} \approx g + \frac{1}{\gamma}\left(r^f_{t+1} - \bar{r}^f\right) + \frac{B}{\gamma}\left(\sum_{j=0}^{\infty}\phi^j \Delta c_{t-j} - \frac{g}{1-\phi}\right) + v_{t+1} \tag{43}$$

$$r^f_{t+2} \approx \bar{r}^f + \psi\left(r^f_{t+1} - \bar{r}^f\right) + (\psi - \phi)B\left(\sum_{j=0}^{\infty}\phi^j \Delta c_{t-j} - \frac{g}{1-\phi}\right) - Bv_{t+1} + \gamma u_{t+1}. \tag{44}$$

Equation (40) can be rewritten in the same way:

$$r^f_{t+1} - \bar{r}^f = -\gamma g + \gamma \Delta c_{t+1} - B\left(\sum_{j=0}^{\infty}\phi^j \Delta c_{t-j} - \frac{g}{1-\phi}\right) - \gamma v_{t+1}. \tag{45}$$

In practice, of course, only a finite number of lags of consumption can be used. In the estimation below, the proxy is constructed for 40 lags (10 years) of consumption growth.

Because these equations is used to specify the model, the accuracy of the approximation must be checked. It is not necessary that the approximations be accurate in every respect. All that is necessary is that the true and the approximate model behave the same as far as estimating the equations above. This can be easily done by rerunning the tests on data simulated from the true model. Performing this exercise reveals that, for the purposes of the equations above, the approximation is extremely accurate.

The strategy is to first estimate (43) and (44) without any restrictions on the parameters. The analogue of (40) is also estimated as a robustness check. Finally, the equations are estimated with the restrictions implied by the model using GMM, leading to a full specification of the model.

3.2 Data

Quarterly, post-war consumption and financial data are used to estimate the model. Following standard practice, per-capita consumption of nondurables and services (from Basic Economics) proxies for aggregate consumption. The financial data are from CRSP. The return on the 90-day Treasury bill proxies for the riskfree rate and the value-weighted return on the NYSE and AMEX proxies for the market return. The price-dividend ratio is constructed by summing the current and last three quarters of dividends and dividing by the current price. All variables are adjusted for inflation.

As discussed above, a long-run weighted average of consumption growth, $\sum \phi^j \Delta c_{t-j}$, proxies for habit. In order to capture the long-run nature of s_t, $\sum \phi^j \Delta c_{t-j}$ is estimated out to 40 lags, or 10 years of previous consumption growth. This has the effect of reducing the length of the sample by 10 years, so that 1957 is the starting year, rather than 1947. The last column of Table 4 presents summary statistics for this sample period.

An important summary statistic is the first-order autocorrelation of the price-dividend ratio. This turns out to equal ϕ, the persistence of the surplus consumption ratio. In quarterly terms, this autocorrelation equals 0.969 (Table 4 reports annualized values). The analysis that follows takes this parameter as given.

3.3 Unrestricted Estimation

First, (43) is estimated. The purpose of the unrestricted estimation is to directly test the role of past consumption in the riskfree rate, and to shed light on the GMM estimation below. In theory, the riskfree rate is known, but in the data, of course, the ex post riskfree rate is uncertain. Therefore these equations must be estimated using the instrumental variables approach of Hansen and Singleton (1983). Following previous studies, the instruments are taken to be twice-lagged values of consumption growth, the riskfree return, and the dividend-price ratio. These are valid instruments, because, as time $t-1$ variables, they are uncorrelated with time $t+1$ errors.

The results of the estimation are given in Panel A of Table 1. The coefficient on the interest rate is similar in magnitude to that found in previous studies. More importantly, it is not signficantly different from zero. However, the coefficient on the long-run moving-average of consumption lags is significant at the 0.05 level. This surprising result supports the notion that the riskfree rate depends on past consumption.

Panel B reports the results of interchanging the roles of consumption and interest rates in (43). Under the orthogonality conditions, the estimates of this equation and the other are asymptotically equivalent, but might differ in finite samples. The coefficient on lagged consumption growth is signficant at the 0.01 level. Interestingly, the coefficient on contemporaneous consumption growth is significant at the .1 level, which differs from previous studies which fail to find significance. This equation shows directly that long lags of consumption growth predict the interest rate.

Figure 1 plots the history of average past consumption growth ($\sum_{t=0}^{40} \phi^j \Delta c_{t-j}$), the log price-dividend ratio, and the fitted value of the interest rate from the first-stage regression. The negative relationship betwen consumption and interest rates as implied by (43) and (45) is apparent throughout the sample period. Moreover, past consumption has a surprising ability to explain both short term and long term fluctuations in the real interest rate. Figure 1 also reveals that changes in average past consumption growth mirror changes in the price-dividend ratio. This striking fact provides direct support for the assumptions behind habit utility.

3.4 Generalized Method of Moment Estimation

In this section, the parameters of the model are estimated using the Generalized Method of Moments. The reason to use this approach is that (43) and (44) are overidentified: the coefficient on lagged consumption in the second equation is a function of the rest of the parameters. Moreover, B is itself a nonlinear function of the parameters. In addition, the approach allows standard errors to be obtained for σ_v, σ_u, and ρ. The system to be estimated is (43) and (44) together with three equations that determine σ_v, σ_u, and ρ:

$$G\left(g, \gamma, b, \bar{r}^f, \psi, \sigma_v, \sigma_u, \rho\right) = \begin{pmatrix} \frac{1}{T} H' \epsilon_1 \\ \frac{1}{T} H' \epsilon_2 \\ \frac{1}{T} \epsilon_1' \epsilon_1 - \sigma_v^2 \\ \frac{1}{T} \epsilon_2' \epsilon_2 - \left(B^2 \sigma_v^2 + \gamma^2 \sigma_u^2 - 2B\gamma\rho\sigma_u\sigma_v\right) \\ \frac{1}{T} \epsilon_1' \epsilon_2 - \left(-B\sigma_v^2 + \gamma\rho\sigma_u\sigma_v\right) \end{pmatrix}. \tag{46}$$

where T is the sample length (in this case, about 165), H is the $T \times 3$ matrix of instruments and ϵ_1 and ϵ_2 are the errors from (43) and (44). B is defined in terms of the primitive parameters by (42) and (12).

While (43) and (45) are asymptotically equivalent under the orthogonality conditions, there is reason to believe that (43) has better finite-sample properties. This is because consumption growth is only weakly predictable by the instruments compared to the interest rate. As shown in Table 1, the R^2 for the first-stage regression is much lower for consumption than for interest rates. Therefore, for the GMM analysis below, interest rates are chosen as the right-hand-side variable. A consequence of this choice is that b/γ is estimated, rather than b alone. Because γ is estimated with noise, estimating b results in large standard errors. This is not a problem: from (14), and (12), either b or b/γ could be interpreted as the more primitive parameter.[9]

The results of the GMM estimation are reported in Table 2. The ratio b/γ is found to be statistically different from zero, confirming the earlier results. The persistence coefficient ψ is found to be near 1, similar to previous studies wich fail to reject a unit root for the interest rate. As before, γ is not statistically different from zero. The point estimate and the standard errors are nearly identical to those in the unrestricted problem.

This estimation implies that the riskfree rate in the model should depend on surplus consumption with a negative sign. Economically, this means that the intertemporal substitution effect of habit dominates the precautionary savings effect. When surplus consumption is low, investors borrow to allow consumption to catch up to habit.

4 Model Evaluation

The model is solved numerically using the results of Section 2, and the estimation in the previous section. Section 4.1 discusses parameter choices, Section 4.2 plots the price-dividend ratio and the yield spread as a function of the variables s_t and z_t, and the remaining sections discuss the implications of the model for unconditional moments and return predictability, using 100,000 quarters of simulated data.

[9] An alternative would be to estimate (45) instead of (39), and suggested by Panel B in Table 1, this results in highly significant estimate for b. On the other hand, this method results in noisier estimates of g and σ_v.

4.1 Parameter Choices

The starting point in calibrating the model are the estimates in Table 2. Simply using the point estimates would be a mistake, however. For example, the curvature parameter γ is measured with substantial noise. Moreover, γ controls the level of the Sharpe ratio in the model, and using the point estimate would result in a Sharpe ratio that is unrealistically high. As in Campbell and Cochrane (1999), γ is set in order to match the Sharpe ratio in the 1957-1998 sample exactly. This implies a value for γ that is lower than, but well within one standard error of the point estimate. Setting γ in this manner has the additional advantage that it standardizes the effects of considering a range of choices for the other parameters.

To determine the range of results implied by the model, several sets of parameter values are considered. The quantitative behavior of bond prices in the model turns out to be sensitive to the choice of b and of σ_u. Increasing ρ or ψ results in similar effects as increasing σ_u. To limit the number of variable combinations, only b and σ_u are varied. Four sets of parameter values are examined. In the first, b/γ and σ_u equal their estimates from the data. In the second, b/γ equals its estimate from the data, and σ_u is reduced by 1 standard deviation. In the third, b/γ is reduced by one standard deviation, and σ_u is at its point estimate. Finally, both b/γ and σ_u are reduced by one standard deviation. The symmetric operation, namely adding one standard deviation, has qualitatively similar effects.

The remaining parameters, g, $\bar{r}^f$, ψ, σ_v, and ρ are set equal to their point estimates. Then $S_{\max}$, and $\bar{S}$ are derived from these "primitive parameters" using (11) and (12). δ is derived by taking unconditional expectations in the equation for the riskfree rate (14). Table 2 and Table 3 summarize the parameter choices.

4.2 The Price-Dividend Ratio and the Yield Spread

Figures 2 and 3 plot the price-dividend ratio and the yield spread as a function of s_t, for $z_t = g$, and $z_t = g \pm \frac{\sigma_u}{\sqrt{1-\psi^2}}$[10]. These figures establish a simple, but important, result: the price-dividend ratio is increasing in s_t, while the yield spread is decreasing in s_t. A second result is that the price-dividend ratio depends almost exclusively on s_t, while the yield spread depends both on s_t and z_t. Fama and French (1989),

[10] $\frac{\sigma_u}{\sqrt{1-\psi^2}}$ equals the unconditional standard deviation of z_t.

Ferson and Merrick (1987), and Lettau and Ludginvson (2000) discuss the relationship between the price-dividend ratio and the business cycle. Taking the price-dividend ratio to be a measure of the business cycle, Figure 2 implies that s_t is a pro-cyclical variable. Moreover, this interpretation and Figure 3 imply that the yield spread is countercyclical.

The price-dividend ratio behaves almost identically for each of the four sets of parameter values, while the yield spread changes dramatically. The yield spread varies much more with z_t than does the price-dividend ratio. The larger the standard deviation σ_u, the greater the variation becomes. Varying b also has large effects. As b falls, the yield spread shifts down for every value of s_t and z_t. Figure 3 also shows that the model is capable of generating a downward-sloping yield curve.

To summarize, the price-dividend ratio is pro-cyclical, and depends almost entirely on s_t alone. The yield spread is countercyclical, and depends on both s_t, the business cycle variable, and z_t, the long-run consumption trend. These basic facts help to interpret the simulation results below.

4.3 Statistics for the Stock Market, the Riskfree Rate, and Aggregate Consumption

Table 4 describes the first and second moments of the excess return on stocks, the riskless return, and the price-dividend ratio for each value of ρ. The model provides a remarkably good fit to the mean and standard deviation of stock returns. This is not a mechanical feature of the model: γ was chosen to match ratio of the mean to the standard deviation, not the individual levels. Thus the model can fit the equity premium puzzle of Mehra and Prescott (1985).

In addition, the model produces moments very close to the mean and the standard deviation of the price-dividend ratio. In this sense, the model improves on Campbell and Cochrane (1999), as their model matched the standard deviation, but not the mean. It is surprising that this model, designed to provide a better fit to consumption and interest rate data, actually results in a superior fit to stock market data. The high volatility of the price-dividend ratio, as well as stock returns demonstrates that the model fits the volatility puzzle described by Shiller (1981). Stock returns and price-dividend ratios are highly volatile even though the dividend process is calibrated to the extremely smooth postwar consumption data.

Finally, the model matches the mean and provides a realistic standard deviation for consumption and interest rates. Thus the model resolves the riskfree rate puzzle of Weil (1989) in addition to the equity premium and volatility puzzles above. The low mean of interest rates follows from the fact that the γ required to fit the Sharpe ratio is very low, unlike in the traditional power utility model.

Unlike the means of consumption growth and interest rates, the volatilities do not appear as parameters to be estimated in the GMM equations. The volatility of consumption is very similar, or somewhat higher, than that found in the data, depending on the parameters chosen. The volatility of the riskfree rate is almost always lower than that found in the data. It is a positive feature of the model that the riskfree rate in the model is less volatile than the one in the data as the riskfree rate in the data is measured with noise due to inflation. The low volatility of the riskfree rate is due not only to the low estimate of γ mentioned above, but to the relatively low estimate of b, the coefficient on surplus consumption. This coefficient is estimated to be significantly negative, but still quite low. Thus the model displays a realistic level of interest rate volatility.

Overall, the aggregate market characteristics vary very little as the parameters vary within the ranges specified by the estimation in the previous section. The success of the model in fitting the stylized facts is derived from the underlying features of the model, not a precise set of parameters.

4.4 Excess Bond Returns and Yields

The parameter values for this model were designed to match the time series behavior of interest rates and consumption. It is therefore surprising that they should also match the cross-sectional behavior bonds. Table 5 shows summary statistics for excess returns on zero-coupon bonds, for each of the four sets of parameters, and for maturities equal to 1, 3, 5 and 10 years. For all parameter values, bonds display positive risk premia that increase slowly with maturity. This pattern results from the estimated coefficent b, and the correlation ρ. If the estimation had implied $b < 0$, bond premia would be negative. Alternatively, ρ might have been found to be negative. This would have increased expected returns on bonds beyond realistic levels.

Instead, the time-series data find a happy medium for the cross-section. Excess bond returns increase with the maturity, but not sufficiently fast to be inconsistent with the data. The standard deviation, Sharpe ratio, and betas on bond returns also increases with the maturity. In all cases, however, they are lower than that for equities. The beta on bonds is relatively low, even though the same factors influence both bond returns and stock market returns. Because bonds load only slightly on s_t, and equities load only slightly on z_t, the two factors are enough to generate independent variation between bonds and stocks.

The differences among the panels of Table 5 makes sense. When bonds load more on s_t (i.e. b is higher), interest rates are more negatively correlated with consumption, bond returns are more highly correlated, and thus are more risky from the point of view of investors. But the higher the risk premium, the more variations in s_t affect bond prices. Thus the standard deviation and beta are higher and rise faster with the maturity. Raising σ_u has the opposite effect. A higher value of σ_u dampens down the correlation between interest rates and consumption, not just because of the added noise, but because the correlation ρ between interest rates and consumption is positive (while b is negative). This lowers expected returns and betas on bonds. Overall, the results for the lower value of b are more consistent with what is found in the data (see, e.g. Campbell and Viceira, 1999), particularly for bonds with high maturities.

The results for means and standard deviations of bond yields in Table 6 mirror that for returns. The average yield curve is upward sloping for all parameter values. The slope is similar across all three sets of parameter values, and is similar to that found in the data. Moreover, for σ_u equal to its higher values, long yields are less volatile than short yields, while for the lower σ_u, the volatility of yields is mostly flat.

The results in this section apply to real bonds, while the evidence on bond returns pertains to nominal bonds. This distinction is likely to be less important for the means of bond returns than for the variances; variances of ex post nominal bond returns will certainly be higher than those of real bond returns. This observation actually helps the model: for $b/\gamma = .004$, the variances of bond returns are below those in the data at all maturities (see Campbell and Viceira (2000) for moments of nominal bond returns). For $b/\gamma = .007$, the variances are either the same or lower, depending on the maturity of the bond.

4.5 Why do Bonds have Risk Premia?

The most surprising feature of the cross-sectional behavior of bonds is that bonds have risk premia, and thus an upward-sloping yield curve, at all. These bonds are not subject to inflation risk, nor are they subject to default risk. Why, then, do bonds have risk premia?

Examining simulation results for the power utility model sheds some light on this question. Because s_t does not enter the riskfree rate, ρ is set equal to -1 to give the power utility model the best shot at producing bond premia and a downward sloping interest rate. Risk aversion $\hat{\gamma} = 4$, and the rate of time preference, $\hat{\delta} = 0.99$.[11]

As shown in Table 7, the risk premia on bonds under power utility are small and decrease with the maturity, and the yield curve is essentially flat. Under power utility, bonds are subject to interest-rate risk, generating a small risk premium. Thus the power utility results fit with the intuition above. Even though interest rates vary substantially, as shown by $\sigma(y_1)$, long bonds are essentially riskless. Bond premia are a feature unique to the habit formation model.

In fact, under the habit formation model, bonds subject to not only interest rate risk, but also business-cycle risk as represented by s_t. Provided there is some risk premium produced by time variation in the interest rate, that risk premium will vary with s_t because of changes in investor's risk aversion. This causes bond prices to vary with s_t as well. Thus, bonds have risk premia because are sensitive to interest rates, and because their prices fall in bad times.

4.6 Forecasting Excess Returns on Bonds and Stocks

Table 8 shows that excess returns on stocks and bonds in the model are forecastable in a pattern closely resembling that of the data. For all sets of parameter values, the price-dividend ratio predicts stock returns with a negative sign. Thus the model

[11]This produces a mean riskfree rate that is much too high; requiring the power utility model to match the mean riskfree rate implies that $\hat{\delta} > 1$. Even if one was not bothered by $\hat{\delta} > 1$, there would be a technical problem. A necessary and sufficient condition for the convergence of

$$P_{\text{market},t} = \sum_{n=1}^{\infty} P_{nt}$$

is that

$$\ln \hat{\delta} + (1-\hat{\gamma})g + \frac{1}{2}(1-\hat{\gamma})^2 \left(\sigma_v^2 + \frac{1}{(1-\psi)^2}\sigma_u^2 + \frac{2}{1-\psi}\rho\sigma_u\sigma_v \right) < 0. \tag{47}$$

Given the other parameter choices, it is not possible to choose $\hat{\delta}$ to satisfy this equation and match the mean of the riskfree rate.

replicates the result of Campbell and Shiller (1988), namely that high price-dividend ratios predict low future stock returns. The magnitude of the R^2 match that found in the data: at the 1-year horizon returns are barely predictable, while the R^2 rise to high levels at the 4-year horizon.

In addition, the model generates the failure of the expectations hypothesis documented by Campbell and Shiller (1991) and Fama and Bliss (1987). The general form of the expectations hypothesis states that the expected returns on long bonds should equal a constant plus the riskfree rate. Campbell and Shiller and Fama and Bliss find instead that high yield spreads forecast high excess bond returns. Table 8 shows that for all parameter values, the model also has this property. Bond returns also display greater predictability, as measured by the R^2, at long horizons. Table 8 also demonstrates a "cross-predictability" effect. The price-dividend ratio predicts excess bond returns, while the yield spread predicts excess bond returns, similar to that found by Fama and French (1989).

All four sets of parameter values display the same qualitative effects, but the quantitative results vary. When σ_u equals its higher value, the yield spread has substantially less ability to predict bond and stock returns. Going from the point estimate to one standard deviation lower for σ_u almost doubles the R^2 in the regression. As illustrated in Figure 3 when σ_u is high, there is more variation in the yield spread due to z_t. Because excess returns depend only on s_t (as shown in Section 2) this variation lowers the R^2.

The predictability generated by this model fits the broader point made by Fama and French (1989): that common business-cycle related variables underly excess returns in stocks and bonds. In the case of the model, the common underlying variable is the surplus consumption ratio s_t. Price-dividend ratios are pro-cyclical (increasing in s_t), while yield spreads and risk premia are counter-cyclical (decreasing in s_t). The model displays these effects for any parameter values such that bonds have positive risk premia. Moreover, for the parameter values in the range suggested by the data, the effects are quantitatively large.

5 Conclusion

This paper demonstrates that stock market and bond market facts can be explained simultaneously in a consumption-based general equilibrium model. In particular, the model accounts for the predictability in bond returns as documented by Campbell and Shiller (1991) and Fama and Bliss (1987), as well as predictability in stock returns and the equity premium. The term structure matches the upward sloping yield curve from the data, downward sloping volatility, and positive and increasing risk premia on bonds.

The model in this paper assumes the utility function from Campbell and Cochrane (1999), but generalizes the endowment and interest rate processes. Rather than fitting the additional parameters to match stylized facts, the parameters are estimated directly from the data. Consumption-based asset pricing models generally imply that the interest rate depends only on expected future consumption growth. The habit formation model in this paper is unique in implying that interest rates may depend not only on future consumption growth, but on past consumption growth as well. Surprisingly, the latter effect turns out to be significant in the data. Thus the model is estimated *from the time series alone*, but succeeds in fitting the stylized facts about the cross-section. The end result is a model for the term structure that depends on factors with analogues in macroeconomics. Thus this paper unites not only stock and bond pricing, but connects them both to underlying macroeconomic behavior.

References

Abel, Andrew, 1990, Asset prices under habit formation and catching up with the joneses, *American Economic Review Papers and Proceedings* 80, 38–42.

Backus, David, Silverio Foresi, and Chris Telmer, 2000, Affine term structure models and the forward premium anomaly, Forthcoming, *Journal of Finance.*

Bekaert, Geert, and Steven R. Grenadier, 1999, Stock and bond pricing in an affine economy, Research Paper No. 1584, Stanford University.

Campbell, John Y., 1986, Bond and stock returns in a simple exchange model, *Quarterly Journal of Economics* 101, 785–803.

Campbell, John Y., 1999, Asset prices, consumption, and the business cycle, in J.B. Taylor, and M. Woodford, eds.: *Handbook of Macroeconomics* (Elsevier Science B.V.,).

Campbell, John Y., and John H. Cochrane, 1995, By force of habit: A consumption-based explanation of aggregate stock market behavior, NBER Working Paper No. 4995.

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 58, 495–514.

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

Campbell, John Y., and Luis Viceira, 2000, Who should buy long-term bonds?, Forthcoming, *American Economic Review* .

Cecchetti, S.G., P.S. Lam, and N.C. Mark, 1990, Mean reversion in equilibrium asset prices, *Americal Economic Review* 80, 398–418.

Chapman, David A., 1998, Habit formation and aggregate consumption, *Econometrica* 66, 1223–1230.

Chou, Ray Y., Robert F. Engle, and Alex Kane, 1992, Measuring risk aversion from excess returns on a stock index, *Journal of Econometrics* 52, 201–224.

Constantinides, George M., 1990, Habit formation: A resolution of the equity premium puzzle, *Journal of Political Economy* 98, 519–543.

Dai, Qiang, 2000, From equity premium puzzle to expectations puzzle: A general equilibrium production economy with stochastic habit formation, Working Paper, New York University.

Dybvig, Philip H., 1995, Dusenberry's ratcheting of consumption: Optimal dynamic consumption and investment given intolerance for any decline in standard of living, *Review of Economic Studies* 62, 287–313.

Epstein, Larry G., and Stanley E. Zin, 1991, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis, *Journal of Political Economy* 99, 263–286.

Fama, Eugene F., and Robert R. Bliss, 1987, The information in long-maturity forward-rates, *American Economic Review* 77, 680–692.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 29, 23–49.

Ferson, Wayne E., and George M. Constantinides, 1991, Habit persistence and durability in aggregate consumption: Empirical tests, *Journal of Financial Economics* 29, 199–240.

Ferson, Wayne E., and John J. Merrick Jr., 1987, Non-stationarity and stage-of-the business-cycle effects in consumption-based asset pricing relations, *Journal of Financial Economics* 19, 127–46.

Hansen, Lars Peter, and Ken Singleton, 1983, Stochastic consumption, risk aversion, and the temporal behavior of asset returns, *Journal of Political Economy* 91, 249–268.

Harvey, Campbell, 1989, Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24, 289=317.

Heaton, John C., 1995, An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications, *Econometrica* 63.

Kandel, Shmuel, and Robert F. Stambaugh, 1991, Asset returns and intertemporal preferences, *Journal of Monetary Economics* 27, 39–71.

Lettau, Martin, and Sydney Ludvingson, 2000, Consumption, aggregate wealth and expected stock returns, Forthcoming, *Journal of Finance.*

Li, Yuming, 2000, Expected returns and habit persistence, Forthcoming, *Review of Financial Studies.*

Mehra, Rajnish, and Edward Prescott, 1985, The equity premium puzzle, *Journal of Monetary Economics* 15, 145–161.

Shiller, Robert J., 1981, Do stock prices move too much to be justified by subsequent changes in dividends?, *American Economic Review* 71, 421–436.

Singleton, Ken, 1990, Specification and estimation of intertemporal asset pricing models, in B. Friedman, and F. Hahn, eds.: *Handbook of Monetary Economics* (North-Holland, Amsterdam).

Sun, Tong-Sheng, 1992, Real and nominal interest rates: A discrete-time model and its continuous-time limit, *Review of Financial Studies* 5, 581–611.

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

Vasicek, Oldrich, 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177–188.

Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 402–421.

Appendix

A Solving for Prices under Habit Formation

From the investor's Euler equation, it follows that prices obey the recursion

$$\frac{P_{nt}}{C_t^\theta} = E_t\left[M_{t+1}\frac{P_{n-1,t+1}}{C_{t+1}^\theta}\right] \tag{48}$$

with boundary coundition

$$\frac{P_{0t}}{C_t^\theta} = 1. \tag{49}$$

Substituting for M_{t+1} from (7) and for $P_{n-1,t+1}/C_{t+1}^\theta$ from (25) implies

$$\begin{aligned}
\frac{P_{nt}}{C_t^\theta} &= E_t\left[\delta\exp\{(\theta-\gamma)(z_t+v_{t+1})+A(n-1)(z_{t+1}-g)+B(n-1)\}\left(\frac{S_{t+1}}{S_t}\right)^{-\gamma}F(s_{t+1},n-1)\right] \\
&= \delta\exp\{(\theta-\gamma)z_t+A(n-1)\psi(z_t-g)+B(n-1)\} \\
&\quad\times E_t\left[\exp\{(\theta-\gamma)v_{t+1}+A(n-1)u_{t+1}\}\left(\frac{S_{t+1}}{S_t}\right)^{-\gamma}F(s_{t+1},n-1)\right].
\end{aligned} \tag{50}$$

By conditioning on v_{t+1} and using the law of iterated expectations, u_{t+1} may be integrated out of the expectation. From the properties of conditional distributions, it follows that

$$A(n-1)u_{t+1}|v_{t+1} \sim N\left(\frac{\sigma_u}{\sigma_v}\rho A(n-1), \sigma_u^2(1-\rho^2)A(n-1)^2\right).$$

Therefore

$$\begin{aligned}
\frac{P_{nt}}{C_t^\theta} &= \delta\exp\left\{(\theta-\gamma)z_t+A(n-1)\psi(z_t-g)+\frac{\sigma_u^2}{2}(1-\rho^2)A(n-1)^2+B(n-1)\right\} \\
&\quad\times E_t\left[\exp\left\{(\theta-\gamma)v_{t+1}+\frac{\sigma_u}{\sigma_v}\rho A(n-1)v_{t+1}\right\}\left(\frac{S_{t+1}}{S_t}\right)^{-\gamma}F(s_{t+1},n-1)\right].
\end{aligned} \tag{51}$$

Comparing (51) with (25) implies that

$$\begin{aligned}
A(n) &= (\theta-\gamma)+\psi A(n-1) \\
B(n) &= \ln\delta+(\theta-\gamma)g+\frac{\sigma_u^2}{2}(1-\rho^2)A(n-1)^2+B(n-1)
\end{aligned}$$

and

$$F(s_t, n) =$$

$$E_t \left[\exp \left\{ \left[\frac{\sigma_u}{\sigma_v} \rho A(n-1) + (\theta - \gamma) \right] v_{t+1} \right\} \left(\frac{S_{t+1}}{S_t} \right)^{-\gamma} F(s_{t+1}, n-1) \right].$$

Solving backwards as in the power utility case yields the expressions for A, B, and F given in the text.

B Solving for Prices under Power Utility

Substituting for M_{t+1} from (31), and for $P_{n-1,t+1}/C_t^\theta$ from (34) in equations (48) and (18) implies

$$\frac{P_{nt}}{C_t^\theta} = E_t \left[\hat{\delta} \exp \left\{ (\theta - \hat{\gamma})(z_t + v_{t+1}) + \hat{A}(n-1)(z_{t+1} - g) + \hat{B}(n-1) \right\} \right]. \quad (52)$$

Substituting in for z_{t+1} using (3), it follows that

$$\frac{P_{nt}}{C_t^\theta} = \hat{\delta} \exp \left\{ (\theta - \hat{\gamma}) z_t + \hat{A}(n-1)\psi(z_t - g) + \hat{B}(n-1) \right\}$$
$$E_t \left[\exp \left\{ (\theta - \hat{\gamma}) v_{t+1} + \hat{A}(n-1) u_{t+1} \right\} \right]. \quad (53)$$

Integrating out the expectation and comparing the resulting formula with (34) implies the following formulas for A and B:

$$\begin{aligned} \hat{A}(n) &= (\theta - \hat{\gamma}) + \psi \hat{A}(n-1) \\ \hat{B}(n) &= \hat{B}(n-1) + \ln \hat{\delta} + (\theta - \hat{\gamma}) g \\ &\quad + \frac{1}{2} \left[\hat{A}(n-1)^2 \sigma_u^2 + 2\hat{A}(n-1)(\theta - \hat{\gamma}) \rho \sigma_u \sigma_v + (\theta - \hat{\gamma})^2 \sigma_v^2 \right]. \end{aligned}$$

From the boundary condition (49) it follows that $\hat{A}(0) = \hat{B}(0) = 0$. Solving backwards,

$$\hat{A}(n) = (\theta - \hat{\gamma})(1 + \psi + \cdots + \psi^{n-1}) = (\theta - \hat{\gamma}) \frac{1 - \psi^n}{1 - \psi}.$$

Substituting into the formula for $\hat{B}(n)$ above yields (36).

C Approximation for Habit in terms of Past Consumption

Near the steady state, the transition equation for s_t is approximately

$$s_{t+1} - \bar{s} \approx \phi(s_t - \bar{s}) + \lambda(\bar{s})(\Delta c_{t+1} - z_t) \tag{54}$$

$$\approx \phi(s_t - \bar{s}) + \lambda(\bar{s})(\Delta c_{t+1} - g) \tag{55}$$

The first line follows from setting $s_t \approx \bar{s}$, the second from $z_t \approx g$ in the variance term in (5). Solving forward produces[12]

$$s_t - \bar{s} = \lambda(\bar{s}) \left(\sum_{j=0}^{\infty} \phi^j \Delta c_{t-j} - \frac{g}{1-\phi} \right). \tag{56}$$

The main purpose of these approximations, besides intuition, is to derive a convenient way of matching parameters to the data. It is not necessary that the approximations be accurate in every respect, just that the moments of the approximate model that are used to match the data are close to those of the exact model. This can be easily verified by rerunning the regressions using simulated data from the exact model.

A first-order approximation around $s_t = \bar{s}$ implies

$$x_t \approx c_t + \ln(1 - e^{\bar{s}}) - (s_t - \bar{s}) \frac{e^{\bar{s}}}{1 - e^{\bar{s}}}$$

$$s_t - \bar{s} \approx \left(1 - \frac{1}{\bar{S}}\right)(x_t - c_t - h), \tag{57}$$

where $h = \ln(1 - \bar{S})$. It follows from (56) that

$$x_{t+1} \approx h + \frac{g}{1-\phi} + (1-\phi) \sum_{j=0}^{\infty} \phi^j c_{t-j}.$$

[12]Without any approximations,

$$s_t - \bar{s} = \sum_{j=0} \phi^j \lambda(s_{t-1-j}) v_{t-j}.$$

Imposing only $s_t \approx \bar{s}$, we have

$$s_t - \bar{s} = \lambda(\bar{s}) \sum_{j=0} \phi^j v_{t-j}$$

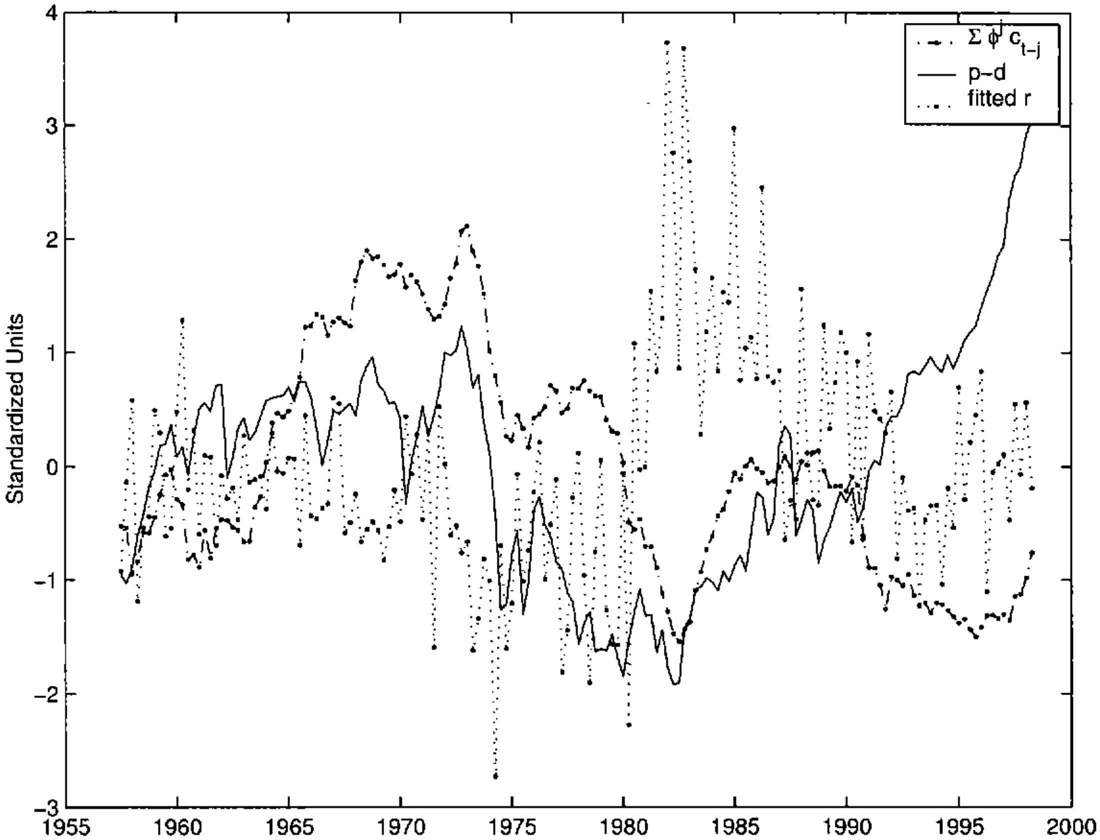

Figure 1: This figure plots the history of average past consumption growth $\sum_{t=0}^{40} \phi^j c_{t-j}$, the log price-dividend ratio, and the fitted value of the interest rate from the first-stage regression described in Section 3.3. Variables are de-meaned and standardized

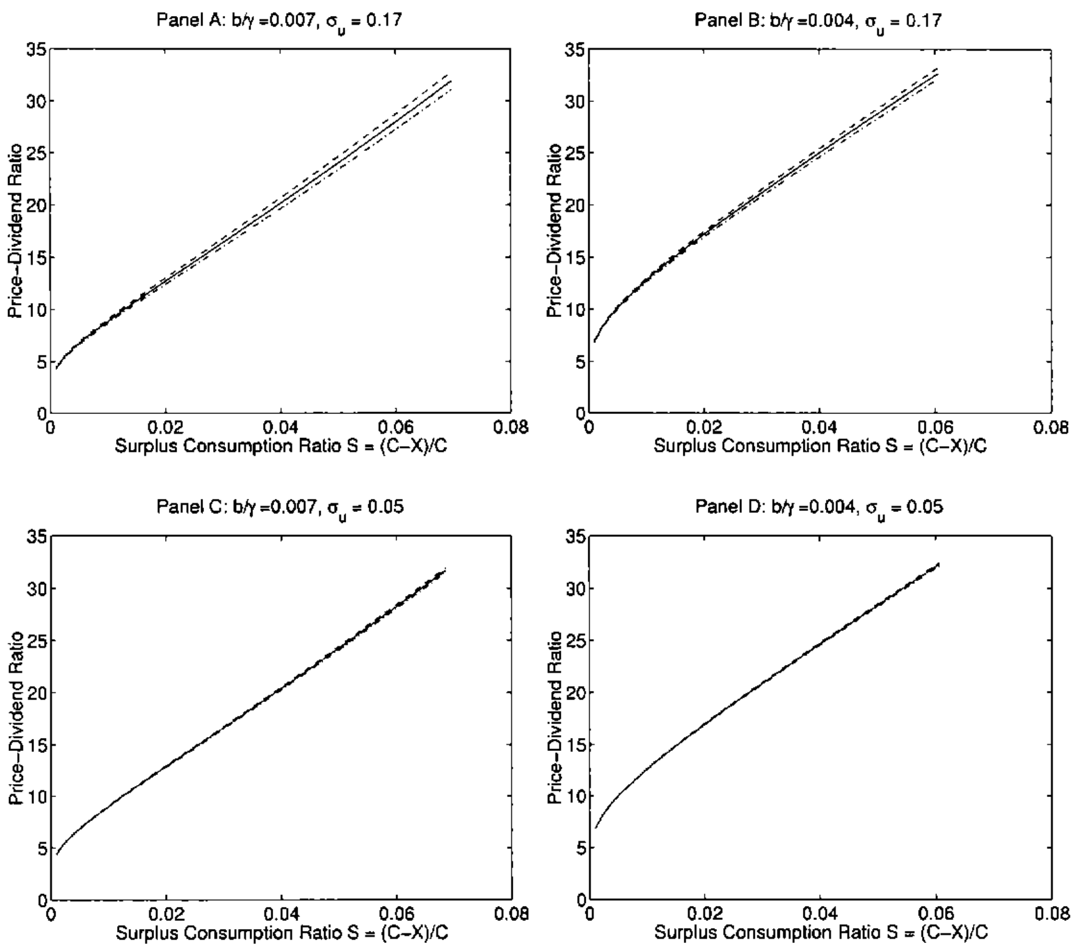

Figure 2: The Price-Dividend Ratio as a Function of S_t. The price-dividend ratio refers to the ratio of aggregate wealth to aggregate consumption. $z_t = g + \sigma_u/\sqrt{1-\psi^2}$ $(-.-)$, $z_t = g$ $(-)$, and $z_t = g - \sigma_u/\sqrt{1-\psi^2}$ $(--)$.

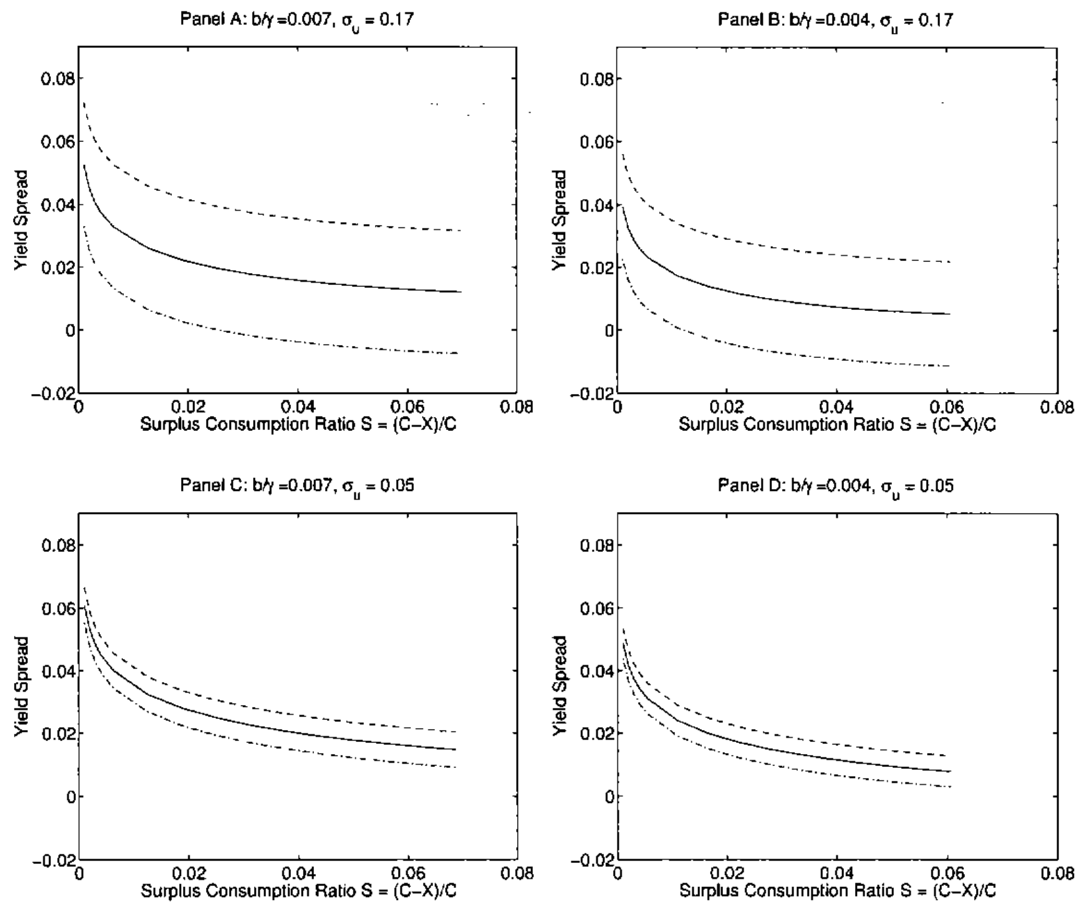

Figure 3: The Yield Spread as a Function of S_t. The yield spread refers to the yield on the 10-year zero less that of the riskless asset. $z_t = g+\sigma_u/\sqrt{1-\psi^2}$ (−.−), $z_t = g$ (−), and $z_t = g - \sigma_u/\sqrt{1-\psi^2}$ (−−).

Panel A: Interest Rates and Habit on RHS

	Const.	r^f_{t+1}	$\sum \phi^j \Delta c_{t-j}$
Δc_{t+1}	0.005	0.234	0.092
	(0.000)	(0.157)	(0.042)
R^2	-	0.256	0.227

Panel B: Consumption and Habit on RHS

	Const.	Δc_{t+1}	$-\sum \phi^j \Delta c_{t-j}$
r^f_{t+1}	-0.008	1.583	-0.270
	(0.004)	(0.848)	(0.070)
R^2	-	0.060	0.227

Table 1: Unrestricted instrumental variable estimates of (43) (Panel A) and (45) (Panel B). Standard errors are in parentheses. R^2 is for the first-stage regression of the RHS variables onto the instruments.

Parameter	Estimate	S. E.
g	0.0051	0.0004
γ	4.545	3.195
b/γ	0.0070	0.0030
$\bar{r}^f$	0.0049	0.0006
ψ	0.908	0.143
σ_v	0.0051	0.0005
σ_u	0.0017	0.0012
ρ	0.310	0.124

Table 2: Generalized method of moments estimation of (43) and (44). The persistence of s_t, ϕ, is set equal to 0.969 (quarterly units).

	Panel A: $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.17$	Panel B: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.17$
Utility curvature γ	1.650	1.400
Discount factor δ	0.984	0.983
Habit coefficient b	0.012	0.006
Persistence ϕ	0.969	0.969
Steady-state surplus cons. $\bar{S}$	0.042	0.037
Max. surplus cons. $S_{\max}$	0.070	0.061

	Panel C: $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.05$	Panel D: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.05$
Utility curvature γ	1.600	1.400
Discount factor δ	0.984	0.983
Habit coefficient b	0.011	0.006
Persistence ϕ	0.969	0.969
Steady-state surplus cons. $\bar{S}$	0.042	0.037
Max. surplus cons. $S_{\max}$	0.069	0.061

Table 3: Utility Parameters. The model is simulated at a quarterly frequency; parameters are in natural units. b refers to minus the coefficient on surplus consumption in the riskfree rate (14).

	Panel A : $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.17$	Panel B: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.17$	Data
$E(r^m - r^f)$	4.92	4.33	4.91
$\sigma(r^m - r^f)$	16.28	14.22	16.34
Sharpe*	0.30	0.30	0.30
$E(P/D)$	22.44	25.20	27.91
$\sigma(P/D)$	0.32	0.27	0.29
Corr(P/D)*	0.97	0.97	0.97
$E(r^f)$	2.00	1.94	1.97
$\sigma(r^f)$	1.52	1.17	1.49
$E(\Delta c_{t+1})$*	2.02	2.02	2.06
$\sigma(\Delta c_{t+1})$	1.30	1.31	0.96

	Panel C : $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.05$	Panel D: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.05$	Data
$E(r^m - r^f)$	4.93	4.39	4.91
$\sigma(r^m - r^f)$	16.51	14.53	16.34
Sharpe*	0.30	0.30	0.30
$E(P/D)$	22.61	24.73	27.91
$\sigma(P/D)$	0.31	0.28	0.29
Corr(P/D)*	0.97	0.97	0.97
$E(r^f)$	1.96	1.99	1.97
$\sigma(r^f)$	1.05	0.65	1.49
$E(\Delta c_{t+1})$*	2.03	2.04	2.06
$\sigma(\Delta c_{t+1})$	1.05	1.05	0.96

Table 4: Summary Statistics. Means and standard deviations of returns and consumption growth are in annualized percentages. The Sharpe ratio, in annual units, is the ratio of the first row to the second. The first-order autocorrelation of the price-dividend ratio is in quarterly units. (*) denotes a moment that the model is designed to fit exactly.

	Panel A: $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.17$				Panel B: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.17$			
mat	$E(r-r^f)$	$\sigma(r-r^f)$	Sharpe	Beta	$E(r-r^f)$	$\sigma(r-r^f)$	Sharpe	Beta
4	0.15	1.53	0.10	0.03	0.04	1.26	0.03	0.01
12	0.70	4.34	0.16	0.14	0.28	3.38	0.08	0.07
20	1.44	6.49	0.22	0.27	0.69	4.69	0.15	0.15
40	3.55	12.40	0.29	0.68	2.20	8.17	0.27	0.45

	Panel C: $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.05$				Panel D: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.05$			
mat	$E(r-r^f)$	$\sigma(r-r^f)$	Sharpe	Beta	$E(r-r^f)$	$\sigma(r-r^f)$	Sharpe	Beta
4	0.26	0.82	0.32	0.04	0.14	0.52	0.26	0.02
12	1.04	3.04	0.34	0.16	0.58	1.86	0.31	0.10
20	1.89	5.49	0.34	0.31	1.13	3.36	0.34	0.20
40	4.01	12.44	0.32	0.73	2.77	8.24	0.34	0.53

Table 5: Excess Bond Returns. Summary statistics are reported for zero-coupon bonds maturing in 1,4,12,20, and 40 quarters (0.25, 1,3,5,10 years). Means and standard deviations are in annualized percentages. The Sharpe ratio (annualized) is the ratio of the first column to the second.

Panel A: $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.17$

mat	$E(y_n)$	$\sigma(y_n)$
1	2.00	1.52
4	2.07	1.41
12	2.31	1.25
20	2.63	1.21
40	3.58	1.27

Panel B: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.17$

mat	$E(y_n)$	$\sigma(y_n)$
1	1.94	1.17
4	1.95	1.05
12	2.05	0.85
20	2.20	0.77
40	2.79	0.80

Panel C: $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.05$

mat	$E(y_n)$	$\sigma(y_n)$
1	1.96	1.05
4	2.09	1.06
12	2.46	1.11
20	2.87	1.17
40	3.92	1.29

Panel D: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.05$

mat	$E(y_n)$	$\sigma(y_n)$
1	1.99	0.65
4	2.06	0.65
12	2.27	0.68
20	2.51	0.73
40	3.24	0.88

Table 6: Bond Yields. Maturities are reported in quarters. Means and standard deviations are in annualized percentages.

Power Utility Bond Returns

mat	$E(r^b - r^f)$	$\sigma(r^b - r^f)$	Sharpe	Beta
4	0.20	3.75	0.05	0.38
12	0.23	9.76	0.02	0.99
20	0.13	12.55	0.01	1.28
40	-0.00	14.60	-0.00	1.48

Power Utility Bond Yields

mat	$E(y_n)$	$\sigma(y_n)$
1	11.64	3.40
4	11.75	2.96
12	11.84	2.12
20	11.82	1.59
40	11.75	0.91

Table 7: The Term Structure Under Power Utility. The first panel reports summary statistics for excess bond returns, The second panel reports summary statistics for bond yields. All bonds are zero-coupon and mature at 1, 4, 12, 20, and 40 quarters. Means and standard deviations are in annualized percentages. The Sharpe ratio (annualized) is the ratio of the first column to the second

		Panel A: $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.17$				Panel B: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.17$			
		$r^m - r^f$		$r^b - r^f$		$r^m - r^f$		$r^b - r^f$	
		β	R^2	β	R^2	β	R^2	β	R^2
$\log(P/D)$	4	-0.42	0.05	-0.59	0.09	-0.48	0.06	-0.65	0.11
	8	-0.80	0.09	-1.10	0.16	-0.89	0.11	-1.22	0.19
	16	-1.42	0.16	-1.97	0.27	-1.59	0.20	-2.18	0.31
$y_n - y_1$	4	0.20	0.01	0.26	0.02	0.21	0.01	0.28	0.02
	8	0.37	0.02	0.49	0.03	0.39	0.02	0.53	0.04
	16	0.65	0.03	0.85	0.05	0.69	0.04	0.94	0.06

		Panel C: $\frac{b}{\gamma} = 0.007$, $\sigma_u = 0.05$				Panel D: $\frac{b}{\gamma} = 0.004$, $\sigma_u = 0.05$			
		$r^m - r^f$		$r^b - r^f$		$r^m - r^f$		$r^b - r^f$	
		β	R^2	β	R^2	β	R^2	β	R^2
$\log(P/D)$	4	-0.43	0.05	-0.60	0.10	-0.48	0.06	-0.72	0.13
	8	-0.80	0.09	-1.14	0.18	-0.90	0.11	-1.37	0.25
	16	-1.42	0.16	-2.02	0.31	-1.62	0.20	-2.45	0.42
$y_n - y_1$	4	0.31	0.03	0.45	0.05	0.37	0.04	0.58	0.09
	8	0.58	0.05	0.84	0.10	0.70	0.07	1.09	0.16
	16	1.03	0.08	1.50	0.17	1.27	0.13	1.97	0.27

Table 8: Forecasting Excess Returns on Bonds and Stocks. Bond and stock returns, measured at horizons of 4, 8, and 16 quarters, are regressed on the price-dividend ratio (first three columns) and the yield spread (second three columns). Each variable is divided by its standard deviation. The yield spread and the excess bond return are calculated using the 10-year bond.