Time-Varying Risk Aversion, Unexpected Inflation, and the Dynamics of the Term Structure*

Michael W. Brandt

The Wharton School
University of Pennsylvania
2300 Steinberg Hall-Dietrich Hall
Philadelphia, PA 19104-6367

Email: *brandtm@wharton.upenn.edu*

and NBER

Kevin Q. Wang

Rotman School of Management
University of Toronto
105 St. George Street
Toronto, Ontario M5S 3E6

Email: *kwang@mgmt.utoronto.ca*

July 1999[†]

* We thank John Cochrane, George Constantinides, Kent Daniel, and Rajnish Mehra, as well as seminar participants at the University of Chicago and the Third New York Area Macroeconomics Workshop at New York University for helpful comments and suggestions. We are solely responsible for any errors.

[†] The latest revision of this paper is available at: *http://brandt.wharton.upenn.edu*

Time-Varying Risk Aversion, Unexpected Inflation, and the Dynamics of the Term Structure

We formulate and empirically evaluate a consumption based asset pricing model in which aggregate risk aversion is time-varying, not only in response to news about aggregate consumption growth (as in a habit formation model), but also in response to news about inflation. We estimate our model and explore its pricing implications on the term structure of nominal interest rates, always comparing the results to those for the standard models with and without habit formation. The empirical results unambiguously support our hypothesis that aggregate risk aversion varies in response to news about inflation. The induced time-variation in aggregate risk aversion is both economically and statistically significant. Our model captures the stylized features of the nominal term structure and even explains the apparent rejection of the expectations hypothesis.

1 Introduction

There is now ample evidence that expected and unexpected inflation affect real asset prices.[1] Economic models designed to capture this empirical fact typically rely on nominal frictions [e.g. Lucas (1990), Christiano and Eichenbaum (1992,1995), or King and Watson (1996)], on money illusions [e.g. Shafir, Diamond, and Tversky (1997)], or just assume that consumers derive utility from holding nominal money balances [e.g. Farmer (1997)]. In this paper, we provide an alternative modeling approach. We formulate and empirically evaluate a consumption based asset pricing model in which aggregate risk aversion is time-varying, not only in response to news about aggregate consumption growth (as in a habit formation model), but also in response to news about inflation. We interpret our model either as capturing cyclical time-variations in aggregate preferences or as an approximate reduced form for a more elaborate model with nominal frictions, money illusions, or money preferences.

Whether it is rational or not, consumers dislike inflation. Shiller's (1996) survey shows that non-economists worry about the adverse real effects of excessive inflation. For instance, the majority of US non-economist respondents to the survey fully agree with the statement:

> When I see projections about how many times more a college education will cost, or how many times more the cost of living will be in coming decades, I feel a sense of *uneasiness*; these inflation projections make me *worry* that my own income will not rise as much as such costs will (emphasis added).

The survey further reveals that consumers worry about sticky wages, the correlation of inflation with some unspecified real factors, rigidities in the nominal tax system, and even the effect of inflation on social morale. Regardless of why consumers dislike inflation, the fact that they become anxious when inflation increases unexpectedly, even if their real endowments are unaffected, is well captured by a temporary increase in aggregate risk aversion.

The idea of time-varying risk aversion is well accepted in the literature, although it enjoys popularity under the disguise of habit formation [e.g. Sundaresan (1989), Abel (1990,1999), Constantinides (1990), Heaton (1995), or Campbell and Cochrane (1999)]. Consumption based models with habit formation are popular because of their success in at least partially explaining Mehra and Prescott's (1985) equity premium puzzle and Weil's (1989) risk-free rate puzzle. In a habit formation model, the representative agent's risk aversion varies

[1] For instance, Campbell and Shiller (1996) and Barr and Campbell (1997) document a strong negative correlation between unexpected inflation and real short-term interest rates using data on indexed and nominal government bonds. Pennacchi (1991) and Campbell and Ammer (1993) find the same correlation using very different research methodologies and data. In addition, Amihud (1996) and Campbell and Shiller (1996) extend the negative correlation evidence to real stock returns.

with some measure of the distance between aggregate consumption and the agent's habit. This raises the correlation between marginal utility and asset returns, while the correlation between consumption and returns remains low. We extend the notion of time-varying risk aversion to explain the observed correlation between unexpected inflation and asset prices.

The setup of our model resembles closely that of Campbell and Cochrane's (1999) habit formation model. However, rather than specifying the agent's habit formation process, as they do, we model the dynamics of aggregate risk aversion. Although the two approaches turn out to be mathematically equivalent, they have somewhat different interpretations. In addition, Campbell and Cochrane's approach is more restrictive than ours. Habit formation models only allow aggregate preferences to change deterministically or in response to news about aggregate consumption growth. By directly modeling time-varying risk aversion, in contrast, we can explicitly specify why and how aggregate preferences vary through time.

In our model, aggregate preferences vary in an economically sensible way. Just as in a habit formation model, good news about aggregate consumption growth (i.e. an unexpected increase in aggregate consumption growth) temporarily lowers aggregate risk aversion. Bad news about aggregate consumption growth, in contrast, temporarily raises risk aversion. In addition, to capture the anxiety about excessive inflation that Shiller (1996) documents, bad news about inflation (i.e. an unexpected increase in inflation) temporarily raises and good news about inflation temporarily lowers aggregate risk aversion.

Our approach is worthwhile even if one feels that modeling time-variations in aggregate preferences is somewhat ad-hoc, since we can always interpret our model as a reduced form for a more elaborate model with nominal frictions, money illusions, or money preferences. Because the evidence on the cause of concern about inflation is not clearcut, it is not obvious which structural extension of the standard model is best. With our approach, however, we can still study the asset pricing implications of time-variations in aggregate risk aversion in response to news about inflation. Furthermore, we formulate our model such that if the data does not support our hypothesis, its pricing implications just collapse to those of Campbell and Cochrane's habit formation model.

We estimate our model and explore its pricing implications on the term structure of nominal interest rates, always comparing the results to those for the standard models with and without habit formation. We focus on interest rates, rather than stock returns, for two reasons. Firstly, the relation between unexpected inflation and real returns has been documented predominantly for bonds, rather than stocks, since bond prices are significantly less noisy than stock prices. Secondly, we already know that Campbell and Cochrane's (1999) model, which is nested in our model, does an impressive job at characterizing stock returns.

However, it is also well known that the traditional consumption based asset pricing models cannot explain the stylized features of nominal interest rates.[2] Therefore, fitting the term structure of nominal interest rates is a more challenging objective for our model than just matching the conditional moments of stock returns.

We estimate our model and its competitors using an intuitive least squares approach. Since two of the three models we consider do not permit closed-form bond prices, we use simulations to evaluate the least squares criterion. The resulting simulated least squares estimator is quite unique. Compared to standard simulation based estimators that are consistent only in the limit as the simulation size tends to infinity, our approach generates consistent parameter estimates for any fixed simulation size.

The empirical results unambiguously support our hypothesis that aggregate risk aversion varies in response to news about inflation. The induced time-variation in risk aversion is both economically and statistically significant. Not only is the estimated range of aggregate risk aversion realistic, but the time-variation in risk aversion also relates to observed business conditions. Aggregate risk aversion rises during the recessions of 1973-1975 and 1979-1982, but it remains constant or decreases during the other subperiods since 1970.

Time-varying risk aversion offers an interesting characterization of the cyclical dynamics of the nominal term structure. The aggregate risk aversion from our model is correlated with observed and fitted yields, yield spreads, realized real returns, and expected term premiums. These correlations are consistent with simple economic intuition about risk aversion and the expected reward required for generating sufficient demand for risky securities.

Consumption based asset pricing models with additively separable preferences cannot explain the stylized features of nominal interest rates. In particular, Backus, Gregory, and Zin (1989) demonstrate that these models produce negative expected term premiums. They also find that the expected term premiums are virtually constant, which implies that the expectations hypothesis holds for these models. In the data, in contrast, the average term premiums are reliably positive, at least at the short end of the term structure, and Fama and Bliss' (1987) term premium regressions strongly reject the expectations hypothesis.

Our model generates time-varying expected term premiums that are positive on average. Even more impressively, the term premium regressions with the fitted data from our model replicate almost exactly Fama and Bliss' results, suggesting that the forward-spot spreads track one-for-one the time-variations in expected term premiums. The surprising result,

[2]Shiller (1990) observes that existing term structure models leave some important features of the data, like the cyclical time-variation of the level and shape of the yield curve and the rejection of the expectations hypothesis, largely unexplained.

however, is that our model implies a strong *negative* correlation between the forward-spot spreads and the expected term premiums, which contradicts the implication of the term premium regressions. This suggests that the Fama and Bliss regressions are spurious, which is confirmed when we correct the regressions for the persistence of the residuals. Therefore, our model provides a novel explanation of the controversial Fama and Bliss regressions and their implied rejection of the expectation hypothesis.[3]

We formulate our model in Section 2. We then explain our simulated least squares approach for estimating term structure models in Section 3. The empirical results for our model and its competitors are in Section 4. We conclude in Section 5

2 Model

2.1 Aggregate Preferences

Consider an endowment economy with an infinitely lived representative agent who maximizes the conditional expectation of the following utility of life-time consumption:

$$u(C_0, C_1, C_2, \ldots) = \sum_{t=0}^{\infty} \delta^t u(C_t - X_t), \tag{1}$$

where

$$u(C_t - X_t) = \begin{cases} \dfrac{(C_t - X_t)^{1-\alpha} - 1}{1-\alpha} & \text{if } \alpha > 1 \\ \ln(C_t - X_t) & \text{if } \alpha = 1. \end{cases} \tag{2}$$

The coefficient δ is a subjective discount factor, C_t is real consumption at time t, and $X_t \leq C_t$ is a subjective reference level against which consumption is measured. For example, X_t may be an external consumption habit formed through past aggregate consumption to capture a "catching up with the Joneses" source of utility [Abel (1990,1999)]. In principle, the only restriction on X_t is that it is not in the agent's choice set, but rather is exogenous.

The coefficient α measures the curvature of the representative agent's utility function with respect to its argument $C_t - X_t$. Unless consumption is referenced to $X_t = 0$, α does not correspond to the coefficient of relative risk aversion. Relative risk aversion, which

[3]The Fama and Bliss regressions are controversial because, together with the pro-cyclical time-variations in the forward-spot spreads, they imply that expected returns vary pro-cyclically. This contradicts both economic intuition and the results of Fama and French (1989) and Fama (1990).

measures the curvature of the utility function with respect to consumption, is time-varying:

$$RRA_t = -\frac{C_t u_{cc}(C_t - X_t)}{u_c(C_t - X_t)} = \alpha \frac{C_t}{C_t - X_t} = \alpha \frac{1}{S_t}, \tag{3}$$

where u_c and u_{cc} denote the first and second derivatives of the utility function with respect to consumption, respectively. The state variable $S_t = (C_t - X_t)/C_t$ is called the consumption surplus ratio and measures the fraction of consumption that actually yields utility. The remaining fraction $1 - S_t$ of consumption is used to maintain the reference level. Notice that for the second equality we use the fact that the reference level does not depend on current consumption, since it is exogenous, so that $\partial X_t / \partial C_t = 0$.

We complete our description of the aggregate preferences by specifying the dynamics of log relative risk aversion $\gamma_t \equiv \ln RRA_t$.[4] Specifically, we assume that log relative risk aversion evolves as a mean-reverting AR(1) process:

$$\gamma_{t+1} = \bar{\gamma}(1 - \phi) + \phi \gamma_t - e_{t+1}, \tag{4}$$

where $\bar{\gamma}$ is the average log relative risk aversion, $\phi \in [0, 1)$ measures the speed of mean reversion, and e_{t+1} are the innovations that drive time-variations in risk aversion. Notice that we subtract, not add, the innovations. This way, we can interpret a positive e_t as good economic news that temporarily lowers the representative agent's relative risk aversion. A negative innovation represents bad news that temporarily raises aggregate risk aversion.

What causes aggregate risk aversion to change? In a habit formation model, where the reference level is only a function of lagged aggregate consumption, relative risk aversion changes in response to news about aggregate consumption growth. Holding habit constant, a negative shock to consumption growth lowers the consumption surplus ratio S_t because the representative agent's consumption approaches that of "the Joneses." Given the definition of relative risk aversion in equation (3), such decrease in the consumption surplus ratio is equivalent to an increase in the representative agent's relative risk aversion.

We postulate that aggregate risk aversion varies not only in response to news about consumption growth, but also in response to news about inflation. Formally, we assume that the innovations e_t to log relative risk aversion γ_t are:

$$e_{t+1} = \lambda(\gamma_t)\varepsilon^g_{t+1} - \theta(\gamma_t)\varepsilon^\pi_{t+1}, \tag{5}$$

[4] Specifying a process for the log relative risk aversion γ_t is mathematically equivalent to Campbell and Cochrane's (1999) approach of specifying a process for the log consumption surplus ratio $s_t \equiv \ln S_t$.

with

$$\varepsilon^{g}_{t+1} = g_{t+1} - \mathrm{E}_t\big[g_{t+1}\big] \quad \text{and} \quad \varepsilon^{\pi}_{t+1} = \pi_{t+1} - \mathrm{E}_t\big[\pi_{t+1}\big], \tag{6}$$

where g_{t+1} denotes consumption growth (i.e. $g_{t+1} = \ln C_{t+1} - \ln C_t$) and π_{t+1} denotes inflation (i.e. $\pi_{t+1} = \ln CPI_{t+1} - \ln CPI_t$) from time t to $t+1$. The innovations ε^{g}_{t+1} and ε^{π}_{t+1} have zero mean and reflect news about consumption growth and inflation, respectively. The two functions $\lambda(\gamma_t)$ and $\theta(\gamma_t)$ are the so-called sensitivity functions because they determine the sensitivity of the representative agent's log risk aversion to news about g_t and π_t. Notice that when both sensitivities are positive, unexpected consumption growth and unexpected disinflation represent good news that temporarily lowers aggregate risk aversion.

2.2 Sensitivity Functions

We parameterize the sensitivity functions in equation (5) as follows:

$$\lambda(\gamma_t) = \exp(\gamma_t) - 1 \quad \text{and} \quad \theta(\gamma_t) = \theta\big[\exp(\gamma_t) - 1\big]. \tag{7}$$

As long as $\theta \geq 0$ and $\gamma \geq 0$, both sensitivities are non-negative and increase linearly in the level of relative risk aversion.[5] It is intuitive to assume that the sensitivities do not change signs just because the agent becomes more or less risk averse. Likewise, it makes sense to think that a more risk averse agent is also more responsive to news about consumption growth and inflation than a less risk averse, but otherwise identical, agent.

The choice of $\lambda(\gamma_t)$ is motivated by Campbell and Cochrane's (1999) sensitivity function for time-variations in the log consumption surplus ratio. In our notation, the sensitivity of log relative risk aversion to shocks in consumption growth implied by their model is:

$$\lambda(\gamma_t) = \begin{cases} \exp(\bar{\gamma})\sqrt{1 + 2(\gamma_t - \bar{\gamma})} - 1 & \text{if } \gamma_t \geq \bar{\gamma} + \frac{1}{2}\big[\exp(-2\bar{\gamma}) - 1\big] \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

They choose this sensitivity function because it satisfies three criteria: (i) the real riskfree rate is constant, (ii) the agent's reference level is predetermined in the steady state of the economy, and (iii) the reference level is also predetermined near the steady state. The first criterion is convenient to limit the variability of marginal utility, but it is not crucial for the success of their model. The second and third criteria are more important. They permit the

[5] In our discrete time setting, we cannot guarantee that log relative risk aversion is non-negative, unless we restrict the distribution of the innovations ε^{g}_{t+1} and ε^{π}_{t+1}. However, in a continuous time formulation, where log relative risk aversion evolves as a diffusion, our choice of sensitivity functions guarantees non-negativity.

interpretation of the agent's reference level as being exogenous. This is because criteria (ii) and (iii) imply that at and near the steady state $\partial X_t / \partial C_t = 0$.

Our $\lambda(\gamma_t)$ is a first-order approximation (in relative risk aversion) of Campbell and Cochrane's sensitivity function around the steady state $\gamma_t = \bar{\gamma}$. It satisfies their criteria (ii) and (iii), but does not generate a constant real riskfree rate, which of course would be awkward in a term structure application. Furthermore, our specification avoids the non-differentiability at $\gamma_t = \bar{\gamma} + 1/2[\exp(-2\bar{\gamma}) - 1]$ that could complicate the estimation.

The function $\theta(\gamma_t)$ is linear in relative risk aversion, which is nice for interpretation and aesthetics, and can always be interpreted as a first-order approximation of a more complicated nonlinear sensitivity function. The coefficient θ allows us to formally test the hypothesis that aggregate risk aversion is insensitive to inflation news (i.e. $\theta = 0$). When $\theta = 0$, our model collapses to Campbell and Cochrane's habit formation model.

2.3 Term Structure Implications

The price p_t^n of an n-period default-free real discount bond is just the conditional expectation of the real pricing kernel [see Campbell, Lo, and MacKinlay (1997) or Cochrane (1998)]:

$$p_t^n = \mathrm{E}_t\big[m(t, t+n)\big], \tag{9}$$

where m denotes the real pricing kernel, or the intertemporal marginal rate of substitution, from time t to $t+n$. In our model, the real pricing kernel is:

$$m(t, t+n) = \delta^n \exp\Big\{\gamma_{t+n} - \gamma_t - \alpha\big(g_{t+1} + \dots g_{t+n}\big)\Big\}. \tag{10}$$

Likewise, the price P_t^n of the corresponding nominal discount bond is:

$$P_t^n = \mathrm{E}_t\big[M(t, t+n)\big], \tag{11}$$

where M is the the nominal pricing kernel:

$$M(t, t+n) = \delta^n \exp\Big\{\gamma_{t+n} - \gamma_t - \alpha\big(g_{t+1} + \dots g_{t+n}\big) - \big(\pi_{t+1} + \dots \pi_{t+n}\big)\Big\}. \tag{12}$$

The only difference between the pricing kernels of our model and those of the standard power utility model is the change in log relative risk aversion term $\gamma_{t+n} - \gamma_t$. If relative risk aversion is constant over the next n periods, the term structure implications of our model are identical to those of the standard power utility model.

2.4 Factor Model Representations

We can derive both an unconditional and a conditional factor model representation of our model under the assumption of joint normality. [6] The factor model representations help generate some more intuition about the sources of risk premiums in our model. They also illustrate further the differences between our model, Campbell and Cochrane's model, and the standard power utility model.

For the unconditional factor model representation, assume that real returns r^i_{t+1} on any security i, consumption growth g_{t+1}, and inflation π_{t+1} are jointly normal. The unconditional pricing relation:

$$\mathrm{E}\big[m(t,t+1)r^i_{t+1}\big] = 1 \tag{13}$$

can be written as:

$$\mathrm{E}\big[r^i_{t+1}\big] = \frac{1}{\mathrm{E}\big[m(t,t+1)\big]} - \frac{\mathrm{Cov}\big[m(t,t+1), r^i_{t+1}\big]}{\mathrm{E}\big[m(t,t+1)\big]}. \tag{14}$$

Then, applying Stein's lemma yields the unconditional factor model:[7]

$$\mathrm{E}\big[r^i_{t+1}\big] = b_0 + b_1\mathrm{Cov}\big[g_{t+1}, r^i_{t+1}\big] + b_2\mathrm{Cov}\big[\gamma_{t+1} - \gamma_t, r^i_{t+1}\big], \tag{15}$$

where

$$b_0 = \frac{1}{\mathrm{E}\big[m(t,t+1)\big]}, \quad b_1 = \alpha, \quad \text{and} \quad b_2 = -1. \tag{16}$$

In the unconditional consumption CAPM with power utility, which corresponds to the case $b_2 = 0$, expected returns depend linearly on the return covariance with consumption growth. When we allow for time-variations in relative risk aversion, the return covariance with changes in relative risk aversion adds a risk premium. Since b_2 is negative, our model just says that securities which tend to generate low returns in times of rising risk aversion are undesirable and hence need to offer an expected return premium to attract investors.

Given the dynamics of aggregate risk aversion, we can also derive a conditional factor model representation, in which conditional expected returns depend linearly on the security's

[6] Even without the assumption of joint normality, the factor model representations of our model can be obtained by linearizing the real pricing kernel with a first-order Taylor series expansion.

[7] Stein's lemma states that if two random variables x and y are jointly normal and g is a differentiable function of one variable, such that $\mathrm{E}|g'(x)|$ exists, then $\mathrm{Cov}[g(x), y] = \mathrm{E}[g'(x)]\mathrm{Cov}(x, y)$.

return covariances with innovations to consumption growth and inflation. Assume that real returns r^i_{t+1} and the innovations ε^g_{t+1} and ε^π_{t+1} are jointly normal, conditional on the information at time t. The conditional pricing relation:

$$\mathrm{E}_t\left[m(t,t+1)r^i_{t+1}\right] = 1 \tag{17}$$

can be written as

$$\mathrm{E}_t\left[r^i_{t+1}\right] = \frac{1}{\mathrm{E}_t\left[m(t,t+1)\right]} - \frac{\mathrm{Cov}_t\left[m(t,t+1), r^i_{t+1}\right]}{\mathrm{E}_t\left[m(t,t+1)\right]}, \tag{18}$$

which, using Stein's lemma again, yields the conditional factor model:

$$\mathrm{E}_t\left[r^i_{t+1}\right] = b_{0,t} + b_{1,t}\mathrm{Cov}_t\left[\varepsilon^g_{t+1}, r^i_{t+1}\right] + b_{2,t}\mathrm{Cov}_t\left[\varepsilon^\pi_{t+1}, r^i_{t+1}\right], \tag{19}$$

where

$$b_{0,t} = \frac{1}{\mathrm{E}_t\left[m(t,t+1)\right]}, \quad b_{1,t} = \lambda(\gamma_t) + \alpha, \quad \text{and} \quad b_{2,t} = -\theta(\gamma_t). \tag{20}$$

When both sensitivities are zero, so that $b_{1,t} = \alpha$ and $b_{2,t} = 0$, this conditional factor model corresponds to the standard conditional consumption CAPM with power utility, since $\mathrm{Cov}_t[\varepsilon^g_{t+1}, r^i_{t+1}] = \mathrm{Cov}_t[g_{t+1}, r^i_{t+1}]$. As we said before, when $\lambda(\gamma_t) > 0$ and $\theta(\gamma_t) = 0$, our model collapses to Campbell and Cochrane's habit formation model. Their model raises the risk premium, relative to the standard power utility model, for a positive return covariance with consumption growth. This helps explain the equity premium puzzle without relying on an unreasonably high curvature parameter α. In our model, where both sensitivities are positive, a security also receives a risk premium for having a negative return covariance with inflation. Securities which tend to generate low returns in times of unexpected inflation are unattractive and hence need to offer an expected return premium to attract investors.

2.5 Consumption Growth and Inflation

To price nominal bonds, we require the joint dynamics of consumption growth and inflation. Following Hansen and Singleton (1983) and Labadie (1989), we assume that consumption growth and inflation follow an exogenous vector autoregressive (VAR) process, which allows for interdependencies between the variables through their lagged levels and their innovations. Some recent empirical studies [e.g. Boudoukh (1993) or Pennacchi (1991)] emphasize the importance of accounting for interdependencies between consumption growth and inflation in characterizing the term structure of nominal interest rates.

Formally, we assume that the representative agent's information set is made up of an exogenous history of consumption growth rates $\{g_t, g_{t-1}, \ldots g_{t-L}\}$ and inflation rates $\{\pi_t, \pi_{t-1}, \ldots \pi_{t-L}\}$ and that g_t and π_t follow a stationary VAR(L) process:

$$\begin{bmatrix} g_t \\ \pi_t \end{bmatrix} = \Phi_0 + \sum_{l=1}^{L} \Phi_l \begin{bmatrix} g_{t-l} \\ \pi_{t-l} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^g \\ \varepsilon_t^\pi \end{bmatrix}, \tag{21}$$

where Φ_0 is a 2×1 vector and the Φ_l are 2×2 matrices. The innovations ε_t^g and ε_t^π are jointly normal with covariance matrix Σ.

3 Econometric Approach

3.1 Simulated Least Squares Estimation

To illustrate our econometric approach, we first assume that the consumption growth and inflation process is known. That is, we take the order L and the parameters $[\Phi_0, \ldots, \Phi_L, \Sigma]$ of the VAR for g_t and π_t as given and describe only how to estimate the five preference parameters $\beta = \left[\alpha, \bar{\gamma}, \phi, \theta, \delta\right]$. Later, we will incorporate the estimation of the VAR.

The data is a panel of T price observations $\{P_t = [P_t^{n_1}, P_t^{n_2}, \ldots, P_t^{n_N}]'\}_{t=1}^{T}$ for a cross-section of N default-free nominal discount bonds with different but constant maturities n_i. The most natural way to estimate the preference parameters β is by least squares. The least squares estimator minimizes:

$$\frac{1}{T} \sum_{t=1}^{T} \left[P_t - P(t,\beta)\right]' W_t \left[P_t - P(t,\beta)\right] = \frac{1}{T} \sum_{t=1}^{T} q(t,\beta), \tag{22}$$

where

$$P(t,\beta) = \mathrm{E}_t\left[M(t,t+n_1), M(t,t+n_2), \ldots, M(t,t+n_N)\right]' \tag{23}$$

and W_t is an $N \times N$ weighting matrix. $P(t,\beta)$ denotes the vector of theoretical bond prices from the model [see equation (11)], corresponding to the vector of observed bond prices P_t, as a function of the preference parameters β.

Least squares estimation is intuitive, although not necessarily efficient. In words, the least squares estimator minimizes the weighted sum of squared pricing errors of the model across maturities and through time. The weighting matrix W_t is used to correct the estimator for heteroscedasticity in the pricing errors.

Unfortunately, when the sensitivities of aggregate risk aversion to economic news are time-varying, we cannot analytically evaluate the conditional expectations of the pricing kernel in equation (23).[8] Computing them numerically is also infeasible because it involves as many as n_N nested integrals. Thus, we propose evaluating the conditional expectations, and thereby the pricing errors of the model, through conditional simulations.

Our simulated least squares estimator minimizes:

$$\frac{1}{T}\sum_{t=1}^{T}\big[P_t-\hat{P}_a(t,\beta)\big]'W_t\big[P_t-\hat{P}_b(t,\beta)\big]=\frac{1}{T}\sum_{t=1}^{T}\hat{q}(t,\beta), \tag{24}$$

where

$$\begin{aligned}\hat{P}_a(t,\beta)&=\frac{1}{S}\sum_{s=1}^{S}\big[\hat{M}_s(t,t+n_1),\hat{M}_s(t,t+n_2),\ldots,\hat{M}_s(t,t+n_N)\big]'\\ \hat{P}_b(t,\beta)&=\frac{1}{S}\sum_{s=S+1}^{2S}\big[\hat{M}_s(t,t+n_1),\hat{M}_s(t,t+n_2),\ldots,\hat{M}_s(t,t+n_N)\big]'.\end{aligned} \tag{25}$$

Both $\hat{P}_a(t,\beta)$ and $\hat{P}_b(t,\beta)$ approximate the conditional expectations $P(t,\beta)$ with an average over S simulated realizations $\hat{M}_s(t,t+n_i)$ of the nominal pricing kernels from time t to $t+n_i$, for $i=1,2,\ldots,N$. The simulations are conditional on the information available at time t.

To simulate one realization of the nominal pricing kernel from time t to $t+n_i$, we first evaluate the log relative risk aversion γ_t using the transition equation (4) with $\gamma_0=\bar{\gamma}$ and the forecast errors ε_t^g and ε_t^π from the VAR. Given this γ_t and the observed $\{g_{t-l},\pi_{t-l}\}_{l=0}^{L}$, we then simulate a path $\{\hat{g}_{t+s},\hat{\pi}_{t+s},\hat{\gamma}_{t+s}\}_{s=1}^{n_i}$ according to the transition equations (4) and (21), using n_i draws from a bivariate normal distribution with zero mean and covariance matrix Σ. Finally, we construct the simulated pricing kernel:

$$\hat{M}_s(t,t+n_i)=\delta^{n_i}\exp\Big\{\hat{\gamma}_{t+n_i}-\gamma_t-\alpha\big(\hat{g}_{t+1}+\ldots+\hat{g}_{t+n_i}\big)-\big(\hat{\pi}_{t+1}+\ldots+\hat{\pi}_{t+n_i}\big)\Big\}. \tag{26}$$

A subtle but important feature of our estimator is that it involves products of the two pricing errors $P_t-\hat{P}_a(t,\beta)$ and $P_t-\hat{P}_b(t,\beta)$, rather than just squared pricing errors. The

[8] This complication raises the question of why we use time-varying, rather than constant, sensitivity functions. With constant sensitivities, the model not only looks more elegant, but it also provides closed-form bond prices. Unfortunately, constant sensitivities cannot produce time-variations in expected excess returns [see equations (19) and (20)] and, in a term structure application, generate negative term premiums.

key to understanding the importance of using these products is to recognize that the two pricing errors are conditionally independent and identically distributed, since they come from different simulations.[9] Furthermore, both pricing errors have the same conditional mean $P_t - P(t,\beta)$, which then implies that the simulated $\hat{q}(t,\beta)$ are unbiased:

$$\mathrm{E}\big[\hat{q}(t,\beta)\big] = \mathrm{E}\big[q(t,\beta)\big]. \tag{27}$$

This unbiasedness allows the law of large numbers to operate across the time-series dimension of the sample to control the joint effect of the simulation errors $\hat{q}(t,\beta) - q(t,\beta)$. More precisely, the second term on the right hand side of the identity:

$$\frac{1}{T}\sum_{t=1}^{T} \hat{q}(t,\beta) \equiv \frac{1}{T}\sum_{t=1}^{T} q(t,\beta) + \frac{1}{T}\sum_{t=1}^{T} \big[\hat{q}(t,\beta) - q(t,\beta)\big] \tag{28}$$

vanishes as $T \to \infty$, irrespective of S. As a result, our estimator and the true weighted least squares estimator converge to the same value for any fixed simulation size.[10]

Of course, in practice the consumption growth and inflation process is not really known. We thus employ the following two-stage procedure. In the first stage, we use standard maximum likelihood to estimate the order L and the parameters $[\Phi_0, \ldots, \Phi_L, \Sigma]$ of the VAR for g_t and π_t. In the second stage, we then obtain the simulated least squares estimates of the preference parameters β, conditional on the parameters and innovations of the estimated VAR from the first stage. Naturally, the asymptotic standard errors of the second-stage estimates take into account the estimation errors from the first stage.

The resulting estimates of the preference parameters are consistent and asymptotically normal for any fixed simulation size S. We present and further discuss the asymptotics of our simulated least squares method in a more general setting in the Appendix.

3.2 Weighting Matrix

Bond prices are notoriously heteroscedastic across maturities. The unconditional variance of five-year bond prices is roughly five times that of one-year bond prices. This means that without an appropriate weighting matrix W_t, our estimator places a disproportionate emphasis on fitting long-term bond prices, relative to fitting short-term bond prices. In

[9] The bond prices $P_a(t,\beta)$ are computed from the simulations $s = \{1, 2, \ldots, S\}$, while for $P_b(t,\beta)$ we use the simulations $s = \{S+1, S+2, \ldots, 2S\}$, which by construction are independent of the first S simulations.

[10] In contrast, the usual simulation based estimators produce inconsistent parameter estimates for a fixed simulation size. See McFadden (1989) or Danielsson (1994) for examples and discussions of the biases in simulation based estimation.

contrast, yields are quite homoscedastic. The unconditional variance of five-year yields is nearly the same as that of one-year yields.

Because of their homoscedasticity property, we generally prefer least squares estimation with yields over least squares estimation with bond prices. Unfortunately, with yields we cannot prove the consistency of our econometric method for a fixed simulation size. The problem is that the transformation from simulated bond prices to simulated yields introduces a systematic bias in the approximate yields that only vanishes as $S \to \infty$.

We resolve this issue by using the following diagonal weighting matrix:

$$W_t = \begin{bmatrix} \frac{1}{n_1^2} & & & \\ & \frac{1}{n_2^2} & & \\ & & \cdots & \\ & & & \frac{1}{n_N^2} \end{bmatrix}. \tag{29}$$

This weighting matrix transforms our least squares estimation with bond prices into a first-order approximation of least squares estimation with yields.[11] In dividing each pricing error by the bond's maturity, we remove most of the heteroscedasticity in bond prices.

4 Empirical Results

4.1 Consumption Growth and Inflation

We collect monthly data on aggregate consumption and consumer prices for January 1969 through June 1997 from the DRI/CITIBASE database. The consumption data is per capita real expenditures on nondurable goods and services [i.e. $C_t = (\texttt{GMCNQ}+\texttt{GMCSQ})/\texttt{POP}$], and the price level data is the implicit price deflater for expenditures on nondurable goods and services [i.e. $CPI_t = (\texttt{GMCNQ}\times\texttt{GMDCN}+\texttt{GMCSQ}\times\texttt{GMDCS})/(\texttt{GMCNQ}+\texttt{GMCSQ})$]. Consumption growth g_t and inflation π_t are the changes in log consumption and in the log price level, respectively.

Table 1 presents maximum likelihood estimates of a VAR(2) for consumption growth and inflation. The estimates in Panel A are for the full sample. Panels B and C show estimates for two subsamples; January 1969 through September 1979 and October 1979 through June 1997, respectively. The reason for splitting the sample in October 1979 is the appointment of Paul Volcker as Federal Reserve chairman and the accompanying change in the Federal

[11] The least squares estimator with yields minimizes $\sum_{t=1}^{T}[Y_t - Y(t,\beta)]^2$. Expanding $Y = -\ln P/n$ around $P = 1$, shows that minimizing $\sum_{t=1}^{T}[P_t/n - P(t,\beta)/n]^2$ approximates least squares estimation with yields.

Reserve policy toward inflation at that time [e.g. Friedman and Kuttner (1996) or Clarida, Gali, and Gertler (1998)].

We estimate but decisively reject a VAR(1) for consumption growth and inflation because the residuals are strongly autocorrelated and cross-autocorrelated. The VAR(2) describes the data reasonably well and represents a considerable improvement over a VAR(1). Most elements in the residual cross-autocorrelation matrices for one through twelve lags (not shown in Table 1) are within two standard errors from zero. Extending the specification to a VAR(3) does not significantly improve the fit. Likewise, allowing for autoregressive conditional heteroscedasticity (ARCH) of the residuals, in the spirit of Boudoukh's (1993) VAR with stochastic volatility of inflation, does not significantly help describe the data.

The results in Table 1 are reasonable. For the whole sample, monthly consumption growth is negatively autocorrelated and responds negatively to lagged inflation. Inflation, in contrast, is highly persistent, both at the first and second lags, and does not significantly relate to lagged consumption growth. The innovations to consumption growth are about four times as variable as the innovations to inflation, and the correlation between the two innovations is only -0.08.

The differences between the two subsamples are interesting. Before the change in Federal Reserve policy toward inflation (in Panel B), consumption growth is autocorrelated not only at the first, but also at the second lag. More interestingly, consumption growth has a strong negative correlation with the second lag of inflation. After the policy change (in Panel C), the second lag of consumption is insignificant and the correlation with lagged inflation shifts from the second lag to the first. In other words, when the Federal Reserve starts managing inflation more pro-actively, aggregate consumption growth becomes less persistent and reacts more quickly to realized inflation.

4.2 Preference Parameters

We collect monthly prices of nominal U.S. Treasury bills with one, three, and six months to maturity and of nominal U.S. Treasury discount bonds with one, three, and five years to maturity from CRSP. The Treasury bill data are from the Fama Treasury term structure file `FPRIAVE6.DAT`, and the artificial discount bond data are from the Fama and Bliss discount bond file `FAMABLISPRI.DAT`. The sample period is again January 1969 through June 1997.

Table 2 presents simulated least squares estimates of the preference parameters for our model, Campbell and Cochrane's model (our model with $\theta = 0$), and the standard power utility model (our model with $X_t = 0$), from the above described panel of 342×6 bond prices. The estimation is based on the VAR(2) for consumption growth and inflation in Panel A of Table 1, and it uses $S=200$ simulated paths of the pricing kernels to compute the pricing errors.[12] The standard errors in parentheses are adjusted for autocorrelation and heteroscedasticity of the residuals using Newey and West's method with 12 lags.

The estimates for our model in Panel A are economically plausible. The constraint $\alpha \geq 1$ is binding, which from the conditional factor representation (19) and (20) and our specification of the sensitivity function (7) implies that the reward $\beta_{1,t}=(RRA_t-1)+\alpha$ for a security's return covariance with consumption growth is dominated by the time-varying relative risk aversion RRA_t.[13] The average level of relative risk aversion, implied by the estimate of $\bar{\gamma}$, is 1.67 with a standard error of 0.38. The estimate of ϕ implies that an innovation to log relative risk aversion has a half-life of more than five years. Finally, the estimated subjective discount factor δ is, as expected, very close to one.

The most important result in Panel A is that the estimate of θ is positive and more than three standard errors from zero. This suggests that aggregate risk aversion indeed responds in an economically plausible and statistically significant way to news about inflation.[14]

When we constrain $\theta=0$, for Campbell and Cochrane's model in Panel B, the persistence of log relative risk aversion increases further, to a half-life of more than ten years, and the average level of relative risk aversion drops to 1.56. When we further constrain $X_t=0$, for the standard power utility model in Panel C, the curvature parameter α, which then takes on the interpretation of the relative risk aversion, remains indistinguishable from one.

Since the VAR estimates are sensitive to the sampling period, we separately estimate the preference parameters for the two subsamples. The estimates in Panels A and B of Table 3 are based on the VAR(2) in Panels B and C of Table 1, respectively. The subsample results support the conclusion that aggregate risk aversion varies in response to news about inflation. The estimates of θ are both positive and more than two standard errors from zero. There is, however, a significant quantitative difference in the sensitivities of log relative risk aversion to news about inflation. In the first subsample, the steady state sensitivity $\theta(\gamma_t)$ is $15.29[\exp(0.45)-1]=8.62$. In the second subsample, it is only $10.79[\exp(0.27)-1]=3.36$.

[12] We use this seemingly small simulation size to keep the numerical minimization of the least squares criterion feasible. However, doubling the simulation size has surprisingly little effect on the estimates.

[13] The fact that the constraint is binding is also consistent with the finding of Dun and Singleton (1986), Cochrane and Hansen (1992), and Cecchetti, Lam, and Mark (1994), that aggregate log utility can only be marginally rejected with data on monthly or quarterly stock and bond returns.

[14] To verify that our inferences are not driven by our choice of sensitivity functions (7), we also estimate the model with Campbell and Cochrane's specification for $\lambda(\gamma_t)$. The results are qualitatively identical.

4.3 Time-Varying Risk Aversion

Table 4 describes the dynamics of aggregate risk aversion implied by the estimates of our model and of Campbell and Cochrane's model. Panel A presents summary statistics of the relative risk aversion RRA_t, the innovations e_t to log relative risk aversion, the fraction X_t/C_t of consumption devoted to maintaining the reference level, and the changes $\Delta X_t/C_t$ in the reference level relative to consumption. Panel B shows correlations of relative risk aversion, changes in relative risk aversion ΔRRA_t, and innovations to log relative risk aversion with each other, with consumption growth g_t and its innovations ε_t^g, with inflation π_t and its innovations ε_t^π, and with the realized real returns r_t^1 on a one-month Treasury bill.

Our model generates economically plausible time-variations in risk aversion. Relative risk aversion is 1.82 on average, but it ranges from 1.41 to 3.07 in the sample. Campbell and Cochrane's model, in contrast, delivers virtually constant relative risk aversion of 1.54. As the parameter estimates in Tables 2 and 3 foreshadow, relative risk aversion is highly persistent in both models, with monthly autocorrelations of 0.98 and 0.99, respectively.

Another way to examine time-variations in risk aversion is through the fraction X_t/C_t of consumption devoted to maintaining the reference level, since $X_t/C_t = 1 - 1/RRA_t$. For our model, this statistic is 43 percent on average, but it ranges from 29 to 67 percent. If we interpret the reference level as a consumption habit, more than one-third, and at times more than two-thirds of consumption is devoted to "catching up with the Joneses." Notice that although the fraction of consumption devoted to maintaining the reference level varies considerable throughout the sample, the reference level itself is fairly stable. From one month to the next, the reference level relative to consumption changes by no more than five percent.

The changes in relative risk aversion are negatively correlated with the innovations to log relative risk aversion.[15] In our model, these innovations, in turn, are positively correlated with unexpected consumption growth (correlation of 0.22) and are negatively correlated with unexpected inflation (correlation of -0.91). The difference in magnitudes of the latter correlations suggests that news about inflation dominates news about consumption growth in generating the time-variations in relative risk aversion documented in Panel A. This means that the sensitivity of aggregate risk aversion to news about inflation is not only statistically significant (see Tables 2 and 3), but it is also economically important.

For our model, the correlation of relative risk aversion with inflation is almost as strong as the correlation of changes in relative risk aversion with inflation (correlations of 0.44 and 0.56, respectively). This means that times of high and unexpectedly rising inflation are associated

[15] The only reason this correlation is not perfect is the nonlinear log transformation $\gamma_t = \ln RRA_t$

with high and rising aggregate risk aversion, just as Shiller's (1996) survey would predict. For Campbell and Cochrane's model, in contrast, relative risk aversion is negatively correlated with inflation (correlation of -0.42) and changes in relative risk aversion are completely uncorrelated with inflation. [16] Their model suggests, somewhat counter-intuitively, that times of high inflation are associated with low aggregate risk aversion.

Another interesting feature of our model is that the correlations of relative risk aversion and changes in relative risk aversion with the realized real returns on the one-month Treasury bill are of opposite signs (correlations of 0.17 and -0.48, respectively). These correlations imply that high aggregate risk aversion tends to be associated with high real returns, but high real returns tend to drive down risk aversion. This mechanism is intuitive. When aggregate risk aversion is high, consumers demand greater real returns to hold risky securities. However, when real returns are unexpected high, which happens when inflation is unexpectedly low, aggregate risk aversion falls due to the embedded good news about inflation. Surprisingly, in Campbell and Cochrane's model, the realized real returns are unrelated to risk aversion.

Since for our model aggregate risk aversion relates to inflation and real returns, both of which are counter-cyclical indicators of economic conditions [e.g. Fama (1981,1982)], it is reasonable to suspect a business cycle pattern in risk aversion. Intuitively, we expect periods of strong economic growth to be associated with low or falling aggregate risk aversion, while recessions are associated with high or rising risk aversion.

Indeed, we find that aggregate risk aversion varies with the observed business cycles in the sample. Figure 1 plots the time-series of relative risk aversion for our model (solid line) and for Campbell and Cochrane's model (dashed line). For our model, risk aversion rises during the recessions of 1973-1975 and 1979-1982. In periods of stable or growing economic activity, risk aversion drops sharply (1983-1986), decrease slightly (1969-1973, 1986-1997), or remains constant (1975-1979). For Campbell and Cochrane's model, in contrast, relative risk aversion appears to vary pro-cyclically, which again is somewhat counter-intuitive.

4.4 Dynamics of the Term Structure

We now explore how well our model explains the dynamics of the term structure, always comparing the results to those obtained from Campbell and Cochrane's model and from the standard power utility model. Throughout this section, we compute bond prices by simulations, as neither our model nor Campbell and Cochrane's model permit closed form bond prices. Just as in the estimation, we approximate the n-period bond price with an

[16] The correlation between relative risk aversion in their model and in our model is -0.51.

average over S simulated realizations of the nominal pricing kernel from time t to $t+n$. The only difference is that we can now afford a large simulation size of $S = 50,000$, since we need to evaluate each bond price just once. The simulations are based on the VAR(2) for consumption growth and inflation in Panel A of Table 1 and the simulated least squares estimates of the preference parameters of the three models in Table 2.

Yields

Panel A of Table 5 describes the observed yields $Y_t = -\ln P_t/n$. The unconditional means increase from six percent for the one-month Treasury bill to eight percent for the five-year discount bond. The standard deviations differ little across maturities, which justifies our choice of weighting matrix in Section 3.2. All yields are severely autocorrelated, with long-term yields being somewhat more persistent than short-term yields.

The yields are positively correlated with relative risk aversion RRA_t from our model and with inflation, but they are negatively correlated with relative risk aversion RRA_t^{cc} from Campbell and Cochrane's model and with consumption growth. The correlations of the yields with relative risk aversion strengthen with maturity. In contrast, the correlations with consumption growth and inflation become weaker with maturity.

Panels B, C, and D of Table 5 describe the fitted yields from our model, Campbell and Cochrane's model, and the standard power utility model, respectively. The unconditional means for our model decrease slightly with maturity, while for the other two models they are constant across maturities. Otherwise, judging by how closely the standard deviations and autocorrelations of the fitted yields match those of the observed yields in Panel A, our model clearly outperforms its competitors.

As another measure of how well the models captures the dynamics of the term structure, we compute correlations between the fitted yields and the observed yields. For our model, these correlations are positive and surprisingly strong, ranging from 0.68 for the one-month Treasury bill to 0.87 for the five-year Treasury bond. For the other models, in contrast, the correlations are low for the short-term yields and decrease even further for the long-term yields. Interestingly, judging by this measure, the standard power utility model explains the dynamics of long-term yields better than Campbell and Cochrane's habit formation model.

Our fitted yields are highly correlated with relative risk aversion, with the correlations rising from 0.82 for short-term yields to 0.98 for long-term yields. This result is intuitive. In times of high or increasing risk aversion, high expected real returns are required to induce consumers to hold bonds. For expected real returns to be high, yields must also be high, for bond prices to be low. Conversely, in times of low or decreasing risk aversion, lower yields and hence lower expected real returns are sufficient to generate demand for bonds.

Expected Term Premiums

Backus, Gregory, and Zin (1989) illustrate that the class of representative agent models with additively separable preferences cannot generate positive expected term premiums, where we define an expected term premium as the expected return of holding an $(n+m)$-period bond for m periods in excess of the return of just holding an m-period bond for m periods. In the data, however, the average term premiums are reliably positive, at least at the short end of the term structure [e.g. Roll (1970), Startz (1982), Fama (1984a,1984b,1990), or Fama and Bliss (1987)].[17] Therefore, we next judge the three models by the sign and magnitude of the expected term premiums they generate.

Like Backus, Gregory, and Zin, we work with simple returns, rather than continuously compounded returns, and define the m-period expected term premium of an $(n+m)$-period bond as:[18]

$$\mathrm{E}_t\big[h(n+m, n : t+m)\big] - r(m : t), \tag{30}$$

where we use Fama and Bliss' (1987) notation:

$$h(n+m, n : t+m) = P^n_{t+m}/P^{n+m}_t \quad \text{and} \quad r(m : t) = 1/P^m_t. \tag{31}$$

To evaluate the expected return by simulation, we further simplify the expectation of the n-period bond price in m periods using the law of iterated expectations:

$$\mathrm{E}_t\big[P^n_{t+m}\big] = \mathrm{E}_t\Big[\mathrm{E}_{t+m}\big[M(t+m, t+m+n)\big]\Big] = \mathrm{E}_t\big[M(t+m, t+m+n)\big]. \tag{32}$$

Table 6 summarizes the expected term premiums generated by the three models. For our model in Panel A, both the means and the standard deviations of the premiums rise with maturity. More importantly, the means are all positive and above one standard deviation from zero. Given Backus, Gregory, and Zin's negative results for class of representative agent models with additively separable preferences, the ability of our model to generate positive

[17] For instance, in our sample the average excess return of holding a six-month Treasury bill for three months is 0.63 percept per annum, with an autocorrelation adjusted standard error of only 0.11 percent.

[18] The use of simple returns eases the computation of the expected term premiums. The expected term premiums with continuously compounded returns involve $\mathrm{E}_t[\ln P^n_{t+m}] = \mathrm{E}_t\big[\ln \mathrm{E}_{t+m}[M(t+m, t+m+n)]\big]$, which nests a conditional expectation nonlinearly in another conditional expectation. Unfortunately, we are not aware of any practical simulation methods to reliably evaluate such nonlinearly nested expectations.

and statistically significant expected term premiums is encouraging. Both Campbell and Cochrane's model in Panel B and the standard power utility model in Panel C produce small negative premiums that are constant throughout the sample.[19]

Unfortunately, the average expected term premiums generated by our model are still too small, when compared to the average term premiums observed in the data. For example, the average excess return of holding a six-month Treasury bill for three months exceeds 0.6 percent per annum in the sample, which is more than ten times the corresponding average premium generated by our model. However, relative to the average magnitude of the expected term premiums, the variability of the premiums is quite substantial. This finding is consistent with numerous empirical studies that suggest significant time-variations in expected term premiums [e.g. Fama and Bliss (1987) or Startz (1982)].

The expected term premiums generated by our model are strongly positively correlated with the representative agent's relative risk aversion. This indicates that time-varying risk aversion is instrumental in explaining the expected term premiums. Since the premiums are interpreted as compensation for risk, the positive correlations with risk aversion are intuitive. When consumers are more risk averse, they demand higher risk premiums for holding risky securities. Conversely, smaller risk premiums are sufficient to generate demand for risky securities in times of low aggregate risk aversion.

Recall from Section 4.3 that the time-variation in aggregate risk aversion is counter-cyclical. The strong positive correlations of the expected term premiums with relative risk aversion then imply that the expected excess returns tend to be low or decreasing when economic conditions are strong, but are high or increasing during recessions. This implication of our model is consistent with the findings of Fama and French (1989) and Fama (1990), that time-variations in expected excess returns on stocks and bonds are counter-cyclical.

Rejection of the Expectations Hypothesis

We already established that our model generates time-varying expected term premiums, which represents a rejection of the expectations hypothesis.[20] An important empirical task is then to find observable variables that forecast time-variations in the premiums. Much of the evidence on the time-variation in expected term premiums, or on the rejection of the expectations hypothesis, is based on forward-spot spread regressions. For example, Fama

[19]It is somewhat surprising that Campbell and Cochrane's model does not produce positive expected term premiums. Both Gregory and Voss (1991) and Salyer (1995) claim that habit formation can generate positive expected term premiums. Nevertheless, the premiums in their models are still virtually constant.

[20]The expectations hypothesis is one of the most popular and simple models of the term structure. It says that expected excess returns on bonds with different maturities are constant through time. See Lutz (1940) and Campbell, Lo, and MacKinlay (1997) for further details and alternatives to the expectations hypothesis.

and Bliss (1987) regress one-year excess returns of two- to five-year bonds on the spreads between the corresponding one-year forward and spot yields:

$$\mathrm{E}_t\big[h(n+12, n : t+12)\big] - r(12 : t) = a + b\big[f(n+12, n : t) - r(12 : t)\big], \tag{33}$$

where

$$f(n+12, n : t) = P_t^{n+12}/P_t^n \quad \text{and} \quad n = \{12, 24, 36, 48\}. \tag{34}$$

The slope coefficients b of these regressions turn out to be statistically different from zero, hence the expectations hypothesis is rejected, with point estimates that are close to one. Fama and Bliss thus conclude that the forward-spot spreads track almost one-for-one the time-variations in the one-year expected term premiums.

Panel A of Table 7 replicates the Fama and Bliss regressions for our sample. Panels B through D of the table show the regressions for the fitted bond prices from our model, Campbell and Cochrane's model, and the standard power utility model. The standard errors in parentheses are adjusted for autocorrelation and heteroscedasticity of the residuals using Newey and West's method with 12 lags.

It is impressive how closely the regressions for the fitted bond prices from our model match the regressions for the data. The estimated slope coefficients are all close to one and more than two standard errors from zero. Apparently, our model fits the data so well that it accounts for Fama and Bliss' rejection of the expectations hypothesis.

In contrast, Campbell and Cochrane's model in Panel C and the standard power utility model in Panel D fail to replicate the Fama and Bliss regressions. The slopes coefficients are less than 0.5. Oddly, the slope coefficients are still more than two standard errors from zero and thereby reject the expectations hypothesis, although the results in Panels B and C of Table 6 suggest that the models generate constant and negligible term premiums.

Fama and Bliss conclude that the forward-spot spreads track almost one-for-one the time-variation in expected term premiums. However, they also find that, at least after 1970, the forward-spot spreads relate to business cycles. Positive spreads tend to be associated with expansions and negative spreads tend to occur during recessions. Therefore, if the forward-spot spreads really captures the time-variation in expected term premiums, this *pro-cyclical* variation of the spreads implies that expected excess returns tend to be high when economic conditions are strong and low when they are weak. Unfortunately, this implication contradicts the more intuitive finding of Fama and French (1989) and Fama (1990), that time-variations in expected excess returns on stocks and bonds are *counter-cyclical.*

Our model offers a simple explanation of the apparently contradicting relations between business cycles and expected excess returns. It suggests that the Fama and Bliss regressions are spurious, probably because the forward-spot spreads and the excess returns are too persistent.[21] On one hand, our model fits the data well enough to replicate the Fama and Bliss results, predicting that the forward-spot spreads track almost one-for-one the time-variation in expected term premiums. On the other hand, the model generates expected term premiums that are strongly *negatively* correlated with the forward-spot spreads.

Table 8 shows in Panel A the correlations between the observed excess returns and the forward-spot spreads. It shows in Panel B the correlations between the fitted excess returns and the forward-spot spreads from our model (first row) and the correlations between the corresponding expected term premiums and the forward-spot spreads (second row). The conclusion from this table is clearcut. The Fama and Bliss regressions in Panel B of Table 7 are spurious. The true relation between the expected term premiums and the forward spot spreads is far from what the excess return regressions predict.

Inspired by the finding that for our model the Fama and Bliss regressions are spurious, a more careful examination of the results for the data is in order. We first notice that both in the data and for our model the excess returns and forward-spot spreads are highly persistent, with slowly decaying autocorrelations that suggest unit (or near-unit) roots.[22] Since the regressions in Table 7 do not account for this persistence in the dependent and independent variables, the residuals are also highly persistent, with autocorrelations around 0.9.

Although it is beyond the scope of this paper to formally examine whether the persistence in the residuals of the Fama and Bliss regressions causes the slope coefficients to be biased and their asymptotic standard errors to be invalid, we briefly check whether the regressions are robust to standard autocorrelation adjustments. If not, there is reason to believe that the Fama and Bliss results are just as spurious for the data as they are for our model.

Table 9 presents feasible generalized least squares (FGLS) estimates of the Fama and Bliss regressions [see Cochrane and Orcutt (1949)]. For Panel A, we assume that the least squares residuals follow a random walk. For Panel B, we assume that they follow an AR(1) process. The differences between the results in this table and in Panel A of Table 7 are striking. The

[21]The idea of spurious regressions originates from Yule (1926) and Granger and Newbold (1974). They discuss the problems in regressing a persistent dependent variable on persistent explanatory variables. More recently, Ferson, Sarkissian, and Simin (1998) and Stambaugh (1998) extend spurious regressions to finance. They are concerned with the problems in regressing less persistent returns on persistent explanatory variables.

[22]The one-, six-, and twelve-month autocorrelations of the 24-, 36-, 48-, and 60-month excess returns and forward-spot spreads are $\rho_1=\{0.92, 0.93, 0.93, 0.92\}$, $\rho_6=\{0.59, 0.57, 0.57, 0.55\}$, $\rho_{12}=\{0.18, 0.13, 0.10, 0.07\}$, and $\rho_1=\{0.86, 0.90, 0.88, 0.89\}$, $\rho_6=\{0.58, 0.62, 0.67, 0.65\}$, $\rho_{12}=\{0.41, 0.46, 0.49, 0.39\}$, respectively. Not surprisingly, given how well our model fits the data, the autocorrelations for our model are very similar.

estimated slope coefficients are all negative and statistically indistinguishable from zero, the R^2 coefficients are close to zero, and the regression residuals are almost uncorrelated through time. This suggests that the Fama and Bliss results are indeed spurious.

4.5 Robustness

4.5.1 Generalized Method of Moments Estimation

To verify that our conclusions are not an artifact of our simulated least squares estimation method, we also estimate the three models using a sequential generalized method of moments (GMM) procedure [see Ogaki (1993)]. In the first stage, we use maximum likelihood to estimate the structural parameters of the VAR. In the second stage, we over-identify the six moment conditions (12) with two sets of lagged instruments.[23] The first set of instruments consists of a constant, yields on the one- and six-months Treasury bills, and yields on the one-, three-, and five-year Treasury bonds. The second set of instruments consists of a constant, yields on the one-month Treasury bill, yields on the one- and five-year Treasury bonds, consumption growth, and inflation. We use Ogaki's weighting matrix with a 12 lag Newey and West adjustment for heteroscedasticity and autocorrelation.

Overall, the sequential GMM results verify our simulated least squares results. Panel A of Table 10 presents the two sets of GMM estimates of the preference parameters for our model. It confirms that the sensitivity of aggregate risk aversion to news about inflation is both economically and statistically significant. The estimates of θ are more than two standard errors from zero and are remarkably similar to the simulated least squares estimate in Table 2. In addition, the resulting time-variation in relative risk aversion closely resembles that in Table 4 and Figure 1. The other parameter estimates are also sensible.

The GMM estimates of the preference parameters for Campbell and Cochrane's model in Panel B and for the standard power utility model in Panel C are also reasonable. For Campbell and Cochrane's model, the estimates of $\bar{\gamma}$ are somewhat higher and the estimates of ϕ are slightly lower than the simulated least squares estimates in Table 2. This results in higher and more variable relative risk aversion, ranging from 2.7 to 3.1 for the first set of instruments and from 2.4 to 2.7 for the second set of instruments. Unfortunately, this does not improve the model's ability to characterize the dynamics of the term structure.

[23] Using the moment conditions $\mathrm{E}_t[M(t,t+n,\beta)B_{t+1}^{n-1}/B_t^n]=1$ yields qualitatively similar results, but the point estimates are more sensitive to the choice of instruments and weighting matrix. This sensitivity is obviously due to the strong correlations between the moments for different maturities n.

4.5.2 Equity Returns

Finally, to check that the time-variation in aggregate risk aversion in response to news about inflation is not entirely driven by the term structure data, we estimate our model with monthly and annual returns on ten NYSE size-decile portfolios. We again use Ogaki's sequential GMM procedure, with the moment conditions $\mathrm{E}_t\left[M(t,t+1)R^i_{t+1}\right]=1$, where R^i_{t+1} is the gross return on the ith size-decile portfolio, and a set of popular lagged instruments. The instruments are a constant, the aggregate dividend to price ratio, the default premium, the term premium, the yield on a one-month Treasury bill, and the CRSP market return.

Table 11 presents the two sets of GMM estimates of the preference parameters. Compared with the results in Tables 2 and 10, the estimates of $\bar{\gamma}$ are considerably higher.[24] The estimates of θ have the correct sign but are much smaller in magnitude. Furthermore, the estimate is statistically significant only for annual returns, not for monthly returns.

However, the difference in magnitude of the θ estimates is misleading. In Panel A of Table 2, the estimates $\bar{\gamma} = 0.5$ and $\theta = 14.4$ imply that the steady state sensitivity of log relative risk aversion to news about inflation is 9.3. In contrast, the estimates $\bar{\gamma} = 1.9$ and $\theta = 5.7$ in Table 11 (with annual returns) imply that it is 32.4. In other words, although the point estimates of θ in Table 11 are about a third of those in Tables 2 and 10, the resulting sensitivity of log relative risk aversion to news about inflation is three times as large.

In summary, the GMM results with stock returns confirm, at least with the annual data, that the sensitivity of aggregate risk aversion to news about inflation is both economically and statistically significant. It is not surprising that the inferences are weaker with stock returns than with bond returns, since the former is considerably more noisy than the later. The fact that the stock return data supports our model is very encouraging.

5 Conclusion

We formulated a consumption based asset pricing model in which aggregate risk aversion is time-varying, not only in response to news about aggregate consumption growth (as in a habit formation model), but also in response to news about inflation. The setup of our model resembles that of Campbell and Cochrane's (1999) habit formation model. However, rather than specifying the representative agent's habit formation process, we model the dynamics of aggregate risk aversion. This leads to a different interpretation and a more general framework for modeling time-varying risk aversion. We interpret our model either as

[24] The constraint $\alpha \geq 1$ is again binding.

capturing cyclical time-variations in aggregate preferences or as an approximate reduced form for a more elaborate model with nominal frictions, money illusions, or money preferences.

The empirical results unambiguously support our hypothesis that aggregate risk aversion varies in response to news about inflation. The induced time-variation in risk aversion is economically and statistically significant. Furthermore, the estimated range of aggregate risk aversion is realistic, and the time-variation in preferences relates to observed business conditions. Finally, we showed that our model captures the stylized features of the nominal term structure and even explains the apparent rejection of the expectations hypothesis.

Not only are our empirical results robust to alternative econometric methods, but they are also broadly supported by different data sets. However, a careful empirical analysis of our model with stock return data, as opposed to the term structure data we use here, is left to future research. Another topic of future research is to use our theoretical framework and empirical approach to investigate other sources of time-variation in aggregate preferences. For example, the model of Barberis, Huang, and Santos (1999) suggests that risk aversion varies in response to aggregate wealth shocks.

References

Abel, Andrew B., 1990, Asset pricing under habit formation and catching up with the Joneses, *American Economic Review* 80, 38–42.

Abel, Andrew B., 1999, Risk premia and term premia in general equilibrium, *Journal of Monetary Economics* 43, 3–33.

Amihud,Yakov, 1996, Unexpected inflation and stock returns revisited–evidence from Israel, *Journal of Money, Credit and Banking* 28, 22–33.

Andrews, Donald W. K., 1987, Consistency in nonlinear econometric models: A generic uniform law of large numbers, *Econometric Theory* 55, 1465–1471.

Backus, David K., Allan W. Gregory, and Stanley E. Zin, 1989, Risk premiums in the term structure: Evidence from artificial economies, *Journal of Monetary Economics* 24, 371–399.

Barberis, Nicholas, Ming Huang, and Tano Santos, 1999, Prospect theory and asset prices, Working Paper, University of Chicago.

Barr, David G., and John Y.Campbell, 1997, Inflation, real interest rates, and the bond market: A study of UK nominal and index-linked government bond prices, *Journal of Monetary Economics* 39, 361–383.

Billingsley, Patrick, 1961, The Lindeberg-Lévy theorem for martingales, *Proceedings of the American Mathematical Society* 12, 788–792.

Boudoukh, Jacob, 1993, An equilibirum model of nominal bond prices with inflation output correlation and stochastic volatility, *Journal of Money, Credit, and Banking* 25, 637–665.

Campbell, John Y., and John Ammer, 1993, What moves the stock and bond markets? A variance decomposition for long-term asset returns, *Journal of Finance* 48, 3–37.

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, 1996, A scorecard for indexed government debt, Working Paper, NBER #5587.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

Cecchetti, Stephen G., Pok-Sang Lam, and Nelson C. Mark, 1993, The equity premium and the risk-free rate: Matching the moments, *Journal of Monetary Economics* 31, 21–45.

Christiano, Lawrence J., and Martin Eichenbaum, 1991, Liquidity effects and the monetary transmission mechanism, *American Economic Review* 82, 346–353.

Christiano, Lawrence J., and Martin Eichenbaum, 1995, Liquidity effects, monetary policy, and the business cycle, *Journal of Money, Credit and Banking* 27, 1113–1136.

Clarida, Richard, Jordi Gali, and Mark Gertler, 1998, Monetary policy rules and macroeconomic stability: Evidence and some theory, Working Paper, New York University.

Cochrane, David, and Guy H. Orcutt, 1949, Application of least squares regression to relationships containing autocorrelated error terms, *Journal of the American Statistical Association* 44, 32–61.

Cochrane, John H., 1998, *Asset Pricing*, Unpublished Manuscript, University of Chicago.

Cochrane, John H., and Lars Peter Hansen, 1992, Asset pricing explorations for macroeconomics, in Oliver J. Blanchard and Stanley Fischer, ed., *NBER Macroeconomics Annual*, MIT Press, Cambridge, Massachusetts.

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

Crockett, Jean A., 1998, Rational expectations, inflation and the nominal interest rate, *Journal of Econometrics* 83, 349–363.

Danielsson, John, 1994, Stochastic volatility in asset prices: Estimation with simulated maximum likelihood, *Journal of Econometrics* 64, 375–400.

Dunn, Kenneth B., and Kenneth J. Singleton, 1986, Modeling the term structure of interst rates under non-separable ultility and durability of goods, *Journal of Financial Economics* 17, 27–55.

Fama, Eugene F., 1981, Stock returns, real activity, inflation, and money, *American Economic Review* 71, 545–565.

Fama, Eugene F., 1982, Inflation, output, and money, *Journal of Business* 55, 201–231.

Fama, Eugene F., 1984a, The information in the term structure, *Journal of Financial Economics* 13, 509–528.

Fama, Eugene F., 1984b, Term premiums and default premiums in money markets, *Journal of Financial Economics* 13, 529–546.

Fama, Eugene F., 1990, Term-structure forecasts of interests rates, inflation, and real returns, *Journal of Monetary Economics* 25, 59–76.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Fama, Eugene F., and Michael R. Gibbons, 194, A comparison of inflation forecasts, *Journal of Monetary Economics* 13, 327–348.

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

Farmer, Roger E. A., 1997, Money in a real business cycle model, *Journal of Money, Credit and Banking* 29, 568–611.

Ferson, Wayne E., Sergei Sarkissian, and Timothy Simin, 1998, Spurious regressions in financial economics?, Working Paper, University of Washington.

Friedman, Benjamin M., and Kenneth N. Kuttner, 1996, A price target for U.S. monetary policy? Lessons from experience with money growth targets, *Brookings Papers on Economic Activity* 0, 77–125.

Granger, Clive W. J., and Paul Newbold, 1974, Spurious regression in economics, *Journal of Econometrics* 4, 111–120.

Gregory, Allan W., and Graham M. Voss, 1991, The term structure of interest rates: Departures from time-separable expected ultility, *Canadian Journal of Economics* 24, 923–939.

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

Heaton, John C., 1995, An empirical investigation of asset pricing with temporally dependent preference specifications, *Econometrica* 63, 681–717.

King, Robert G., and Mark W. Watson, 1996, Money, prices, interest rates and the business cycle, *Review of Economics and Statistics* 78, 35–53.

Labadie, Pamela, 1989, Stochastic inflation and the equity premium, *Journal of Monetary Economics* 24, 277–298.

Lucas, Robert E., Jr., 1990, Liquidity and interest rates, *Journal of Economic Theory* 50, 237–264.

Lutz, Friedrich, 1940, The structure of interest rates, *Quarterly Journal of Economics* 55, 36–63.

McFadden, Daniel, 1989, A method of simulated moments for estimation of discrete response models without numerical integration, *Econometrica* 57, 995–1026.

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

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Ogaki, Masao, 1993, Generalized method of moments: Econometric applications, in G. S. Maddala, C. Radhakrishna Rao, and Hrishikesh D. Vinod, ed., *Handbook of Statistics, Vol. II*, 455–488, Amsterdam: North-Holland.

Pennacchi, George G., 1991, Identifying the dynamics of real interest rates and inflation: Evidence using survey data, *Review of Financial Studies* 4, 53–86.

Roll, Richard, 1970, *The Behavior of Interest Rates: An Application of the Efficient Market Model to U.S. Treasury Bills*, Basic Books, New York, New York.

Salyer, Kevin D., 1995, Habit persistence and the nominal term premium puzzle: A partial resolution, *Economic Inquiry* 33, 672–691.

Shafir, Elder and Peter Diamond, and Amos Tversky, 1997, Money illusion, *The Quarterly Journal of Economics* 112, 341–374.

Shiller, Robert J., 1990, The term structure of interest rates, in Frank H. Hahn and Benjamin M. Friedman, ed., *Handbook of Monetary Economics*, 627–672, Elsevier Science, New York, New York.

Shiller, Robert J., 1996, Why do people dislike inflation?, Working Paper, NBER #5539.

Stambaugh, Robert F., 1998, Predictive regressions, Working Paper, University of Pennsylvania.

Startz, Richard, 1982, Do forecast errors on term premia really make the difference between long and short rates?, *Journal of Financial Economics* 10, 323–329.

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

Weil, Philippe, 1989, The equity premium puzzle and the risk-free rate puzzle, *Journal of Monetary Economics* 24, 401–421.

Yule, G. Udny, 1926, Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series, *Journal of the Royal Statistical Society* 89, 1–69.

A Appendix

This appendix provides further details on our simulated least squares estimation approach. Since the estimator is far less specialized than the discussion in Section 3.1 suggests, we first formalize our approach again in a more general setting. Then, we derive and discuss its asymptotic properties.

A.1 Simulated Least Squares Estimation

Let x_t be a vector of state variables with transition function h, such that:

$$x_{t+1} = h(x_t, \varepsilon_{t+1}, \gamma), \tag{35}$$

where $\gamma \in \Gamma$ is a vector of structural parameters and $\{\varepsilon_{t+1}\}$ is a sequence of i.i.d. innovations.

The prices f of a set of nominal discount bonds are equal to the conditional expectation of the nominal pricing kernel g:

$$f(x_t, \beta, \gamma) = \mathrm{E}\big[g(w_{t+n}, \beta)\big|\ x_t, \gamma\big], \tag{36}$$

where $w_{t+n} = [x'_{t+1}, x'_{t+2}, \ldots, x'_{t+n}]'$, for some positive integers n (i.e. bond maturities). The conditional expectation is taken with respect to the joint distribution of the random vector $\eta_{t+n} = [\varepsilon'_{t+1}, \varepsilon'_{t+2}, \ldots, \varepsilon'_{t+n}]'$, given that w_{t+n} is generated by the recursion (35). $\beta \in \mathcal{B}$ is a vector of preference parameters.

We consider an econometric model:

$$y_t = f(x_t, \beta, \gamma) + u_t, \tag{37}$$

where $\{y_t\}_{t=1}^T$ are vectors of observed bond prices. The errors $\{u_t\}_{t=1}^T$ can be interpreted either as measurement errors or as pricing errors.

For notational convenience, define $z_t = [y'_t, x'_t]'$, $z = [z'_1, z'_2, \ldots, z'_T]'$, and $x = [x'_1, x'_2, \ldots, x'_T]'$. Furthermore, let $\theta \in \Theta$ denote the distinct elements of β and γ.

The econometrician's objective function is:

$$Q_T(\theta) = \sum_{t=1}^{T} q(z_t, \theta), \tag{38}$$

where

$$q(z_t,\theta) = \big[y_t - f(x_t,\beta,\gamma)\big]' W_t \big[y_t - f(x_t,\beta,\gamma)\big], \tag{39}$$

and $\{W_t\}_{t=1}^T$ is a sequence of exogenous weighting matrices.

We generate an i.i.d. sequence of vectors $\{\hat\eta_t = [\hat\eta'_{t+n,1}, \hat\eta'_{t+n,2}, \ldots, \hat\eta'_{t+n,2S}]'\}_{t=1}^T$, where each subvector $\hat\eta_{t+n,i} = [\hat\varepsilon'_{t+1,i}, \hat\varepsilon'_{t+2,i}, \ldots, \hat\varepsilon'_{t+n,i}]'$ is formed with independent draws $\{\hat\varepsilon_{\tau,i}\}_{\tau=t+1}^{t+n}$ that are identically distributed as $\{\varepsilon_\tau\}_{\tau=t+1}^{t+n}$. Then, for a given set of structural parameters γ, we recursively compute:

$$\hat w_{t+n,i} \equiv \big[\hat x'_{t+1,i}, \hat x'_{t+2,i}, \ldots, \hat x'_{t+n,i}\big]', \tag{40}$$

using $\hat x_{\tau+1,i} = h(\hat x_{\tau,i}, \hat\varepsilon_{\tau+1,i}, \gamma)$ with the initial condition $\hat x_{t,i} = x_t$, for $\tau = \{t, t+1, \cdots, t+n-1\}$ and $i = \{1, 2, \cdots, 2S\}$. In this way, we construct for every date t a set of $2S$ independent random vectors $\{\hat w_{t+n,i}\}_{i=1}^{2S}$, each of which is independent and identically distributed as w_{t+n}, conditional on the state vector x_t and the structural parameters γ.

Finally, we define the simulated least squares estimator $\hat\theta_T$ of θ as:

$$\hat\theta_T = \arg\min_{\theta\in\Theta} \hat Q_T(\theta), \tag{41}$$

where

$$\hat Q_T(\theta) = \sum_{t=1}^T \hat q(z_t,\theta) \tag{42}$$

and

$$\hat q(z_t,\theta) = \Big[y_t - \frac{1}{S}\sum_{s=1}^S g(\hat w_{t+n,s},\beta)\Big]' W_t \Big[y_t - \frac{1}{S}\sum_{s=S+1}^{2S} g(\hat w_{t+n,s},\beta)\Big]. \tag{43}$$

Intuitively, the reason for using summations over two different sets of simulations is that conditional on z_t, the first set of pricing errors $y_t - 1/S\sum_{s=1}^S g(\hat w_{t+n,s},\beta)$ is independent of the second set $y_t - 1/S\sum_{s=S+1}^{2S} g(\hat w_{t+n,s},\beta)$. This mean that $\hat q(z_t,\theta)$ is an unbiased estimate of $q(z_t,\theta)$, and that, as a result, the law of large numbers controls the approximation errors introduced by the simulations across the time-series dimension of the data.

Formally, the errors:

$$\delta(z_t, \theta) \equiv \hat{q}(z_t, \theta) - q(z_t, \theta) \tag{44}$$

form a *martingale difference sequence.* This allows us to show that for any θ, the two objective functions $\hat{Q}_T(\theta)/T$ and $Q_T(\theta)/T$ converge to the same value as $T \to \infty$.

When the dimension of the parameter vector θ is large, minimizing the function $\hat{Q}_T(\theta)$ can be problematic. This motivates the following two-stage procedure. In the first stage, we obtain a consistent estimate $\hat{\gamma}_T$ of the structural parameters γ (by maximum likelihood, for example). In the second stage, we then estimate the preference parameters β using our simulated least squares approach:

$$\hat{\beta}_T = \arg\max_{\beta \in \mathcal{B}} \hat{Q}_T(\beta, \hat{\gamma}_T), \tag{45}$$

where

$$\hat{Q}_T(\beta, \hat{\gamma}_T) = \sum_{t=1}^{T} \hat{q}(z_t, \beta, \hat{\gamma}_T), \tag{46}$$

and

$$\hat{q}(z_t, \beta, \hat{\gamma}_T) = \Big[y_t - \frac{1}{S}\sum_{s=1}^{S} g(\hat{w}_{t+n,s}, \beta)\Big]' W_t \Big[y_t - \frac{1}{S}\sum_{s=S+1}^{2S} g(\hat{w}_{t+n,s}, \beta)\Big]. \tag{47}$$

A.2 Asymptotics

We now provide a set of assumptions that are sufficient to establish the consistency and asymptotic normality of the two-stage simulated least squares estimator. We do not attempt to find the most general conditions. Instead, we want to illustrate that the asymptotics of our estimator require only standard assumptions.

The first set of assumptions is:

(A1) $\{z_t\}$ is a stationary and ergodic sequence.

(A2) (i) $\hat{\gamma}_T \xrightarrow{a.s.} \gamma_o$,

(ii) $\mathrm{E}\big[u_t \big| x\big] = 0$, and

(iii) $\mathrm{E}\big|f(x_t, \beta, \gamma_o) - f(x_t, \beta_o, \gamma_o)\big| \neq 0$ for $\beta \neq \beta_o$,

where γ_0 and β_0 denote the *true* parameter values.

Assumption (A2.i) is a prerequisite for the two-stage procedure. It allows us to identify the structural parameters in the first stage. The assumption (A2.ii) is standard for least squares estimators. Finally, assumption (A2.iii) is required for identification.

Proposition 1: Suppose that $\mathcal{B}$ and Θ are compact and that θ_o is interior to Θ. Also, let (A1), (A2), and the following conditions hold:

(i) $\hat{q}(z_t\theta)$ is measurable in $[z_t, \hat{\eta}_{t+n}]$ and is continuously differentiable in θ,

(ii) $\mathrm{E}\Big[\sup_{\theta\in\Theta}\big|\hat{q}(z_t,\theta)\big|\Big] < \infty$, and

(iii) $\mathrm{E}\Big[\sup_{\theta\in\Theta}\big|\hat{q}_\gamma(z_t,\theta)\big|\Big] < \infty$,

where subscripts denote partial derivatives. Then,

$$\hat{\beta}_T \xrightarrow{a.s.} \beta_o. \tag{48}$$

When $\hat{\gamma}_T$ and the least squares estimator of β are asymptotically normally distributed, we can also establish asymptotic normality of our two-stage simulated least squares estimator. For this, we need to add the following set of assumptions:

(A3) (i) $1/\sqrt{T}\,\big(\hat{\gamma}_T - \gamma_o\big) \xrightarrow{\mathcal{D}} \mathrm{N}(0, D)$,

(ii) $f_{\beta\beta}$ exists and $\mathrm{E}\big[q_{\beta\beta}(z_t,\theta_o)\big]$ is finite and nonsingular, and

(iii) $1/\sqrt{T}\,Q_{\beta,T}(\theta_o) \xrightarrow{\mathcal{D}} \mathrm{N}(0, B)$,

where double-subscripts denote second-order partial derivatives.

Proposition 2: Suppose assumptions (A1), (A2), and (A3) hold. Also, let the three conditions in Proposition 1 and the following two conditions hold:

(i) $\hat{q}(z_t,\theta)$ is twice continuously differentiable in θ, and

(ii) $E\Big[\sup_{\theta\in\Theta}\big|\hat{q}_{\theta\theta}(z_t,\theta)\big|\Big] < \infty$ and $\hat{q}_\theta(z_t,\theta_o)$ has finite second moments.

Then,

$$1/\sqrt{T}\,\big(\hat{\beta}_T - \beta_o\big) \xrightarrow{\mathcal{D}} \mathrm{N}\big(0, A^{-1}(B + \Delta + CDC')A^{-1}\big), \tag{49}$$

where $A = \mathrm{E}\big[q_{\beta\beta}(z_t,\theta_o)\big]$, $\Delta = \mathrm{E}\big[\delta_\beta(z_t,\theta_o)\delta_\beta(z_t,\theta_o)'\big]$, and $C = \mathrm{E}\big[\hat{q}_{\beta\gamma}(z_t,\theta_o)\big]$.

A.3 Proofs

Poof of Proposition 1: By the mean value theorem, there exists a $\tilde{\gamma}_T$ between $\hat{\gamma}_T$ and γ_o, such that:

$$\frac{1}{T}\hat{Q}_T(\beta,\hat{\gamma}_T) = \frac{1}{T}\hat{Q}_T(\beta,\gamma_o) + \frac{1}{T}\hat{Q}_{\gamma,T}(\beta,\tilde{\gamma}_T)(\hat{\gamma}_T - \gamma_o). \tag{50}$$

It is straightforward to verify the sufficient conditions for Andrews' (1987) uniform law of large numbers, which in turn yields:

$$\lim_{T\to\infty}\sup_{\theta\in\Theta}\left|\frac{1}{T}\hat{Q}_T(\beta,\gamma) - \mathrm{E}\big[\hat{q}(z_t,\beta,\gamma)\big]\right| \overset{a.s.}{=} 0 \tag{51}$$

$$\lim_{T\to\infty}\sup_{\theta\in\Theta}\left|\frac{1}{T}\hat{Q}_{\gamma,T}(\beta,\gamma) - \mathrm{E}\big[\hat{q}_\gamma(z_t,\beta,\gamma)\big]\right| \overset{a.s.}{=} 0. \tag{52}$$

This guarantees that:

$$\lim_{T\to\infty}\sup_{\beta\in\mathcal{B}}\left|\frac{1}{T}\hat{Q}_T(\beta,\hat{\gamma}_T) - \mathrm{E}\big[\hat{q}(z_t,\beta,\gamma_o)\big]\right| \overset{a.s.}{=} 0. \tag{53}$$

To complete the proof, notice that assumption (A2.ii) implies:

$$\begin{aligned}\mathrm{E}\big[\hat{q}(z_t,\beta,\gamma_o)\big] = \mathrm{E}\big[u_t'W_tu_t\big] + \dots \\ \mathrm{E}\Big[\big[f(x_t,\beta,\gamma_o) - f(x_t,\beta_o,\gamma_o)\big]'W_t\big[f(x_t,\beta,\gamma_o) - f(x_t,\beta_o,\gamma_o)\big]\Big].\end{aligned} \tag{54}$$

Therefore, by assumption (A2.iii), $E\big[\hat{q}(z_t,\beta,\gamma_o)\big]$ attains a unique global minimum at $\beta = \beta_o$.

Proof of Proposition 2: By the mean value theorem:

$$\frac{1}{\sqrt{T}}\hat{Q}_{\beta,T}(\hat{\beta}_T,\hat{\gamma}_T) = \frac{1}{\sqrt{T}}\hat{Q}_{\beta,T}(\beta_o,\hat{\gamma}_T) + \frac{1}{T}\hat{Q}_{\beta\beta,T}(\beta_T^*,\hat{\gamma}_T)\sqrt{T}(\hat{\beta}_T - \beta_o) = 0 \tag{55}$$

and

$$\frac{1}{\sqrt{T}}\hat{Q}_{\beta,T}(\beta_o,\hat{\gamma}_T) = \frac{1}{\sqrt{T}}\hat{Q}_{\beta,T}(\beta_o,\gamma_o) + \frac{1}{T}\hat{Q}_{\beta\gamma,T}(\beta_o,\gamma_T^*)\sqrt{T}(\hat{\gamma}_T - \gamma_o) = 0, \tag{56}$$

for some β_T^* between $\hat{\beta}_T$ and β_o and some γ_T^* between $\hat{\gamma}_T$ and γ_o. Andrews' uniform law of large numbers then implies that:

$$\frac{1}{T}\hat{Q}_{\beta\beta,T}(\beta_T^*,\hat{\gamma}_T) \xrightarrow{a.s.} A \quad\text{and}\quad \frac{1}{T}\hat{Q}_{\beta\gamma,T}(\beta_o,\gamma_T^*) \xrightarrow{a.s.} C. \tag{57}$$

Therefore, the asymptotic distribution of $\hat{Q}_{\beta,T}(\beta_o, \hat{\gamma}_T)/\sqrt{T}$ is the same as the asymptotic distribution of $\hat{Q}_{\beta,T}(\beta_o, \gamma_o)/\sqrt{T} + \sqrt{T}C(\hat{\gamma}_T - \gamma_o)$.

Let $\mathcal{F}_t = \sigma\{z_s, \eta_{s+n}; s \leq t\}$ and let $\hat{\mathcal{F}}_t = \sigma\{z_t, \mathcal{F}_{t-1}\}$. Since

$$\mathrm{E}\big[\delta_\beta(z_t, \theta)\big|\hat{\mathcal{F}}_t\big] = \mathrm{E}\big[\delta_\beta(z_t, , \theta)\big|z_t\big] = 0, \tag{58}$$

we have

$$\mathrm{E}\big[\delta_\beta(z_t, \theta)\big|\mathcal{F}_{t-1}\big] = \mathrm{E}\Big[\mathrm{E}\big[\delta_\beta(z_t, \theta)\big|\hat{\mathcal{F}}_t\big]\Big|\mathcal{F}_{t-1}\Big] = 0. \tag{59}$$

Thus, $\{\delta_\beta(z_t, \theta)\}$ is martingale difference sequence with respect to $\{\mathcal{F}_t\}$. Billingsley's (1961) central limit theorem then provides:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \delta_\beta(z_t, \theta) \xrightarrow{\mathcal{D}} \mathrm{N}(0, \Delta). \tag{60}$$

Finally, since:

$$\mathrm{E}\big[\delta_\beta(z_t, \theta)\big|z_t\big] = 0, \tag{61}$$

$Q_{\beta,T}(\theta_o)/\sqrt{T}$ and $\sum_{t=1}^{T} \delta_\beta(z_t, \theta_o)/\sqrt{T}$ are uncorrelated. Therefore, we have:

$$\frac{1}{\sqrt{T}} \hat{Q}_{\beta,T}(\beta_o, \gamma_o) \xrightarrow{\mathcal{D}} \mathrm{N}(0, B + \Delta), \tag{62}$$

where we used:

$$\frac{1}{\sqrt{T}} \hat{Q}_{\beta,T}(\theta_o) = \frac{1}{\sqrt{T}} Q_{\beta,T}(\theta_o) + \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \delta_\beta(z_t, \theta_o). \tag{63}$$

Given the fact that:

$$\mathrm{E}\big[\hat{Q}_{\beta,T}(\beta_o, \gamma_o)\big|x\big] = \mathrm{E}\big[Q_{\beta,T}(\beta_o, \gamma_o)\big|x\big] = \mathrm{E}\big[-2\sum_{t=1}^{T} f_\beta(x_t, \beta_o, \gamma_o)' W u_t \big|x\big] = 0, \tag{64}$$

we thus know that $\hat{Q}_{\beta,T}(\beta_o, \gamma_o)/\sqrt{T}$ and $\sqrt{T}(\hat{\gamma}_T - \gamma_o)$ are uncorrelated. The remainder of the proof is straightforward.

Table 1

Consumption Growth and Inflation

This table shows maximum likelihood estimates of a VAR(2) for monthly consumption growth g_t and inflation π_t for different sample periods. In parentheses are asymptotic standard errors.

Panel A: January 1969 - June 1997

	Intercept	g_{t-1}	g_{t-2}	π_{t-1}	π_{t-2}	Residual Covariance Matrix	
Consumption Growth g_t	0.336	-0.308	-0.023	-0.239	-0.082	0.125	
	(0.047)	(0.054)	(0.054)	(0.107)	(0.109)		
Inflation π_t	0.086	0.031	0.033	0.524	0.253	-0.005	0.030
	(0.023)	(0.027)	(0.027)	(0.053)	(0.053)		

Panel B: January 1969 - September 1979

	Intercept	g_{t-1}	g_{t-2}	π_{t-1}	π_{t-2}	Residual Covariance Matrix	
Consumption Growth g_t	0.674	-0.368	-0.223	-0.146	-0.566	0.127	
	(0.109)	(0.087)	(0.088)	(0.191)	(0.190)		
Inflation π_t	0.125	-0.012	0.017	0.457	0.323	-0.010	0.026
	(0.050)	(0.039)	(0.040)	(0.087)	(0.086)		

Panel C: October 1979 - June 1997

	Intercept	g_{t-1}	g_{t-2}	π_{t-1}	π_{t-2}	Residual Covariance Matrix	
Consumption Growth g_t	0.304	-0.328	0.036	-0.431	0.043	0.101	
	(0.053)	(0.069)	(0.068)	(0.126)	(0.130)		
Inflation π_t	0.099	0.036	0.028	0.515	0.183	-0.004	0.031
	(0.028)	(0.037)	(0.037)	(0.068)	(0.070)		

Table 2

Simulated Least Squares Estimates of the Preference Parameters

This table shows simulated least squares estimates of the preference parameters of the model, of Campbell and Cochrane's model, and of the standard power utility model. In parentheses are autocorrelation and heteroscedasticity adjusted asymptotic standard errors.

Panel A: The Model

	α	$\bar{\gamma}$	ϕ	θ	δ
Estimate	1.000	0.512	0.989	14.378	0.999
	–[a]	(0.235)	(0.002)	(3.420)	(0.002)

[a] The constraint $\alpha \geq 1$ is binding

Panel B: Campbell and Cochrane's Model

	α	$\bar{\gamma}$	ϕ	θ	δ
Estimate	1.000	0.451	0.996	0.000	0.999
	–[a]	(0.291)	(0.002)	–	(0.002)

[a] The constraint $\alpha \geq 1$ is binding

Panel C: Standard Power Utility Model

	α	$\bar{\gamma}$	ϕ	θ	δ
Estimate	1.006	–	–	–	0.999
	(0.089)				(0.002)

Table 3

Simulated Least Squares Estimates of the Preference Parameters for Subsamples

This table shows simulated least squares estimates of the preference parameters of the model for different sample periods. In parentheses are autocorrelation and heteroscedasticity adjusted asymptotic standard errors.

Panel A: January 1969 - September 1979

	α	$\bar{\gamma}$	ϕ	θ	δ
Estimate	1.000	0.447	0.987	15.291	0.999
	–[a]	(0.314)	(0.003)	(5.291)	(0.003)

[a] The constraint $\alpha \geq 1$ is binding

Panel B: October 1979 - June 1997

	α	$\bar{\gamma}$	ϕ	θ	δ
Estimate	1.000	0.271	0.990	10.794	0.999
	–[a]	(0.300)	(0.003)	(4.485)	(0.003)

[a] The constraint $\alpha \geq 1$ is binding

Table 4

Time-Varying Risk Aversion

This table describes the time-variation in aggregate risk aversion implied by the model and by Campbell and Cochrane's model. RRA_t, ΔRRA_t, and e_t are the relative risk aversion, its change, and the innovation to log relative risk aversion. X_t/C_t and $\Delta X_t/C_t$ are the reference level and its change relative to consumption. ε_t^g is the innovation to consumption growth g_t. ε_t^π is the innovation to inflation π_t. r_t^1 is the realized real return on a one-month Treasury bill.

Panel A: Summary Statistics

	The Model				Campbell and Cochrane's Model			
	RRA_t	e_t	$100\frac{X_t}{C_t}$	$100\frac{\Delta X_t}{C_t}$	RRA_t	e_t	$100\frac{X_t}{C_t}$	$100\frac{\Delta X_t}{C_t}$
Mean	1.82	0.000	43.43	0.03	1.54	0.000	35.14	0.06
Standard Deviation	0.34	0.025	8.57	1.16	0.03	0.002	1.42	0.04
Max	3.07	0.105	67.39	4.50	1.61	0.007	38.03	0.20
Min	1.41	-0.132	29.20	-4.10	1.50	-0.006	33.08	-0.08
Auto-Correlation	0.98	0.032	0.99	-0.03	0.99	0.009	0.99	-0.06

Panel B: Correlations

	ΔRRA_t	e_t	ε_t^g	ε_t^π	g_t	π_t	r_t^1
			The Model				
RRA_t	0.10	-0.17	-0.02	0.17	-0.09	0.46	-0.17
ΔRRA_t		-0.98	-0.21	0.81	-0.20	0.55	-0.48
e_t			0.22	-0.91	0.22	-0.64	0.52
			Campbell and Cochrane's Model				
RRA_t	0.11	-0.15	-0.15	-0.14	-0.04	-0.42	0.02
ΔRRA_t		-0.99	-0.99	0.08	-0.94	0.07	0.00
e_t			0.99	-0.08	0.94	-0.05	0.00

Table 5

Yields

This table describes the observed yields of bonds with different maturities and the corresponding fitted yields for the model, for Campbell and Cochrane's model, and for the standard power utility model. RRA_t and RRA_t^{cc} are the relative risk aversion implied by the model and by Campbell and Cochrane's model. g_t and π_t are the observed consumption growth and inflation.

Panel A: The Data

	Bond Maturity					
	1-Month	3-Month	6-Month	1-Year	3-Year	5-Year
Mean	6.09	6.80	7.11	7.41	7.86	8.09
Standard Deviation	2.48	2.64	2.66	2.58	2.32	2.19
Auto-Correlation	0.89	0.96	0.97	0.97	0.98	0.98
Corrlations:						
RRA_t	0.75	0.80	0.82	0.83	0.85	0.86
RRA_t^{cc}	-0.48	-0.55	-0.56	-0.58	-0.59	-0.59
g_t	-0.12	-0.13	-0.12	-0.11	-0.08	-0.07
π_t	0.46	0.47	0.48	0.44	0.34	0.29

Panel B: The Model

	Bond Maturity					
	1-Month	3-Month	6-Month	1-Year	3-Year	5-Year
Mean	7.65	7.65	7.63	7.61	7.61	7.60
Standard Deviation	3.06	2.63	2.37	2.08	1.65	1.47
Auto-Correlation	0.72	0.91	0.93	0.95	0.97	0.98
Corrlations:						
RRA_t	0.82	0.90	0.93	0.96	0.98	0.98
g_t	-0.53	-0.24	-0.18	-0.14	-0.11	-0.10
π_t	0.71	0.76	0.72	0.65	0.54	0.51
Observed Yields	0.68	0.78	0.81	0.82	0.85	0.87

Table 5 Continued

Panel C: Campbell and Cochrane's Model

	Bond Maturity					
	1-Month	3-Month	6-Month	1-Year	3-Year	5-Year
Mean	7.07	7.07	7.09	7.07	7.07	7.08
Standard Deviation	1.87	1.21	0.91	0.57	0.19	0.12
Auto-Correlation	0.32	0.74	0.77	0.77	0.73	0.74
Corrlations:						
RRA_t^{cc}	-0.22	-0.36	-0.36	-0.31	0.04	0.25
g_t	-0.79	-0.40	-0.31	-0.28	-0.31	-0.32
π_t	0.71	0.95	0.97	0.97	0.88	0.69
Observed Yields	0.38	0.47	0.47	0.41	0.16	-0.04

Panel D: Standard Power Utility Model

	Bond Maturity					
	1-Month	3-Month	6-Month	1-Year	3-Year	5-Year
Man	8.14	8.14	8.15	8.14	8.12	8.12
Standard Deviation	1.90	1.26	0.96	0.62	0.23	0.14
Auto-Correlation	0.34	0.76	0.80	0.81	0.81	0.81
Corrlations:						
g_t	-0.77	-0.38	-0.29	-0.25	-0.23	-0.23
π_t	0.73	0.95	0.97	0.98	0.98	0.98
Observed Yields	0.40	0.50	0.51	0.47	0.36	0.32

Table 6

Expected Term Premiums

This table describes the expected term premiums generated by the model, by Campbell and Cochrane's model, and by the standard power utility model, for different holding periods and bond maturities. RRA_t and RRA_t^{cc} are the relative risk aversion implied by the model and by Campbell and Cochrane's model. g_t and π_t are the observed consumption growth and inflation.

Panel A: The Model

Holding Period:	3 Months			1 Year			
Bond Maturity:	6-Month	9-Month	1-Year	2-Year	3-Year	4-Year	5-Year
Mean	0.043	0.070	0.091	0.070	0.107	0.143	0.175
Standard Deviation	0.024	0.041	0.056	0.047	0.075	0.112	0.149
Corrlations:							
RRA_t	0.99	0.99	0.99	0.98	0.98	0.99	0.99
g_t	-0.09	-0.09	-0.09	-0.08	-0.07	-0.10	-0.10
π_t	0.41	0.41	0.41	0.42	0.41	0.41	0.40

Panel B: Campbell and Cochrane's Model

Holding Period:	3 Months			1 Year			
Bond Maturity:	6-Month	9-Month	1-Year	2-Year	3-Year	4-Year	5-Year
Mean	-0.003	-0.006	-0.007	-0.006	-0.006	-0.006	-0.005
Standard Deviation	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Corrlations:							
RRA_t^{cc}	0.30	0.43	0.78	0.70	0.80	0.82	0.85
g_t	0.00	0.08	0.02	0.09	0.04	0.01	0.00
π_t	-0.40	-0.43	-0.47	-0.76	-0.71	-0.66	-0.63

Table 6 Continued

Panel C: Standard Power Utility Model

Holding Period:	3 Months			1 Year			
Bond Maturity:	6-Month	9-Month	1-Year	2-Year	3-Year	4-Year	5-Year
Mean	-0.003	-0.006	-0.007	-0.006	-0.007	-0.007	-0.007
Standard Deviation	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 7

Term Premium Regressions

This table presents least squares regressions of realized excess bond returns on the corresponding forward-spot spreads for the observed data and for the fitted data of the model, of Campbell and Cochrane's model, and of the standard power utility model. $h(n{+}12, n\!:\!t{+}12)$ is the realized return of holding an $(n{+}12)$-month bond for one year, $f(n{+}12, n\!:\!t)$ is the one-year forward rate in $n{+}12$ months, and $r(12\!:\!t)$ is the yield of a one-year bond. R^2 is the coefficient of determination of the regression and ρ is the first-order autocorrelation of the residuals. In parentheses are autocorrelation and heteroscedasticity adjusted standard errors.

Panel A: The Data

$h(n+12, n:t+12) - r(12:t)$	Intercept	$f(n+12, n:t) - r(12:t)$	R^2	ρ
$n = 12$	0.001	0.910	12.4	0.89
	(0.003)	(0.250)		
$n = 24$	-0.001	1.227	14.0	0.90
	(0.005)	(0.312)		
$n = 36$	-0.004	1.492	15.4	0.89
	(0.007)	(0.380)		
$n = 48$	0.001	1.039	5.1	0.90
	(0.008)	(0.472)		

Panel B: The Model

$h(n+12, n:t+12) - r(12:t)$	Intercept	$f(n+12, n:t) - r(12:t)$	R^2	ρ
$n = 12$	0.000	0.829	18.6	0.78
	(0.001)	(0.191)		
$n = 24$	0.001	0.942	15.0	0.82
	(0.002)	(0.246)		
$n = 36$	0.001	1.002	12.5	0.84
	(0.003)	(0.292)		
$n = 48$	0.001	1.019	10.3	0.85
	(0.003)	(0.328)		

Table 7 Continued

Panel C: Campbell and Cochrane's Model

$h(n+12, n: t+12) - r(12: t)$	Intercept	$f(n+12, n: t) - r(12: t)$	R^2	ρ
$n = 12$	0.000 (0.001)	0.479 (0.105)	20.0	0.69
$n = 24$	0.000 (0.001)	0.481 (0.107)	20.3	0.69
$n = 36$	0.000 (0.001)	0.483 (0.108)	20.4	0.68
$n = 48$	0.000 (0.001)	0.484 (0.110)	20.4	0.68

Panel D: Standard Power Utility Model

$h(n+12, n: t+12) - r(12: t)$	Intercept	$f(n+12, n: t) - r(12: t)$	R^2	ρ
$n = 12$	0.000 (0.001)	0.478 (0.107)	19.8	0.70
$n = 24$	0.000 (0.001)	0.480 (0.110)	19.9	0.70
$n = 36$	0.000 (0.001)	0.480 (0.113)	20.0	0.70
$n = 48$	0.000 (0.001)	0.480 (0.113)	20.0	0.70

Table 8
Correlations of Excess Returns and Term Premiums with Forward-Spot Spreads

noindent

This table shows the correlations of the realized excess bond returns with the corresponding forward-spot spreads for the observed data and for the fitted data of the model. For the model, it also shows the correlations of the expected term premiums with the forward-spot spreads. $h(n{+}12, n\,{:}\,t{+}12)$ is the realized return of holding an $(n{+}12)$-month bond for one year, $f(n{+}12, n\,{:}\,t)$ is the one-year forward rate in $n{+}12$ months, and $r(12\,{:}\,t)$ is the yield of a one-year bond.

Panel A: The Data

	$f(n+12, n:t) - r(t:12)$			
	$n = 12$	$n = 24$	$n = 36$	$n = 48$
$h(n+12, n:t+12) - r(12:t)$	0.35	0.38	0.39	0.23

Panel B: The Model

	$f(n+12, n:t) - r(t:12)$			
	$n = 12$	$n = 24$	$n = 36$	$n = 48$
$h(n+12, n:t+12) - r(12:t)$	0.43	0.39	0.36	0.33
$\mathrm{E}_t\big[h(n+12, n:t+12) - r(12:t)\big]$	-0.65	-0.72	-0.77	-0.78

Table 9

Feasible Generalized Least Squares
Term Premium Regressions

This table presents feasible generalized least squares regressions of realized excess bond returns on the corresponding forward-spot spreads for the observed data. The least squares residuals are assumed to follow either a random walk or an AR(1) process. $h(n{+}12, n\!:\!t{+}12)$ is the realized return of holding an $(n{+}12)$-month bond for one year, $f(n{+}12, n\!:\!t)$ is the one-year forward rate in $n{+}12$ months, and $r(12\!:\!t)$ is the yield of a one-year bond. R^2 is the coefficient of determination of the regression and ρ is the first-order autocorrelation of the residuals. In parentheses are autocorrelation and heteroscedasticity adjusted standard errors.

Panel A: Random Walk Residuals

$h(n+12, n:t+12) - r(12:t)$	Intercept	$f(n+12, n:t) - r(12:t)$	R^2	ρ
$n = 12$	0.000	-0.078	0.2	0.07
	(0.001)	(0.107)		
$n = 24$	0.000	-0.154	0.3	0.12
	(0.001)	(0.154)		
$n = 36$	0.000	-0.037	0.0	0.10
	(0.001)	(0.160)		
$n = 48$	0.000	-0.366	0.7	0.02
	(0.001)	(0.205)		

Panel B: AR(1) Residuals

$h(n+12, n:t+12) - r(12:t)$	Intercept	$f(n+12, n:t) - r(12:t)$	R^2	ρ
$n = 12$	0.001	-0.036	0.0	0.14
	(0.001)	(0.110)		
$n = 24$	0.001	-0.082	0.1	0.18
	(0.001)	(0.156)		
$n = 36$	0.001	0.041	0.0	0.17
	(0.001)	(0.163)		
$n = 48$	0.002	-0.303	0.7	0.08
	(0.001)	(0.208)		

Table 10

Generalized Method of Moments Estimates of the Preference Parameters

This table shows generalized method of moments estimates of the preference parameters of the model, of Campbell and Cochrane's model, and of the standard power utility model for different sets of lagged instruments. It also describes the time-variation in aggregate risk aversion implied by the models. The first set of instruments consists of a constant, yields on the one- and six-months Treasury bills, and yields on the one-, three-, and five-year Treasury bonds. The second set of instruments consists of a constant, yields on the one-month Treasury bill, yields on the one- and five-year Treasury bonds, consumption growth, and inflation. In parentheses are autocorrelation and heteroscedasticity adjusted asymptotic standard errors.

Panel A: The Model

	α	$\bar{\gamma}$	ϕ	θ	δ	J-Statistic
Estimate Instruments I	1.000	0.339	0.998	14.615	0.998	20.06
	–[a]	(0.158)	(0.003)	(5.319)	(0.000)	(0.950)
Estimate Instruments II	1.000	0.418	0.994	11.948	0.999	26.10
	–[a]	(0.135)	(0.002)	(3.312)	(0.000)	(0.759)

[a] The constraint $\alpha \geq 1$ is binding

	Instruments I				Instruments II			
	Mean	Standard Deviation	Max	Min	Mean	Standard Deviation	Max	Min
RRA_t	1.57	0.21	2.22	1.29	1.63	0.19	2.21	1.36

Table 10 Continued

Panel B: Campbell and Cochrane's Model

	α	$\bar{\gamma}$	ϕ	θ	δ	J-Statistic
Estimate Instruments I	1.000	1.046	0.939	0.000	0.998	20.61
	–[a]	(0.123)	(0.016)	–	(0.000)	(0.954)
Estimate Instruments II	1.000	0.927	0.932	0.000	0.999	20.82
	–[a]	(0.123)	(0.014)	–	(0.000)	(0.951)

[a] The constraint $\alpha \geq 1$ is binding

	Instruments I				Instruments II			
	Mean	Standard Deviation	Max	Min	Mean	Standard Deviation	Max	Min
RRA_t	2.85	0.08	3.08	2.70	2.53	0.06	2.69	2.42

Panel C: Standard Power Utility Model

	α	$\bar{\gamma}$	ϕ	θ	δ	J-Statistic
Estimate Instruments I	1.004	–	–	–	0.999	21.48
	(0.062)				(0.000)	(0.953)
Estimate Instruments II	1.005	–	–	–	0.999	21.42
	(0.063)				(0.000)	(0.954)

Table 11

Generalized Method of Moments Estimates of the Preference Parameters with Stock Returns

This table shows generalized method of moments estimates of the preference parameters of the model with stock returns for different data frequencies. It also describes the time-variation in aggregate risk aversion implied by the models. The returns are on on the ten NYSE size-decile portfolios. The lagged instruments are a constant, the aggregate dividend to price ratio, the default premium, the term premium, the yield on a one-month Treasury bill, and the CRSP market return. In parentheses are autocorrelation and heteroscedasticity adjusted asymptotic standard errors.

	α	$\bar{\gamma}$	ϕ	θ	δ	J-Statistic
Estimate Monthly Data	1.000	2.010	0.935	5.070	0.990	24.87
	$-^a$	(0.807)	(0.054)	(2.639)	(0.003)	(0.999)
Estimate Annual Data	1.000	1.906	0.868	5.693	0.993	25.81
	$-^a$	(0.271)	(0.019)	(1.504)	(0.000)	(0.999)

[a] The constraint $\alpha \geq 1$ is binding

	Monthly Data				Annual Data			
	Mean	Standard Deviation	Max	Min	Mean	Standard Deviation	Max	Min
RRA_t	7.83	2.17	20.47	5.29	6.83	1.05	11.72	5.11

Figure 1

Time-Varying Relative Risk Aversion

This figure shows the time-series of relative risk aversion implied by the model (solid line) and by Campbell and Cochrane's model (dashed line).