

Habit Formation and Returns on Bonds and Stocks

Jessica A. Wachter *

May 15, 2002

Abstract

This paper proposes a habit formation model that explains the failure of the expectations hypothesis documented by Campbell and Shiller (1991) and Fama and Bliss (1987). The model also produces positive excess returns on long-term bonds, an upward sloping average yield curve, and allows for realistic levels of time-variation in the mean of consumption growth. The model generates a novel empirical prediction: Long lags of consumption growth predict the short-term interest rate with a negative sign. This prediction is shown to be strongly supported by the data.

*New York University, Stern School of Business. Email: jwachter@stern.nyu.edu, Phone: (212)998-0799. I thank Michael Brandt, Geert Bekaert, John Campbell, John Cochrane, Vassil Konstantinov, Martin Lettau, Anthony Lynch, David Marshall, Christopher Telmer, Lasse Pederson, Andre Perold, Jeremy Stein, Matt Richardson, Stephen Ross, Robert Whitelaw, Yihong Xia, seminar participants at the 2001 NBER Asset Pricing meeting in Los Angeles, the NYU Macro lunch, the New York Federal Reserve, Washington University, and the Wharton School. I thank Lehman Brothers for financial support.

Introduction

The expectations puzzle, documented by Campbell and Shiller (1991) and Fama and Bliss (1987), has long been a challenge for general equilibrium models of the term structure. Backus, Gregory, and Zin (1989) show that a model assuming power utility preferences and time-varying expected consumption growth cannot account for this finding. Although Dai and Singleton (2001) show that a statistical model of the stochastic discount factor can fit the puzzle, this only raises the question of what economic mechanism is at work.

This paper proposes a general equilibrium model that captures the empirical results of Campbell and Shiller (1991) and Fama and Bliss (1987) and links them to investor preferences and aggregate consumption. In particular, Campbell and Shiller run the regression

$$y_{t+1}^{n-1} - y_t^n = \text{constant} + \beta_n \frac{1}{n-1} (y_t^n - y_t^1) + \text{error},$$

where $y_t^n = \frac{1}{n} \ln P_t^n$, and P_t^n is the price of a bond with maturity n . According to the expectations hypothesis, excess returns on bonds are unpredictable, and all the variation in yield spreads is due to variation in future short-term interest rates. In terms of the regression above, this means $\beta_n = 1$ for all n . But Campbell and Shiller show, on the contrary, that β_n is negative and decreasing in n . The model in this paper reproduces both of these findings.¹ The model also implies an upward sloping average yield curve and positive risk premia on bonds, both of which are found in the data.

Two ingredients enable the model to capture these findings. The first is external habit persistence from Campbell and Cochrane (1999). Habit persistence generates time variation in investor preferences. After periods of unusually low consumption growth, the volatility of investors' marginal utility rises, causing them to demand greater premia on risky assets. As a result, the risk premium on the aggregate stock market varies in a countercyclical fashion.

¹There are a number of econometric difficulties with this regression. Non-exogenous regressors bias the coefficients upward, causing the hypothesis to be rejected less strongly than it should be (Stambaugh (1999)), while Peso problems (Bekaert, Hodrick, and Marshall (2001)) result in increased dispersion of the estimates, leading to the model to be rejected too strongly. Bekaert and Hodrick (2001) argue that standard tests tend to reject the null of the expectations hypothesis even when it is true. They find, however, that the data remain inconsistent with the expectations hypothesis, even after adjusting for small-sample properties. Given that the expectations puzzle survives more rigorous statistical analysis, it makes sense to seek a general equilibrium explanation.

Habit utility preferences are clearly not enough: In the model of Campbell and Cochrane (1999), the riskfree rate is constant and the term structure is trivial. The second ingredient is thus a model for the short-term interest rate that makes long-term bonds risky in the first place. Without this ingredient, it is impossible for long-term bonds to have positive, countercyclical risk premia.

In this paper, the short-term interest rate is driven by two factors. The first is expected future consumption growth. When expected consumption growth is high investors borrow against it, thus driving up the interest rate. As shown below, aggregate consumption data provides support for time-varying consumption growth of the type assumed in the model. The second factor is past consumption growth. The model in this paper implies that when past consumption growth is relatively low, investors borrow to give habit a chance to catch up to consumption. The hypothesis is that past consumption growth predicts the short-term interest rate with a negative sign. This hypothesis is confirmed using consumption and interest rate data. It turns out that a long-run moving average of past consumption growth indeed predicts the interest rate with a negative sign. This empirical finding is of interest outside of the context of habit formation; it has the ability to constrain any general equilibrium model that speaks to the relationship between consumption and interest rates.

Besides the empirical literature on the expectations hypothesis, this paper draws on the theoretical habit formation literature (e.g., Abel (1990), Chapman (1998), Constantinides (1990), Dybvig (1995), Heaton (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 that 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. 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 Motivated

² Lately there has been increased interest in the empirical properties of habit formation models. Dai (2000) links the Constantinides model to a model for labor income. Brandt and Wang (2001) study habit preferences over inflation. Menzly, Santos, and Veronesi (2001) use a variant of the Campbell and Cochrane model to explain returns on industry portfolios.

by habit formation models, Li (2001) examines the ability of past consumption growth to predict excess returns on stocks. However, Li does not look at the predictive ability of consumption for short or long-term interest rates, nor does he consider the implications for habit formation for the expectations hypothesis.

An intriguing feature of the model in this paper is the link it produces between asset returns and underlying macroeconomic variables. The price-dividend ratio is procyclical and captures past consumption growth. The yield spread is countercyclical and 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. Moreover, the model preserves the advantages of the original Campbell and Cochrane (1999) framework. It successfully captures the high equity premium for the aggregate market, excess volatility, and predictability of excess stock returns.

The outline of the paper is as follows. Section 1 describes the model, which generalizes that of Campbell and Cochrane (1999) to account for both time-variation in expected consumption growth, and predictability of interest rates by past consumption growth. Section 2 estimates the model from consumption and interest rate data. An unrestricted regression finds that past consumption growth predicts the interest rate with a negative sign. Then the primitive parameters are estimated taking into account the restrictions imposed by the model using the generalized method of moments technique of Hansen and Singleton (1982). The data fails to reject the over-identifying restrictions. Section 3 illustrates, through simulated data, the implications of the model for long-term bonds.

1 Habit Formation

This section describes the model assumed in this paper. Sections 1.1 and 1.2 describe the assumptions for endowments and preferences. Section 1.3 derives equations for pricing long-term bonds and the aggregate market. Section 1.4 relates the model to reduced-form models for the term structure.

1.1 Endowments

Assume that the consumption growth evolves according to:

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

where

$$\begin{bmatrix} v_{t+1} \\ u_{t+1} \end{bmatrix} \sim N\left(0, \begin{bmatrix} \sigma_v^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_u^2 \end{bmatrix}\right).\tag{2}$$

This specification was introduced in Campbell (1999). It is particularly convenient, because it leads to a two-factor model for the interest rate, for which z_t is one of the factors. It nests a number of models that are considered recently: the ARMA model for dividend growth considered by Bansal and Yaron (2000), and the AR(1) considered by Bekaert and Grenadier (2001). In addition, it is related to the Markov switching models of Cecchetti, Lam, and Mark (1990, 1993), Kandel and Stambaugh (1991), and Whitelaw (2000).

It is important to note that time-varying consumption growth plays a very different role in this paper than in the papers described above. Traditionally, time-varying consumption growth has been used to generate predictability in asset returns. That is not its role here. On the contrary, as will be shown later, time-varying consumption growth *moderates* predictability introduced by habit formation, it does not generate it.

1.2 Preferences

Assume that an investor has utility over consumption relative to a reference point X_t :

$$E \sum_{t=0}^{\infty} \delta^t \frac{(C_t - X_t)^{1-\gamma} - 1}{1-\gamma}.\tag{3}$$

Habit, X_t , is defined indirectly, through surplus consumption S_t , where

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

To ensure that X_t never falls below C_t , $s_t = \ln S_t$ is modeled:

$$s_{t+1} = (1 - \phi)\bar{s} + \phi s_t + \lambda(s_t)v_{t+1},\tag{4}$$

The process for s_t is heteroscedastic, and perfectly correlated with innovations in consumption growth. The sensitivity function $\lambda(s_t)$ will be described below.

The investor's habit is *external*: the investor does not take into account the effect that today's consumption decisions have on X_t in the future.³ Because habit is external, the investor's intertemporal marginal rate of substitution is given by:

$$M_{t+1} = \delta \left(\frac{S_{t+1} C_{t+1}}{S_t C_t} \right)^{-\gamma}. \quad (5)$$

From the Euler equation, it follows that the riskfree rate equals

$$\begin{aligned} r_{t+1}^f &= \ln(1/E_t[M_{t+1}]) \\ &= -\ln \delta + \gamma z_t + \gamma(1 - \phi)(\bar{s} - s_t) - \frac{\gamma^2 \sigma_v^2}{2} (1 + \lambda(s_t))^2. \end{aligned} \quad (6)$$

where $z_t = E_t(c_{t+1})$ from the previous section. This riskfree rate has some familiar terms from the power utility case and others that are new to habit formation. As in the power utility model, positive expected consumption growth leads investors to borrow from the future to smooth consumption. This is reflected in the term γz_t (however, γ is not equal to risk aversion as it is under power utility). The second term, proportional to $\bar{s} - s_t$, implies that as surplus consumption falls relative to its long-term mean, investors 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 reflects 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 function $\lambda(s_t)$ is chosen so that the intertemporal substitution and the precautionary savings effect offset each other. The net effect of s_t on the riskfree rate can then be calibrated to the data. For simplicity, $\lambda(s_t)$ is further restricted so that r_{t+1}^f is linear in s_t . In addition, $\lambda(s_t)$ is chosen so that for $s_t \approx \bar{s}$ and $z_t \approx g$, habit, x_t is a deterministic function of past consumption. These considerations imply that

$$\begin{aligned} \lambda(s_t) &= (1/\bar{S})\sqrt{1 - 2(s_t - \bar{s})} - 1 \\ \bar{S} &= \sigma_v \sqrt{\frac{\gamma}{1 - \phi - b/\gamma}}. \end{aligned}$$

³Formally, 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.

More details can be found in Appendix B.⁴ Substituting these equations into (6) reduces the riskfree rate equation to

$$\begin{aligned} r_{t+1}^f &= \left(-\ln \delta + \gamma g - \frac{\gamma(1-\phi) - b}{2} \right) + \gamma(z_t - g) + b(\bar{s} - s_t) \\ &= \bar{r}^f + \gamma(z_t - g) + b(\bar{s} - s_t) \end{aligned} \quad (7)$$

where b is a free preference parameter that will be estimated from the data, and \bar{r}^f equals the unconditional mean of r_{t+1}^f .

The parameter b has an economic interpretation that stems from (6). If $b > 0$, the intertemporal smoothing effect wins out, and an increase in surplus consumption s_t drives down the interest rate. If $b < 0$, the precautionary savings effect wins out. An increase in surplus consumption s_t decreases the sensitivity $\lambda(s_t)$ and drives down the interest rate. Moreover, b has an empirical interpretation. It follows from (4) that

$$s_t - \bar{s} \approx \lambda(\bar{s}) \left(\sum_{j=0}^{\infty} \phi^j \Delta c_{t-j} - \frac{g}{1-\phi} \right), \quad (8)$$

where $\lambda(\bar{s}) = 1/\bar{S} - 1$.⁵ Thus the parameter b can be estimated using consumption and interest rate data.

The model examined by Campbell and Cochrane (1999) is a special case of the model described above. Campbell and Cochrane assume that consumption growth has a constant mean. In addition they choose $\lambda(s_t)$ so that the riskfree rate is constant in s_t (the intertemporal smoothing and precautionary savings effects exactly cancel out). The Campbell and Cochrane model can be obtained from the equations above by setting $z_t = g$, and $b = 0$.⁶ While these choices simplify the model, they are not supported by the data. In contrast, this paper allows z_t to vary, and b to be nonzero. Both effects are estimated from the data and found to be statistically significant.

⁴In order that the quantity within the square root remain positive, $\lambda(s_t)$ is set to be 0 when $s_t > s_{\max}$, for $s_{\max} = \bar{s} + \frac{1}{2}(1 - \bar{S}^2)$. s_t ventures above s_{\max} sufficiently rarely that this feature does not affect the behavior of the model.

⁵ Near the steady state, the transition equation for s_t is approximately

$$\begin{aligned} s_{t+1} - \bar{s} &\approx \phi(s_t - \bar{s}) + \lambda(\bar{s})(\Delta c_{t+1} - z_t) \\ &\approx \phi(s_t - \bar{s}) + \lambda(\bar{s})(\Delta c_{t+1} - g) \end{aligned}$$

Solving forward produces the equation in the text.

⁶Campbell and Cochrane (1999) suggest how to introduce s_t into the riskfree rate. But they do not develop the implications of this alternative model.

While the functional form of $\lambda(s_t)$ is chosen to match the behavior of the riskfree rate, it has important implication for returns on risky assets. It follows from the investor's Euler equation that

$$\frac{E_t(R_{t+1} - R_{t+1}^f)}{\sigma_t(R_{t+1})} = -\rho_t(M_{t+1}, R_{t+1}) \frac{\sigma_t(M_{t+1})}{E_t(M_{t+1})},$$

where R_{t+1} is the return on some risky asset. As a consequence

$$\frac{E_t(R_{t+1} - R_{t+1}^f)}{\sigma_t(R_{t+1})} \approx -\rho_t(M_{t+1}, R_{t+1})\gamma\sigma_v(1 + \lambda(s_t)), \quad (9)$$

which follows from the lognormality of M_{t+1} conditional on time- t information. Because $\lambda(s_t)$ is decreasing in s_t , the ratio of the volatility of the stochastic discount factor to its mean varies countercyclically. This provides a mechanism by which Sharpe ratios, and hence risk premia, vary countercyclically over time.

In the model of Campbell and Cochrane (1999), the mechanism in (9) does not create time-varying risk premia on bonds for the simple reason that bond returns are constant, and equal to the riskfree rate at all maturities. In terms of (9), the Campbell and Cochrane model implies that $\rho_t(M_{t+1}, R_{t+1}) = 0$, when R_{t+1} is the return on a bond. However, the model in this paper generates a time-varying riskfree rate. Therefore $\rho_t(M_{t+1}, R_{t+1})$ is nonzero, and (9) provides a mechanism for risk premia on bonds, as well as risk premia on stocks, to vary through time. Of course, this observation alone does not solve the expectations puzzle. The sign of bond premia, and the magnitude of time-variation will depend on the results of the parameter estimation in Section 2.

1.3 Model Solution

This section calculates the prices of long-term bonds and stocks. Techniques from affine bond pricing are combined with numerical methods to produce a solution technique that is efficient and flexible.

Bond Prices

The bonds in this model are all real; they pay off in future units of the consumption good. Let P_t^n denote the price, at time t , of a bond maturing in n periods. The return from holding the bond between t and $t + 1$ is denoted R_{t+1}^n and is given by

$$R_{t+1}^n = \frac{P_{t+1}^{n-1}}{P_t^n}$$

The log yield to maturity is defined as

$$y_t^n = -\frac{1}{n} \ln P_t^n.$$

Note that $y_t^1 = r_{t+1}^f$.

The investor's Euler equation recursively determines bond prices:

$$P_t^n = E_t \left[\delta \left(\frac{S_{t+1} C_{t+1}}{S_t C_t} \right)^{-\gamma} P_{t+1}^{n-1} \right]. \quad (10)$$

When $n = 0$, the bond is worth one unit of the consumption good. This implies the boundary condition

$$P_t^0 = 1.$$

The distribution of S_{t+1}/S_t and C_{t+1}/C_t at time t depend only on the state variables s_t and z_t . It follows recursively from (10) that bond prices also are functions of s_t and z_t . In principle, (10) can be solved by numerical integration. For this problem, numerical integration is much superior to calculating the expectation by Monte Carlo. In this model, prices of assets are very sensitive to the probabilities of rare events, making simulation unreliable.

While straightforward in principle, solving (10) via numerical integration is complicated by the fact that there are two state variables. Fortunately a simple trick can be used to reduce computation time back to what it would be for a single state variable. Using the law of iterated expectations, and conditioning on realizations of the shock v_{t+1} it can be shown that bond prices must take the form

$$P_t^n = G(z_t, n) F(s_t, n), \quad (11)$$

for functions G and F . Further, G has a closed form solution:

$$G(z_t, n) = A(n) \exp \{ B(n)(z_t - g) \}$$

where

$$B(n) = -\gamma \frac{1 - \psi^n}{1 - \psi}$$

$$A(n) = \delta^n \exp \left\{ -n\gamma g + \frac{\gamma^2}{2} (1 - \rho^2) \sigma_u^2 \sum_{k=1}^{n-1} \left(\frac{1 - \psi^k}{1 - \psi} \right)^2 \right\}.$$

The function $F(s_t, n)$ is determined by one-dimensional numerical integration:

$$F(s_t, n) = E_t \left[\exp \left\{ -\gamma \left[\frac{\sigma_u}{\sigma_v} \rho \frac{1 - \psi^{n-1}}{1 - \psi} + 1 \right] v_{t+1} \right\} \left(\frac{S_{t+1}}{S_t} \right)^{-\gamma} F(s_{t+1}, n - 1) \right], \quad (12)$$

with boundary condition $F(s_t, 0) = 1$. Details can be found in Appendix C.

Besides reducing computation time, these formulas can be used to gain insight into the workings of the model. From these formulas, it follows that yields equal

$$y_t^n = -\frac{1}{n} (\ln A(n) + B(n)(z_t - g) + \ln F(s_t, n)), \quad (13)$$

while bond returns are given by

$$R_{t+1}^n = \frac{F(s_{t+1}, n - 1) A(n - 1)}{F(s_t, n) A(n)} \exp \{ B(n - 1)(z_{t+1} - g) - B(n)(z_t - g) \}. \quad (14)$$

It follows from (14) that when $\rho = 0$ and $b = 0$, $E_t(R_{t+1}^n) = R_{t+1}^f$. The reason is that, for all n , $F(s_{t+1}, n)$ is the same constant multiple of $F(s_t, n - 1)$.⁷ Even though interest rates vary, bonds have zero risk premia. While this may seem surprising, the economic intuition follows from rewriting the investor's Euler equation:

$$E(R_t^n - R_t^1) = -\text{Cov}(R_t^n - R_t^1, M_t) \frac{\sigma(M_t)}{E(M_t)}. \quad (15)$$

When $\rho = 0$ and $b = 0$, r_{t+n}^f does not covary M_{t+1} . Therefore bond returns R_{t+1}^n , which depend on future interest rates, do not covary either. Thus when $\rho = 0$ and $b = 0$ bonds are not risky from the point of view of investors.

Aggregate Wealth

In this economy, the market portfolio is equivalent to aggregate wealth, and the dividend equals aggregate consumption. The price-consumption ratio and the return on the market can be calculated using methods similar to those above, with a small but important modification.

Analogously to the previous section, let $P_{e,t}^n$ denote the price of an asset that pays the endowment C_{t+n} in n periods. The e subscript denotes equity. Because these assets

⁷When $\rho = 0$, the term multiplying $F(s_{t+1}, n - 1)$ in (12) is a constant in n . Because the riskfree rate does not depend on s_t when $b = 0$, this term must be a constant, and in fact equals $e^{-\gamma(1-\phi)/2}$.

pay no coupons, they have the same recursive pricing relation as bonds (10). Of course the prices are different, and this is because there is a different boundary condition:

$$P_{e,t}^0 = C_t.$$

Unlike the case for bonds, $P_{e,t}^n$ is not simply a function of s_t and z_t . It is a function of consumption C_t as well. To avoid introducing an additional variable into the problem, the equations for equity are rewritten in terms of price-consumption ratios, rather than simply prices.

$$\frac{P_{e,t}^n}{C_t} = E_t \left[\delta \left(\frac{S_{t+1}}{S_t} \right)^{-\gamma} \left(\frac{C_{t+1}}{C_t} \right)^{1-\gamma} \frac{P_{e,t+1}^{n-1}}{C_{t+1}} \right]. \quad (16)$$

with boundary condition $\frac{P_{e,t}^0}{C_t} = 1$. Now the problem is truly analogous to that of the previous section. Appendix C describes the functions corresponding to F , G , A and B in the previous section.

Finally, the price-consumption ratio of the market equals the sum of the price-consumption ratio on these zero-coupon securities:

$$\frac{P_t}{C_t} = \sum_{n=1}^{\infty} \frac{P_{e,t}^n}{C_t}. \quad (17)$$

This way of calculating the price-consumption ratio is equivalent to the more traditional fixed-point method used by Campbell and Cochrane (1999). In this endowment economy, (17) also represents the price-dividend ratio.

1.4 Relation to affine models

The sections above described the habit formation model in terms of consumption and preferences. A mathematically equivalent way to describe the model is in terms of the stochastic discount factor M_t . It makes sense to ask how the model for the stochastic discount factor relates to others in the literature. One of the state variables, z_t , follows a Gaussian process as in Vasicek (1977), while s_t follows a the process given for the short-rate in Cox, Ingersoll, and Ross (1985). It is instructive to compare the model to the affine class of Dai and Singleton (2000) and Duffie and Kan (1996), defined by multivariate extensions of the Vasicek and Cox-Ingersoll-Ross models.

In fact, the model is affine in z_t but not in s_t . On one level, this is apparent from the formulas for yields, (13), which shows that yields are nonlinear in s_t . This non-linearity is

not an artifact of the choice of discrete time. Rather, it is fundamental to the assumptions on consumption and preferences.

To see why, consider the equation for the price of a two-year bond. For simplicity consider the case where $z_t \equiv g$, so that there is no second factor. Then

$$\begin{aligned} P_t^2 &= \delta E_t [M_{t+1} P_{t+1}^1] \\ &= \delta \exp \{ \bar{r}^f + b\phi(\bar{s} - s_t) - \gamma(1 - \phi)(\bar{s} - s_t) - \gamma g \} \exp \{ (\gamma + \gamma\lambda(s_t) + b\lambda(s_t))^2 \}. \end{aligned}$$

Recall that

$$\lambda(s_t) = (1/\bar{S})\sqrt{1 - 2(s_t - \bar{s})} - 1.$$

Here, the volatility of the stochastic discount factor is not proportional to the volatility of the riskfree rate, so the square root does not disappear as in Cox, Ingersoll, and Ross (1985). The reason is that the riskfree rate is a linear function of s_t , while the log stochastic discount factor is a linear function of s_t and c_t . For the model to be affine, the volatility of consumption growth would have to be proportional to the volatility of s_t . This is a very strong assumption, and it seems worth accepting non-linearity to avoid it, particularly because the model remains tractable.⁸

To summarize, the model in this paper differs from affine models on several levels. First, because the model starts with preferences and consumption, it generates testable implications about the relationship of consumption growth and asset returns. These implications are investigated in the next section. Second, the model implies a different pricing kernel. If viewed as a reduced-form model for the term structure and fitted to return data, the implications will be different than those of affine models. Finally, the results of the model can be interpreted in terms of the preferences and endowments of a representative agent, and thus can be described in economic terms.

2 Estimation

The model in Section 1 suggests a relation between consumption growth and interest rates that has not previously been investigated. Surplus consumption approximates a long-run moving average of past consumption growth. Thus, the model in this paper

⁸A broader class of affine models is proposed by Duffee (2000) and investigated by Dai and Singleton (2001). In these models, the price-of-risk follows a Gaussian rather than square-root process. Because s_t is non-Gaussian the model in this paper is not nested by this broader class either.

allows long-run average consumption growth to predict the interest rate. The question is therefore whether long-run consumption growth predicts the interest rate, and, if so, with what sign.

This question is of interest for several reasons. First, this parameter is key to determining the characteristics of the term structure.⁹ Second, the relation between past consumption growth and the interest rate is of interest in its own right. Given the importance of consumption growth and the riskfree rate to asset pricing, this result is suggestive of a new restriction on asset pricing models.

Section 2.1 demonstrates that past consumption growth does indeed predict the interest rate, as conjectured in Section 1. Section 2.2 assumes the functional form given by the model. The primitive parameters are estimated using the generalized method of moments approach of Hansen and Singleton (1982). These parameter values are taken as a starting point for the simulations in Section 3.

2.1 Does long-run consumption growth predict the interest rate?

As shown in Section 1.2, the model implies that

$$r_{t+1}^f = \bar{r}^f + \gamma(z_t - g) + b(\bar{s} - s_t)$$

where $z_t = E_t(\Delta c_{t+1})$ and

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

for $s_t \approx \bar{s}$ and $z_t \approx g$. Thus the model suggests regressing a measure of the short-term interest rate on long lags of past consumption growth.¹⁰ In Section 2.2, expected future consumption growth is included in the regression using instrumental variables.

In order to investigate the connection between past consumption growth and the interest rate, returns on the 90-day Treasury Bill (from CRSP) are regressed on a weighted average of past consumption growth. Both variables are inflation-adjusted. The weighted average is constructed by taking the previous 40 quarters (ten years) of consumption

⁹It is possible to generate an interesting term structure purely from time-variation in expected consumption growth. This is the approach taken in Wachter (2000). However, the evidence suggests that surplus consumption does indeed influence the interest rate.

¹⁰Short-run lags of consumption growth do not appear to be statistically significant in predicting the interest rate.

Table 1: Predicting interest rates with past consumption

Regression of the quarterly interest rate r_{t+1}^f on an exponentially weighted average of past consumption data $\sum_1^{40} \phi^j \Delta c_{t-j}$. The regression coefficient is denoted by β . Both variables are de-meanned and standardized.

ϕ	β	tstat	R^2
0.50	-0.08	-1.03	0.01
0.60	-0.10	-1.34	0.01
0.70	-0.14	-1.80	0.02
0.80	-0.19	-2.51	0.04
0.85	-0.23	-3.02	0.05
0.90	-0.28	-3.79	0.08
0.95	-0.36	-5.02	0.13
0.99	-0.43	-6.17	0.19

growth and discounting with discount factor ϕ . Varying the length of the summation between 5 and 15 years has little impact on the pattern or significance of the results. Choosing 40 lags of consumption growth has the effect of reducing the length of the sample by 10 years, so that 1957 is the starting year, rather than 1947. Table 2 presents summary statistics for consumption growth for this sample period.

To summarize, the regression is given by

$$r_{t+1}^f = \text{constant} + \beta \sum_{j=1}^{40} \phi^j \Delta c_{t-j} + \text{error}, \quad (18)$$

where r_{t+1}^f denotes the rate of return between time t and $t + 1$. The weighted average starts with consumption in the quarter preceding the interest rate in order to avoid the effects of time-averaged consumption data. This choice has no impact on the results.

Table 1 reports the results of this regression for values of ϕ ranging from 0.5 to 0.99. The coefficient on the weighted average of consumption growth (β) is negative for all values of ϕ . Furthermore, both the significance and the R^2 rise substantially as ϕ rises. At $\phi = 0.99$, β is significant at the 0.01 level, and variation in expected consumption growth accounts for 19% of the variation in the interest rate. The higher the value of ϕ , the greater weight placed on earlier consumption growth. Table 1 shows that a *long-run* weighted average of past consumption growth is necessary to predict the interest rate.

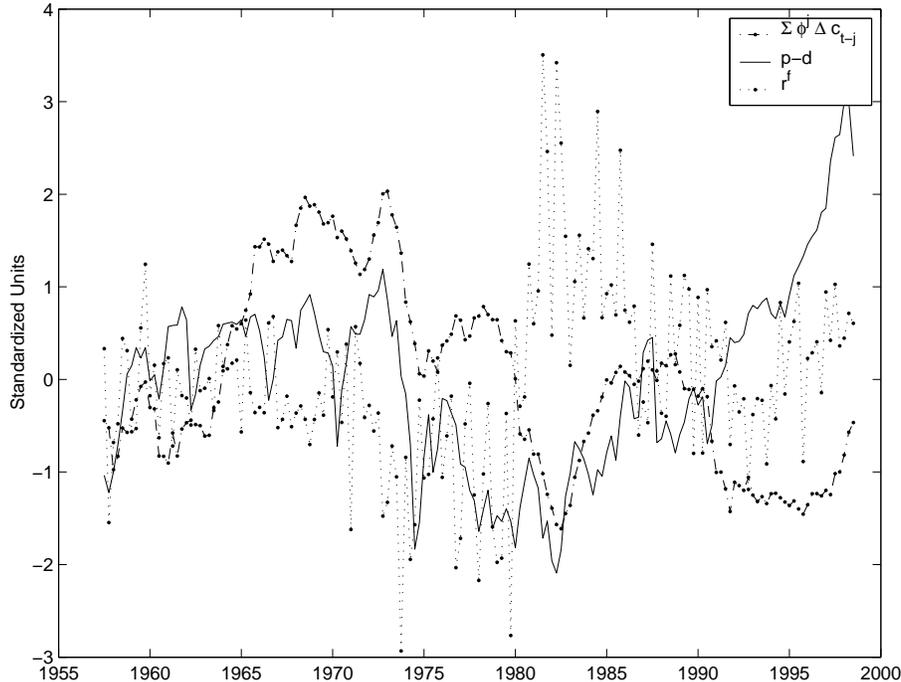


Figure 1: **Past Consumption Growth and the Interest Rate.** This figure plots the history of average past consumption growth $\sum_{j=0}^{40} \phi^j \Delta c_{t-j}$, the log price-dividend ratio, and the continuously compounded rate of return on the 90-day Treasury Bill. The parameter $\phi = 0.96$. Variables are adjusted for inflation, and de-meanned and standardized.

Figure 1 plots the history of average past consumption growth ($\sum_{j=0}^{40} \phi^j \Delta c_{t-j}$), the log price-dividend ratio (from CRSP), and the log 90-day Treasury-Bill rate. So that the persistence of past consumption growth matches that of the price-dividend ratio, $\phi = 0.96$.¹¹ The negative relationship between past consumption and the interest rate is apparent throughout the sample period. Moreover, past consumption has a surprising ability to explain both short and long-term fluctuations in the interest rate. Figure 1 also reveals that changes in average past consumption growth mirror changes in the price-dividend ratio. As will be clear from the analysis below, the relation between past consumption growth and asset prices provides direct support for the assumptions behind the habit formation model.

¹¹As is shown below, the model in this paper implies that the persistence of the weighted average of past consumption growth equals the persistence of the price-dividend ratio in the economy.

Table 2: Consumption moments

The mean, standard error on the mean, standard deviation, and the first and second-order autocorrelations for consumption growth from the period 1957-1998. The data is quarterly and adjusted for inflation. The mean and standard deviation are reported in percentage terms.

$E(\Delta c)$	SE	$\sigma(\Delta c)$	Corr1	Corr2
0.53	0.04	0.47	0.32	0.21

2.2 Generalized Method of Moments Estimation

This section estimates the primitive parameters of the model using consumption and interest-rate data. From the equations for consumption growth (1) and the riskfree rate (7), it follows that the latent variable z_t can be expressed in terms of s_t and r_{t+1}^f :

$$z_t = g + \frac{1}{\gamma} (r_{t+1}^f - \bar{r}^f) + \frac{b}{\gamma} (s_t - \bar{s}). \quad (19)$$

Substituting in for z_t in (1) produces the the system of equations:

$$\Delta c_{t+1} = g + \frac{1}{\gamma} (r_{t+1}^f - \bar{r}^f) + \frac{b}{\gamma} (s_t - \bar{s}) + v_{t+1} \quad (20)$$

$$r_{t+2}^f \approx \bar{r}^f + \psi (r_{t+1}^f - \bar{r}^f) + b(\psi - \phi)(s_t - \bar{s}) - b\lambda(\bar{s})v_{t+1} + \gamma u_{t+1}. \quad (21)$$

Equation (20) nests equations long important to empirical asset pricing. When s_t is constant (equivalently, when investors have power utility preferences), this equation reduces to the equation considered by Hansen and Singleton (1982). This, and many subsequent studies have failed to find a significant relationship between consumption growth and interest rates. The model in this paper suggests a testable modification to the traditional model, namely adding surplus consumption as a right hand side variable. Equation (21) indicates that the assumptions on the process for the latent variable z_t imply a process for the riskfree rate. The equation is written in terms of the riskfree rate at $t+1$ and $t+2$ in order to estimate the covariance matrix for the two-equation system.

Equations (20) and (21) are estimated using the generalized method of moments, with twice-lagged values of consumption growth, the 90-day Treasury Bill return, and the price-dividend ratio as instruments. As usual, instrumental variables are necessary because r_{t+1}^f is measured with noise due to inflation. Along with a vector of ones to estimate g and \bar{r}^f , the four instruments mentioned above imply a total of eight moment

Table 3: GMM estimation

Generalized method of moments estimation of (20) and (21) using quarterly inflation-adjusted consumption and interest rate data from 1957-1998. The instruments are the twice-lagged dividend price ratio, interest rate, and consumption growth. The autocorrelation of the price-dividend ratio forms an additional moment condition, which helps to identify ϕ .

Parameter	Estimate	S. E.
g , %	0.49	(0.04)
\bar{r}^f , %	0.50	(0.06)
σ_v , %	0.50	(0.05)
σ_u , %	0.13	(0.10)
b/γ	0.61×10^{-2}	(0.27×10^{-2})
γ	5.8	(4.9)
ϕ	0.956	(0.022)
ψ	0.904	(0.113)
ρ	0.355	(0.223)
J stat	5.21	($p = 0.16$)

conditions. The errors of (20) and (21) imply three additional moment conditions. The resulting moment conditions are nonlinear functions of the underlying parameters g , \bar{r}^f , b , ψ , ϕ , ρ , σ_u , and σ_v (note that \bar{S} can be expressed in terms of these parameters).¹² Because γ is difficult to estimate, b/γ is estimated, rather than b directly. This is not a problem: from (7) either b or b/γ could be interpreted as the more primitive parameter.

As in the previous section, the long-run weighted average of past consumption growth given in (18) proxies for s_t . Because the proxy for s_t is itself a function of ϕ , the system is highly nonlinear. To make the estimation more stable in ϕ , the autocorrelation of the price-dividend ratio is added as an additional moment condition. Even though the price-dividend ratio, could, in principle, be a function of all the underlying parameters in the economy, the simulation results from the next section show that its first-order autocorrelation is identically equal to ϕ . The resulting system has a total of twelve moment conditions to estimate nine parameters.

¹²For simplicity, \bar{r}^f is regarded as the underlying variable rather than δ .

The results of the GMM estimation are reported in Table 3. The results confirm those found in the previous section. First, b/γ is significantly different from zero. Economically, this means that the intertemporal substitution effect from habit dominates the precautionary savings effect, also arising from habit. Table 3 shows that ψ is relatively high and σ_u is much lower than σ_v . Because there are 12 moment conditions and 9 unknowns, the J -statistic has a $\chi^2(3)$ distribution under the null of a perfect fit. The p -value for this statistic is calculated to be 0.16, so the model is not rejected.

As is well known, estimating the system described above is equivalent to estimating the system where (20) is rearranged so that consumption growth is on the left hand side and interest rates on the right:

$$r_{t+1}^f - \bar{r}^f = -\gamma g + \gamma \Delta c_{t+1} + b(\bar{s} - s_t) - \gamma v_{t+1} \quad (22)$$

The estimation that replaces (20) with (22) is believed to be less reliable because consumption growth is not highly correlated with the twice-lagged instruments. It turns out that replacing (20) by (22) results in larger standard errors, but similar point estimates. Most importantly, the parameter b retains its significance. Even controlling for expected future consumption growth, past consumption growth predicts the interest rate.

The results demonstrate that the model provides a good fit to consumption and interest rate data. A separate question of interest is whether the process assumed for the aggregate endowment fits consumption data alone.¹³ Appendix A estimates the process given in (1) using maximum likelihood. Conditional on consumption data, the process is equivalent to an ARMA(1,1). The results in Appendix A show that the nested alternative of an AR(1), and an i.i.d. model are strongly rejected.¹⁴

¹³It would not be possible to fit the entire model to consumption data. There would be no way to estimate the coefficient b .

¹⁴The interpretation of the results in Appendix A is complicated by known difficulties with consumption data. Christiano, Eichenbaum, and Marshall (1991) shows that time-averaging could bias correlations upward. Heaton (1993) shows that durability biases correlations downward. Ferson and Harvey (1992) show that seasonal adjustment, like time-averaging, could introduce spurious positive correlation. While accounting simultaneously for all of these biases is beyond the scope of this paper, the results in Appendix A show, at the least, that the correlation cannot result from time-averaging alone.

3 Implications for Asset Returns

This section describes the implications of the model for returns on bonds and stocks. Section 3.1 characterizes the price-dividend ratio, the yield spread, and yields themselves as functions of the underlying state variables s_t and z_t . Sections 3.2-3.4 evaluate the model by simulating 100,000 quarters of returns on bonds and stocks. Besides fitting stylized facts about the aggregate market and the riskfree rate, the model is shown to produce positive risk premia on bonds, an upward sloping yield curve. Lastly, the model is shown to replicate the empirical results of Campbell and Shiller (1991), known as the expectations puzzle.

The parameter estimates from Table 3 are used as a starting point for the analysis. A difficulty with simply using the results in Table 3 is the noise in the estimate of γ . In the simulations that follow, γ is not set equal to its point estimate, but rather to the value that allows the model to match the unconditional Sharpe ratio in the data, supposing the other parameters are as stated in Table 3. This results in a value of 1.1 for γ , substantially lower than the point estimate, but well within one standard error. To get a sense of the range of possible results, the model is simulated not only at the point estimates, but at the point estimates with b plus or minus one standard error, and the point estimates with ρ plus or minus one standard error. Varying σ_u and ψ have similar effects to varying ρ , and the remaining parameters are estimated fairly precisely. Thus varying b and ρ gives a good sense of the range of possible values, as well as intuition for the source of the effects.

3.1 Characterizing the Solution

Figure 2 plots the price-consumption ratio and the yield spread on a ten-year zero-coupon bond as a function of s_t , for z_t equal to its unconditional mean g , g plus one standard deviation, and g minus one standard deviation. Figure 2 demonstrates that the price-dividend ratio is increasing in s_t . The price-dividend ratio is often taken to be a measure of the business cycle (e.g. Lettau and Ludvigson (2001)), so this confirms the intuition that s_t is a procyclical variable. Figure 2 also shows that the yield spread is a countercyclical variable, just as in the data (e.g. Fama and French (1989)). Both the price-dividend ratio and the yield spread are decreasing in expected consumption growth z_t . While shocks to expected cash flows increase the price-dividend ratio, shocks to the riskfree rate decrease it. Because $\gamma > 1$, the latter effect dominates.

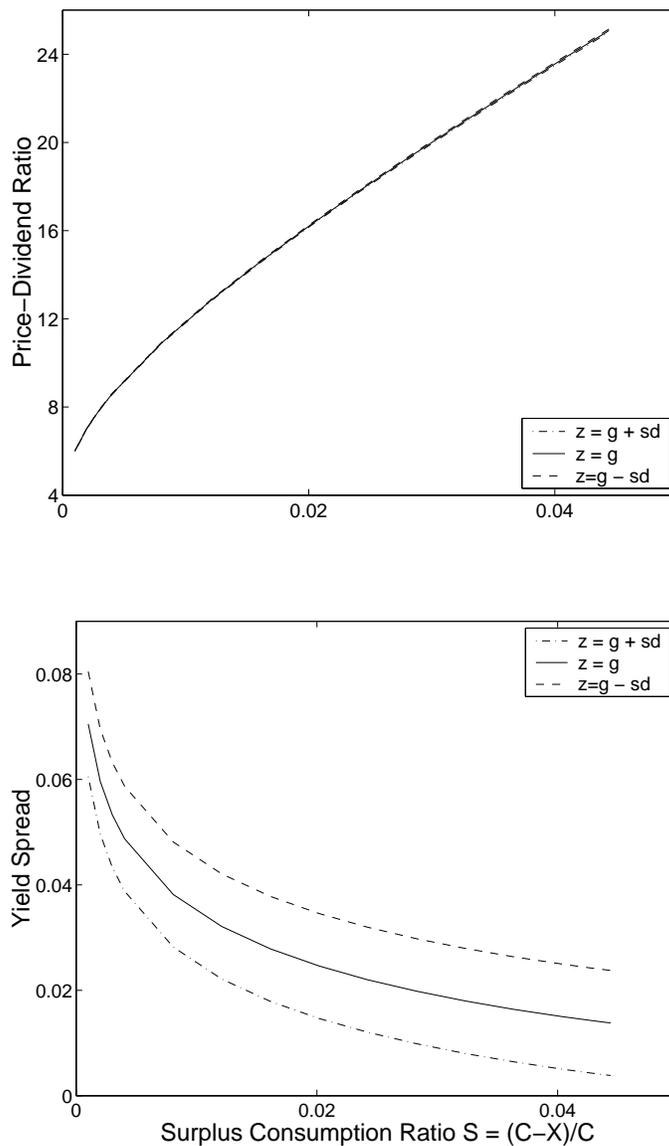


Figure 2: **The Price-Dividend Ratio and the Yield Spread.** This figure plots the price-dividend ratio and the yield spread as functions of S_t , for z_t at and one standard deviation above and below its long-run mean. The price-dividend ratio is the ratio of aggregate wealth to aggregate consumption. The yield spread is the difference between the log yield on a 10-year zero-coupon bond and on a bond maturing in one quarter. The standard deviation $sd = \frac{\sigma_u}{\sqrt{1-\psi^2}}$.

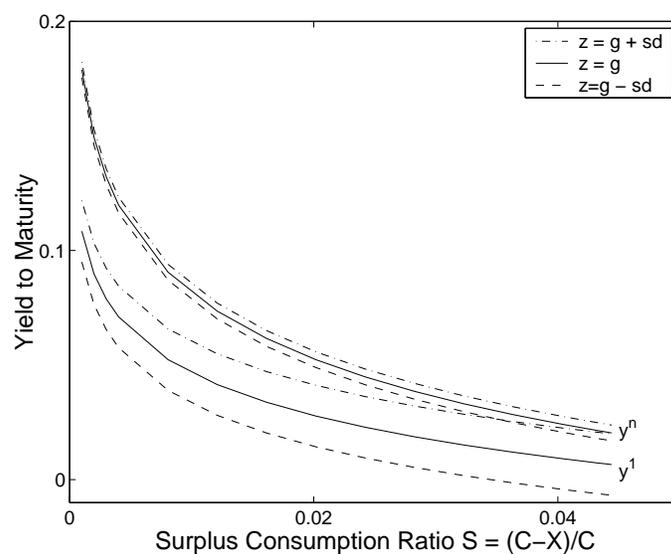


Figure 3: **Bond yields** This figure plots the yield levels as functions of S_t , for z_t at and one standard deviation above and below its long-run mean. y^n is the yield on a 10-year zero-coupon bond; y^1 is the yield on a bond maturing in one quarter. The standard deviation $sd = \frac{\sigma_u}{\sqrt{1-\psi^2}}$.

Figure 2 also demonstrates that, at the parameter values of interest, the price-dividend ratio barely depends on z_t at all. The yield spread, by contrast, depends on both z_t and s_t . Figure 3 illustrates how yield levels vary as functions of s and z . Both long and short-term yields decrease with s and increase with z , but, long yields are much less sensitive to z than short yields, while short yields are much less sensitive to s . These figures show that the price-dividend ratio and the yield spread do indeed represent two different factors in the economy. During recessions, i.e. times of low s_t , both factors depend almost exclusively on s_t , and are more highly correlated. During booms, variation in expected consumption growth z_t becomes more important.

3.2 Aggregate wealth and the riskfree rate

To evaluate the predictions of the model for asset returns, 100,000 quarters of data are simulated. Prices and returns are calculated numerically, using the method described in the previous section. The values of the parameters come from Table 3, with $\gamma = 1.1$ to match the Sharpe ratio. As explained above, the model is simulated not only at the point estimates, but at the point estimates with b plus or minus one standard error, and the point estimates with ρ plus or minus one standard error.

The implications of this model for equity returns are very similar to that of Campbell and Cochrane (1999). As Table 4 shows, 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). 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. In addition, price-dividend ratios have the ability to predict excess returns on equities, just as the data would suggest. Table 4 also shows the model matches the mean and provides a realistic standard deviation for the riskfree rate. Thus the model resolves the riskfree rate puzzle of Weil (1989). The low mean and volatility 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. The volatility of the riskfree rate is almost always lower than that found in the data, a positive feature because the interest rate in the data are measured with inflationary noise.

Table 4: Statistics for the Aggregate Market and the Riskfree Rate

Statistics for the aggregate market and the riskfree rate from simulated quarterly data. The mean and standard deviation of returns are in annualized percentages. The Sharpe ratio is the first column divided by the second. Curvature $\gamma = 1.1$. The first panel varies b plus and minus one standard error while the second panel varies ρ plus and minus one standard error. All other parameters are as in Table 3.

$b (\times 10^{-2})$	0.37	0.67	0.97	data
$E(r^m - r^f)$	5.00	5.32	5.40	4.93
$\sigma(r^m - r^f)$	15.45	17.34	18.63	16.28
Sharpe	0.32	0.31	0.29	0.30
$E(P/D)$	20.78	20.00	20.05	30.92
$\sigma(p - d)$	0.24	0.27	0.30	0.25
$\text{Corr}(p - d)$	0.95	0.95	0.95	0.95
$E(r^f)$	2.02	1.99	1.97	1.96
$\sigma(r^f)$	0.72	0.90	1.18	1.49
ρ	0.13	0.35	0.57	data
$E(r^m - r^f)$	5.32	5.32	5.31	4.93
$\sigma(r^m - r^f)$	17.47	17.34	17.36	16.28
Sharpe	0.30	0.31	0.31	0.30
$E(P/D)$	19.93	20.00	19.96	30.92
$\sigma(p - d)$	0.27	0.27	0.28	0.25
$\text{Corr}(p - d)$	0.95	0.95	0.96	0.95
$E(r^f)$	2.01	1.99	2.04	1.96
$\sigma(r^f)$	1.00	0.90	0.82	1.49

Given that the consequences for equity returns are so similar to those of Campbell and Cochrane (1999), the sections that follow focus on the properties of bond returns. These sections demonstrate the model’s ability to explain features of the bond data.

3.3 Risk Premia on Bonds

Tables 5 and 6 demonstrate that bonds have positive risk premia and that the average yield curve is upward sloping. For all parameter values, expected excess returns and yield spreads are positive and increase with maturity.

The magnitudes of the unconditional moments vary widely across parameter values, despite the fact that the parameters are only adjusted by a standard error. Recall from Section 1.2 that the riskfree rate is given by

$$r_{t+1}^f = \bar{r}^f + \gamma(z_t - g) + b(\bar{s} - s_t).$$

As Table 5 shows, the higher the value of b , the greater the slope of the yield curve, and the higher expected returns. Increasing b also increases the variance of returns, though the variance of yields remains roughly constant. Increasing ρ , the correlation between shocks to z_t and to Δc_t , has the opposite effect. Higher values of ρ lead to lower expected excess returns and lower yield spreads. Comparisons with bond moments from the data can be misleading because the bonds in this model are real, while those in the data are nominal. An informal comparison with the data (e.g. the nominal bond moments in Campbell and Viceira (2001)) suggests that expected returns and yields increase faster than the analogous moments from the data. However, this effect can be mitigated by increasing ρ and decreasing b .

The mechanism behind the effects of changing the parameter values can be understood in terms of the covariance form of the investor’s Euler equation:

$$E(R_t^n - R_t^1) = -\text{Cov}(R_t^n - R_t^1, M_t) \frac{\sigma(M_t)}{E(M_t)}, \quad (23)$$

where R_{t+1}^n is the return on a bond maturing in n periods. Increasing b means that interest rates covary more with s_t . Because bond returns move in the opposite direction to interest rates, the higher b , the greater the positive covariance of bonds with s_t . This means that bonds have high returns in good times and poor returns in bad. Investors demand more of a risk premium to hold them. On the other hand, raising ρ increases the hedging feature of bonds. The higher ρ , the greater the covariance of interest rates with

Table 5: Excess Bond Returns

Means and standard deviations for excess holding period returns on bonds in simulated data. Statistics are in annualized percentages; maturity is reported in quarters. Curvature $\gamma = 1.1$. The first panel varies b plus and minus one standard error, while the second panel varies ρ plus and minus one standard error. All other parameters are as in Table 3.

	$b = 0.0037$		$b = 0.0067$		$b = 0.0097$	
mat	$E(r-r^f)$	$\sigma(r-r^f)$	$E(r-r^f)$	$\sigma(r-r^f)$	$E(r-r^f)$	$\sigma(r-r^f)$
4	0.08	0.78	0.20	0.92	0.30	1.11
12	0.47	2.29	0.93	3.11	1.28	4.00
20	1.07	3.72	1.87	5.59	2.41	7.26
28	1.87	5.61	2.94	8.61	3.56	10.95
40	3.35	9.59	4.60	13.99	5.12	16.83

	$\rho = 0.13$		$\rho = 0.35$		$\rho = 0.57$	
mat	$E(r-r^f)$	$\sigma(r-r^f)$	$E(r-r^f)$	$\sigma(r-r^f)$	$E(r-r^f)$	$\sigma(r-r^f)$
4	0.26	1.03	0.20	0.92	0.15	0.79
12	1.10	3.51	0.93	3.11	0.77	2.72
20	2.09	6.22	1.87	5.59	1.66	5.02
28	3.19	9.41	2.94	8.61	2.71	7.93
40	4.84	14.95	4.60	13.99	4.38	13.25

consumption growth, and the lower the covariance of bond returns with consumption growth. This makes bonds less risky than they would be in a model driven by s_t alone, and decreases the risk premium.

The previous discussion shows that interest rate risk leads bonds to have a positive risk premium. Because of this positive risk premium, there is a feedback effect that further raises the risk, and therefore, the premium on bonds. As shown below, risk premia on bonds vary. Variation in the risk premium itself induces price fluctuations, much like “excess volatility” in the stock market. This excess volatility makes expected returns on bonds larger than they otherwise would be.

This second effect helps in understanding why bonds command risk premia at all. After all, these bonds pay off a fixed amount. Why is it that investors simply do not wait

Table 6: Bond Yield Spreads

Means and standard deviations for the yield spread on long-term bonds in simulated data. Statistics are in annualized percentages; maturity is reported in quarters. Curvature $\gamma = 1.1$. The first panel varies b plus and minus one standard error while the second panel varies ρ plus and minus one standard error. All other parameters are as in Table 3.

mat	$b = 0.0037$		$b = 0.0067$		$b = 0.0097$	
	$E(y^n - y^1)$	$\sigma(y^n - y^1)$	$E(y^n - y^1)$	$\sigma(y^n - y^1)$	$E(y^n - y^1)$	$\sigma(y^n - y^1)$
4	0.04	0.09	0.10	0.10	0.15	0.11
12	0.20	0.28	0.43	0.30	0.61	0.32
20	0.44	0.42	0.83	0.46	1.13	0.47
28	0.74	0.53	1.30	0.58	1.68	0.59
40	1.31	0.69	2.06	0.74	2.51	0.70

mat	$\rho = 0.13$		$\rho = 0.35$		$\rho = 0.57$	
	$E(y^n - y^1)$	$\sigma(y^n - y^1)$	$E(y^n - y^1)$	$\sigma(y^n - y^1)$	$E(y^n - y^1)$	$\sigma(y^n - y^1)$
4	0.13	0.10	0.10	0.10	0.07	0.10
12	0.52	0.30	0.43	0.30	0.34	0.30
20	0.97	0.45	0.83	0.46	0.70	0.46
28	1.47	0.57	1.30	0.58	1.14	0.60
40	2.25	0.71	2.06	0.74	1.89	0.77

until maturity to sell the bond, when the return is fixed? Consistent with this intuition, the power utility model of Backus, Gregory, and Zin (1989) implies that bonds have negative excess returns that are very small in magnitude. In this model, by contrast, bonds are risky because their prices fall during periods of low surplus consumption, namely during recessions. These are the times when investor's marginal utility is the highest, and when, as a result, they most want to increase their consumption. Long-term bonds command a premium not only because of their dependence on the time-varying riskfree rate, but because they do badly in recessions.

Table 7: Campbell-Shiller Long-Rate Regressions

Each column reports the coefficient β_n from the regression

$$y_{t+1}^{n-1} - y_t^n = \alpha_n + \beta_n \frac{1}{n-1} (y_t^n - y_t^1) + \text{error}$$

using simulated quarterly data. Curvature $\gamma = 1.1$. The first panel varies b plus and minus one standard error while the second panel varies ρ plus and minus one standard error. All other parameters are as in Table 3.

mat	$b = 0.0037$	$b = 0.0067$	$b = 0.0097$
4	0.70	0.26	-0.01
12	0.39	-0.23	-0.55
20	-0.04	-0.79	-1.11
28	-0.56	-1.36	-1.61
40	-1.40	-2.09	-2.20

mat	$\rho = 0.13$	$\rho = 0.35$	$\rho = 0.57$
4	0.28	0.26	0.22
12	-0.21	-0.23	-0.29
20	-0.79	-0.79	-0.85
28	-1.37	-1.36	-1.39
40	-2.14	-2.09	-2.07

Table 8: Campbell-Shiller Short-Rate Regressions

Each column reports the coefficient β_n from the regression

$$\sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right) \Delta y_{t+i}^1 = \alpha_n + \beta_n (y_t^n - y_t^1) + \text{error}$$

using simulated quarterly data. Curvature $\gamma = 1.1$. The first panel varies b plus and minus one standard error while the second panel varies ρ plus and minus one standard error. All other parameters are as in Table 3.

mat	$b = 0.0037$	$b = 0.0067$	$b = 0.0097$
4	0.86	0.65	0.51
12	0.76	0.50	0.37
20	0.64	0.36	0.23
28	0.51	0.24	0.12
40	0.35	0.12	0.02

mat	$\rho = 0.13$	$\rho = 0.35$	$\rho = 0.57$
4	0.65	0.65	0.63
12	0.51	0.50	0.48
20	0.37	0.36	0.34
28	0.25	0.24	0.23
40	0.12	0.12	0.11

3.4 Time-Variation in Risk Premia

The previous section pointed to time-variation in the risk premium as a source of variation for long-bond prices. This section shows that risk premia are indeed time-varying, and explains why. Table 7 shows the outcome of regressions

$$y_{t+1}^{n-1} - y_t^n = \text{constant} + \beta_n \frac{1}{n-1} (y_t^n - y_t^1) + \text{error}$$

These “long-rate” regressions are performed by Campbell and Shiller (1991), to test the hypothesis of constant risk premia on bonds, also known as the generalized expectations hypothesis. If risk premia are constant, β should be equal to one. Instead, Campbell and Shiller find a coefficient that is negative at all maturities, and significantly different from one. Moreover, the higher the maturity, the lower β_n .

Table 7 demonstrates that this model replicates the result of Campbell and Shiller (1991) for the entire range of parameter values. At all maturities, the coefficient on the yield spread is less than one, and for maturities greater than one year the coefficient is negative. The coefficient decreases with maturity, just as in the data.

Campbell and Shiller also perform the following “short-rate” regression:

$$\sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right) \Delta y_{t+i}^1 = \text{constant} + \beta_n (y_t^n - y_t^1) + \text{error}$$

The expectations hypothesis predicts that $\beta_n = 1$. Campbell and Shiller find that β_n is positive but significantly less than one. Table 8 shows that the model also replicates this result for all parameter values.

Why drives the model’s success at replicating these empirical findings? The condition $\beta \neq 1$ in both regressions is equivalent to the statement that excess returns on long-term bonds are predictable. It follows from the definition of yields and returns that

$$r_{t+1}^n = y_t^n - (n-1)(y_{t+1}^{n-1} - y_t^n)$$

Rearranging, and taking expectations:

$$E_t (y_{t+1}^{n-1} - y_t^n) = \frac{1}{n-1} (y_t^n - y_t^1) - \frac{1}{n-1} E_t (r_{t+1}^n - y_t^1)$$

Thus the coefficient of a regression of changes in yields on the scaled yield spread produces a coefficient of one only if risk premia on bonds are constant. A similar argument shows that

$$E_t \left(\sum_{i=1}^{n-1} \left(1 - \frac{i}{n}\right) \Delta y_{t+i}^1 \right) = (y_t^n - y_t^1) - \frac{1}{n} \sum_{i=1}^{n-1} E_t (r_{t+i}^{n-i+1} - r_{t+i}^1)$$

Thus a regression of changes in future short-term rates on the yield spread only produces a coefficient of one if risk premia are constant.

In this model, risk premia are not constant. During recessions, the volatility of investor's marginal utility rises, as shown in (9). In Campbell and Cochrane (1999), this mechanism produces a time-varying risk premium on the aggregate market. Here, the same mechanism produces time-varying risk premia on bonds, provided, of course, that bonds are risky assets in the first place. The empirical work in Section 2, combined with the model in Section 1 shows why this is the case.

4 Conclusion

This paper offers an explanation for term structure anomalies based on the preferences of a representative agent. By generalizing a model already known to fit stylized facts about the aggregate stock market, that of Campbell and Cochrane (1999), this paper is able to parsimoniously model both bond and stock returns.

Relative to the model of Campbell and Cochrane (1999), this model has two innovations. The first is a new preference parameter that controls how surplus consumption influences the riskfree rate. In economic terms, this parameter represents a tradeoff between the intertemporal substitution effect and the precautionary savings effect. For lack of evidence on this parameter, Campbell and Cochrane set it to be zero. This paper provides direct evidence that the intertemporal substitution effect dominates. High surplus consumption is shown to imply lower short-term interest rates in post-war quarterly data. That is, long-run past consumption growth is shown to predict the interest rate with a negative sign. This finding is interesting both in the context of the habit formation model and more broadly. In principle it has implications for any model that speaks to the relationship between consumption growth and the riskfree rate.

The second innovation is to introduce a predictable component into consumption growth. Within the context of this model, the consequences for the aggregate market are slight. However, even a small amount of predictability has large implications for bond returns. Interestingly, predictable consumption growth produces lower risk premia on bonds, and less predictability in bond returns than otherwise. For the parameters estimated from the data, bonds provide a hedge against changes in the mean of consumption growth.

Although the model is fitted to consumption and Treasury Bill data, it is able to

replicate empirical findings on long-term bonds. In particular, the model is able to reproduce both the sign and the magnitude of Campbell-Shiller short-rate *and* long-rate regressions. That is, yield spreads predict changes in future short-term rates with a positive sign and a coefficient well below one. Yield spreads predict changes in long-term yields with a negative sign. The coefficient decreases with the maturity. In addition, the model produces positive risk premia on bonds and an upward-sloping yield curve.

The model offers an economic explanation for these term structure anomalies. The riskfree rate depends on surplus consumption with a negative sign. This implies that the short-term rate rises and bond prices fall after periods of unusually low consumption growth. Thus, bonds carry risk premia because their prices decline during “recessions”. Furthermore, the risk aversion of the representative agent varies through time. Thus, risk premia on bonds vary. This implies that the expectations hypothesis is violated in simulated data.

This paper has taken the approach of fitting the model to consumption and short-rate data, and testing the implications for long-term bonds and stocks. A natural extension would be to fit the model to long-term bonds and stocks, and to determine the implications for the riskfree rate and aggregate consumption. This approach would mirror that of the reduced-form term structure literature, with the important modification that this model implies additional testable restrictions on consumption growth. Furthermore, the model can easily accommodate additional factors, such as a process for inflation, or a more complicated endowment process. Thus the model has the potential to unify stock and bond pricing, and to connect 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, A. Gregory, and Stanley Zin, 1989, Risk premiums in the term structure: Evidence from artificial economies, *Journal of Monetary Economics* 24, 371–399.
- Bansal, Ravi, and Amir Yaron, 2000, Risks for the long run: A potential resolution of asset pricing puzzles, NBER Working Paper #8059.
- Bekaert, Geert, and Steven R. Grenadier, 2001, Stock and bond pricing in an affine economy, Working Paper, Columbia University.
- Bekaert, Geert, and Robert J. Hodrick, 2001, Expectations hypothesis tests, *Journal of Finance* 56, 1357–1394.
- Bekaert, Geert, Robert J. Hodrick, and David A. Marshall, 2001, Peso problem explanations for term structure anomalies, *Journal of Monetary Economics* 48, 241–270.
- Brandt, Michael W., and Kevin Q. Wang, 2001, Time-Varying Risk Aversion and Unexpected Inflation, Working Paper, The Wharton School.
- Campbell, John Y., 1999, Asset prices, consumption, and the business cycle, in J.B. Taylor, and M. Woodford, eds.: *Handbook of Macroeconomics, Volume I* (North-Holland, Amsterdam).
- 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, 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, 2001, Who should buy long-term bonds?, *American Economic Review* 91, 99–127.
- Cecchetti, S.G., P.S. Lam, and N.C. Mark, 1990, Mean reversion in equilibrium asset prices, *American Economic Review* 80, 398–418.

- Cecchetti, S.G., P.S. Lam, and N.C. Mark, 1993, The equity premium and the risk-free rate: Matching the moments, *Journal of Monetary Economics* 31, 21–45.
- Chapman, David A., 1998, Habit formation and aggregate consumption, *Econometrica* 66, 1223–1230.
- Christiano, Lawrence J., Martin Eichenbaum, and David Marshall, 1991, The Permanent Income Hypothesis Revisited, *Econometrica* 59, 397–423.
- Constantinides, George M., 1990, Habit formation: A resolution of the equity premium puzzle, *Journal of Political Economy* 98, 519–543.
- Cox, John C., Jonathan C. Ingersoll, and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–408.
- Dai, Qiang, 2000, From equity premium puzzle to expectations puzzle: A general equilibrium production economy with stochastic habit formation, Working Paper, New York University.
- Dai, Qiang, and Kenneth Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.
- Dai, Qiang, and Kenneth Singleton, 2001, Expectations puzzles, time-varying risk premia, and affine models of the term structure, Forthcoming, *Journal of Financial Economics*.
- Duffee, Gregory R., 2000, Term premia and interest rate forecasts in affine models, Working paper, Stanford University.
- Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379–406.
- 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.
- 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 Campbell R. Harvey, 1992, Seasonality and consumption-based asset pricing, *Journal of Finance* 47, 511–552.
- Hansen, Lars Peter, and Ken Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.
- 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.
- Heaton, John C., 1993, The interaction between time-nonseparable preferences and time aggregation, *Econometrica* 61, 353–385.
- 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 Ludvigson, 2001, Consumption, aggregate wealth and expected stock returns, *Journal of Finance* 56, 815–850.
- Li, Yuming, 2001, Expected returns and habit persistence, *Review of Financial Studies* 14, 861–899.
- Mehra, Rajnish, and Edward Prescott, 1985, The equity premium puzzle, *Journal of Monetary Economics* 15, 145–161.
- Menzly, Lior, Tano Santos, and Pietro Veronesi, 2001, Habit Formation and the Cross-Section of Stock Returns, University of Chicago.
- Shiller, Robert J., 1981, Do stock prices move too much to be justified by subsequent changes in dividends?, *American Economic Review* 71, 421–436.
- Stambaugh, Robert, 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.
- 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.
- Wachter, Jessica A., 2000, Essays in Financial Economics, Ph.D. thesis Harvard University.
- Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 402–421.
- Whitelaw, Robert F., 2000, Stock market risk and return: an equilibrium approach, *Review of Financial Studies* 13, 521–547.

A Maximum likelihood estimation of (1)

This section describes the results of estimating the process for consumption (1) from consumption data alone, using the method of maximum likelihood. The following proposition gives the exact likelihood function:

Proposition A.1 *Let*

$$\begin{aligned}\eta &= \rho\sigma_u\sigma_v - \psi\sigma_v^2 \\ \nu &= \sigma_u^2 + \sigma_v^2(1 - \psi^2)\end{aligned}$$

Then

$$\Delta c_{t+1} | \Delta c_t, \dots, \Delta c_0 \sim N(\hat{z}_t, \hat{\sigma}_t^2) \quad (24)$$

where

$$\hat{z}_t = (1 - \psi)g + \psi\Delta c_t + \frac{\eta}{\hat{\sigma}_{t-1}^2}(\Delta c_t - \hat{z}_{t-1}) \quad (25)$$

$$\hat{\sigma}_t^2 = \nu - 2\psi\eta - \frac{\eta^2}{\hat{\sigma}_{t-1}^2} \quad (26)$$

and

$$\hat{z}_0 = g \quad (27)$$

$$\hat{\sigma}_0^2 = \frac{\nu}{1 - \psi^2}. \quad (28)$$

The proof is given at the end of Appendix A.

The form of the likelihood function implies that (1) cannot be identified by consumption data alone. The parameters g and ϕ can be estimated using maximum likelihood, but ρ , σ_u , and σ_v can only be identified in the combinations $\rho\sigma_u\sigma_v - \psi\sigma_v^2$ and $\sigma_v^2 + \frac{\sigma_u^2}{1 - \psi^2}$. A second implication is that (1) has the same likelihood function as an ARMA(1,1) process when only consumption data is observable.¹⁵ Setting $\rho\sigma_u\sigma_v - \psi\sigma_v^2 = 0$ results in the

¹⁵To be precise, suppose

$$\Delta c_{t+1} = (1 - \psi)g + \psi\Delta c_t + \theta\epsilon_t + \epsilon_{t+1}$$

Then θ and $\text{Var}(\epsilon)$ are related to the parameters above via the equations

$$\begin{aligned}\sigma_v^2(1 - \psi^2) + \sigma_u^2 &= ((\theta + \psi)^2 + 1 - \psi^2) \text{Var}(\epsilon) \\ \theta\text{Var}(\epsilon) &= \rho\sigma_u\sigma_v - \psi\sigma_v^2.\end{aligned}$$

Table 9: Maximum likelihood estimation

Maximum likelihood estimates for the system (1) with $\rho = 0.35$. Quasi-maximum likelihood standard errors are in parenthesis. LR(1) is the likelihood ratio test statistic against the AR(1) model; LR(2) is the test statistic against the random walk.

Parameter	Estimate	SE
$g, \%$	0.51	(0.07)
ψ	0.72	(0.11)
$\sigma_u, \%$	0.19	(0.05)
$\sigma_v, \%$	0.40	(0.03)
LR(1)	4.72	($p < 0.02$)
LR(2)	24.75	($p < 0.001$)
$\text{Var}(\hat{z})/\text{Var}(\Delta c)$	0.14	

AR(1) process assumed by Mehra and Prescott (1985). Further setting $\psi = 0$ results in the random walk model of consumption assumed by Campbell and Cochrane (1999).

Table 9 describes the results of maximum likelihood estimation of system (1). So that the parameter values can be compared to those in Section ?, ρ is set equal to 0.35, and the consumption data is from 1957-1998. The first four entries are estimates of the parameters, with asymptotic standard errors in parentheses. The second panel of Table 9 reports likelihood ratio statistics for tests of (1) against the AR(1) model and the random walk model. The AR(1) model, which implies one restriction, is rejected below the 2% level. The random walk model, which implies two restrictions, is rejected below the 0.1% level. Clearly the predictability in consumption growth is statistically significant. Moreover, the predictability does not solely result from the first lag of consumption growth.

The next question to ask is whether the amount of predictability is economically significant. The best forecast for Δc_{t+1} based on past consumption data is

$$\hat{z}_t = (1 - \psi)g + \psi \Delta c_t + \left(\frac{\rho \sigma_u \sigma_v - \psi \sigma_v^2}{\sigma_{t-1}^2} \right) (\Delta c_t - \hat{z}_{t-1}). \quad (29)$$

Table 9 reports that variation in \hat{z} can explain 14% of the variation in total consumption growth. The predictability in consumption growth is both statistically and economically

significant.

Proof of Proposition A.I:

PROOF: It is sufficient to show, by induction, that

$$z_t | \Delta c_t, \dots, \Delta c_0 \sim N(\hat{z}_t, \hat{\sigma}_t^2 - \sigma_v^2) \quad (30)$$

Given (30), the Proposition follows from the independence of v_{t+1} and z_t .

Equations (27) and (28) represent the steady-state mean and variance. This proves the case for $t = 0$. Now assume by way of induction that

$$z_{t-1} | \Delta c_{t-1}, \dots, \Delta c_0 \sim N(\hat{z}_{t-1}, \hat{\sigma}_{t-1}^2 - \sigma_v^2) \quad (31)$$

holds. Because the distribution of $(v_t \ u_t)'$ is independent of all $t - 1$ variables,

$$\begin{bmatrix} z_{t-1} - \hat{z}_{t-1} \\ v_t \\ u_t \end{bmatrix} | \Delta c_{t-1}, \dots, \Delta c_0 \sim N \left(0, \begin{bmatrix} \hat{\sigma}_{t-1}^2 - \sigma_v^2 & 0 & 0 \\ 0 & \sigma_v^2 & \rho\sigma_u\sigma_v \\ 0 & \rho\sigma_u\sigma_v & \sigma_u^2 \end{bmatrix} \right)$$

Because

$$\begin{aligned} \Delta c_t &= \hat{z}_{t-1} + (z_{t-1} - \hat{z}_{t-1}) + v_t \\ z_t &= (1 - \psi)g + \psi\hat{z}_{t-1} + \psi(z_{t-1} - \hat{z}_{t-1}) + u_t, \end{aligned}$$

it follows that

$$\begin{bmatrix} \Delta c_t \\ z_t \end{bmatrix} | \Delta c_{t-1}, \dots, \Delta c_0 \sim N \left(\begin{bmatrix} \hat{z}_{t-1} \\ (1 - \psi)g + \psi\hat{z}_{t-1} \end{bmatrix}, \begin{bmatrix} \hat{\sigma}_{t-1}^2 & \psi\hat{\sigma}_{t-1}^2 + \eta \\ \psi\hat{\sigma}_{t-1}^2 + \eta & \psi^2(\hat{\sigma}_{t-1}^2 - \sigma_v^2) + \sigma_u^2 \end{bmatrix} \right).$$

By the properties of the conditional normal distribution:

$$z_t | \Delta c_t, \dots, \Delta c_0 \sim N \left((1 - \psi)g + \psi\hat{z}_{t-1} + \beta_t(\Delta c_t - \hat{z}_{t-1}), \psi^2(\hat{\sigma}_{t-1}^2 - \sigma_v^2) + \sigma_u^2 - \frac{(\psi\hat{\sigma}_{t-1}^2 + \eta)^2}{\hat{\sigma}_{t-1}^2} \right) \quad (32)$$

where

$$\beta_t = \frac{\psi\hat{\sigma}_{t-1}^2 + \eta}{\hat{\sigma}_{t-1}^2}.$$

Equation (30) follows from rearranging the terms in (32) and comparing with (25) and (26). \square

B. Deriving the sensitivity function $\lambda(s_t)$

The sensitivity function $\lambda(s_t)$ is reverse-engineered to produce a riskfree rate that is linear in s_t , as in (7). Setting this equation equal to the expression (6) produces the following general form for λ .

$$\lambda(s_t) = \frac{\sqrt{2}}{\gamma\sigma} \left(-\ln \delta + \gamma g + \gamma(1 - \phi)(\bar{s} - s_t) - b(s_t - \bar{s}) - \bar{r}^f \right)^{\frac{1}{2}} - 1. \quad (33)$$

Campbell and Cochrane (1999) further impose the conditions

$$\lambda(\bar{s}) = \frac{1}{\bar{S}} - 1 \quad (34)$$

$$\lambda'(\bar{s}) = -\frac{1}{\bar{S}} \quad (35)$$

They show that these conditions are equivalent to requiring that for $s_t \approx \bar{s}$, x_t is approximately a deterministic function of past consumption. Their reasoning still holds in this model, along as z_t is near g . Equations (33) - (35) lead to the expressions for \bar{r}^f and \bar{S} that are given in the text.

C. Solving for Prices

Bonds and stocks can be put into a more general pricing framework by thinking of them as levered claims to consumption C_t^θ . Bonds correspond to the case where $\theta = 0$, while stocks correspond to the case of $\theta = 1$. From the investor's Euler equation, it follows that prices obey the recursion

$$\frac{P_t^n}{C_t^\theta} = E_t \left[\left(\frac{S_{t+1}}{S_t} \right)^{-\gamma} \left(\frac{C_{t+1}}{C_t} \right)^{\theta-\gamma} \frac{P_{t+1}^{n-1}}{C_{t+1}^\theta} \right] \quad (36)$$

with boundary condition

$$\frac{P_t^0}{C_t^\theta} = 1. \quad (37)$$

By way of induction, suppose that

$$\frac{P_{t+1}^{n-1}}{C_{t+1}^\theta} = A(n-1) \exp \{ B(n-1)(z_{t+1} - g) \} F(s_{t+1}, n-1)$$

Applying (36) and moving constants outside the expectation, it follows that

$$\frac{P_t^n}{C_t^\theta} = A(n-1)\delta \exp\{(\theta - \gamma)z_t + B(n-1)\psi(z_t - g)\} \times \\ E_t \left[\exp\{(\theta - \gamma)v_{t+1} + B(n-1)u_{t+1}\} \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} F(s_{t+1}, n-1) \right]. \quad (38)$$

By conditioning on v_{t+1} and using the law of iterated expectations, u_{t+1} may be integrated out of the expectation. It follows from the joint distribution of u_{t+1} and v_{t+1} that

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

Therefore

$$\frac{P_t^n}{C_t^\theta} = \delta A(n-1) \exp\left\{(\theta - \gamma)z_t + B(n-1)\psi(z_t - g) + \frac{\sigma_u^2}{2}(1-\rho^2)B(n-1)^2\right\} \\ \times E_t \left[\exp\left\{(\theta - \gamma)v_{t+1} + \frac{\sigma_u}{\sigma_v}\rho B(n-1)v_{t+1}\right\} \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} F(s_{t+1}, n-1) \right]. \quad (39)$$

Comparing (39) with (11) implies that

$$F(s_t, n) = E_t \left[\exp\left\{\left[\frac{\sigma_u}{\sigma_v}\rho B(n-1) + (\theta - \gamma)\right]v_{t+1}\right\} \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} F(s_{t+1}, n-1) \right] \quad (40)$$

and

$$B(n) = (\theta - \gamma) + \psi B(n-1) \\ A(n) = \delta A(n-1) \exp\left\{(\theta - \gamma)g + \frac{\sigma_u^2}{2}(1-\rho^2)B(n-1)^2\right\}.$$

The expressions for A and B are obtained by solving these equations backwards, starting from $A(0) = 1$ and $B(0) = 0$.