

State Uncertainty Aversion and the Term Structure of Interest Rates

Frode Brevik*
VU Amsterdam

This version: Feb 2008
(First version: Sep 2007)

Abstract

This paper proposes a novel explanation for the empirical finding that yields on risk-free bonds are increasing with their maturity (the term premium). The key ingredient in the explanation is that investors not only dislike risk, but also dislike uncertainty about the current trend growth rate of the economy. The model setup is one where investors observe consumption growth rates and use these observations to estimate the current level of a mean reverting trend growth rate. At a given point in time, uncertainty about the state is given by the variance of the estimate. Disliking uncertainty, investors bias their estimate of the current trend downwards. On average this lowers short term interest rates relative to long run interest rates. The model can account quantitatively for the observed term premium in the US data and correctly predicts the flattening of the real yield curve since the early nineties.

*Department of Finance, VU Amsterdam, De Boelelaan 1105, 1081HV Amsterdam, Netherlands. Email: fbrevik@gmail.com. The research in this paper is financed by a postdoctoral fellowship from the Swiss National Science Fund. I thank Thomas Sargent, Paul Söderlind, Fabio Trojani, and seminar participants at Brown University, Cambridge University, New York University, VU Amsterdam, University of Leicester, and University of St. Gallen for helpful discussions and comments and the NYU Economics department for its hospitality.

1 Introduction

The premise of this paper is that investors dislike uncertainty about the prevailing state of the economy and that aversion carries over to the relative prices of risk-free bonds of different maturities. The fundamental question asked is whether uncertainty aversion can help account for the term premium puzzle: the robust empirical finding that long term bonds on average pay higher yields than short term bonds.

As shown by Backus et al. (1989) a standard model with time varying expected consumption growth rates and power utility actually predicts that the average yield curve should be downward sloping. The reason is that spot interest rates—which are prices for shifting consumption from future periods to today—are high whenever current consumption is scarce relative to its expected future level. Bad news about consumption growth causes interest rates to decrease and bonds to rally. This makes real bonds a hedge against bad consumption growth rate news. The longer the maturity of the bond, the better the hedge. As long as this is the only mechanism that acts on the unconditional yield curve, it will be downward sloping.

The traditional explanation for the term premium is that it is a compensation for inflation risk that investors require to hold nominal bonds. However, structural models need to press this mechanism very hard to match the empirical term premium. Piazzesi and Schneider (2006) shows that with Epstein-Zin preferences, a risk aversion parameter of 210 is needed to match the term premium in the US data for realistic correlations between consumption innovations and inflation. This explanation is also at odds with the term premium apparent in the short sample we have on inflation indexed US government bonds.

This paper shows that state uncertainty aversion in the flavor recently introduced by Hansen and Sargent (2007) by itself produces a real term premium. It counters the effect identified by Backus et al. (1989) by pushing down the short end of the yield relative to the long end curve. If investors are sufficiently averse to state uncertainty, a positive real term premium results. The mechanism is the following: Investors' expected utility depends on a hidden growth state. Uncertain about the current realization of this state, investors make a conservative adjustment to their estimate by biasing it downward. The size of the adjustment is proportional to the variance of the estimate. This downward bias carries over to expectations for future periods, decaying at the same rate as the growth state converges to its mean. As a result, adjusted expected consumption growth rates will on average be increasing with the horizon. Because the interest rate for a given maturity

is an increasing function of the expected average growth rate between today and that maturity, the short end of the yield curve is shifted down more than the long end.

Apart from the preference specification, the model setting is a familiar one from the literature on asset pricing in Hidden Markov Models:¹ Consumption growth rates are the sum of independent shocks (increments to a Brownian motion) and a hidden, time varying trend. Using their knowledge on the hyper-parameters of the economy and the record of historic consumption growth rates, investors form posterior beliefs about the current growth state.

The growth state is specified as an Ornstein-Uhlenbeck process, yielding a well known filtering problem. If the process parameters are constant, the variance of the estimate of the state converges quickly to a constant steady-state value. This would imply a time-invariant shift of the yield curve, so it would be hard to test the plausibility of the mechanism causing it. Fortunately, from the perspective of testing the model, the process parameters do not appear to be constant over the whole sample period. In particular, consumption volatility is markedly lower in the late part of the sample. Moreover, the lower volatility appears to be due to regime switching which makes it straightforward to model. Lower consumption volatility means that consumption growth rates will cling closer to the time varying trend and inference about the growth state will be preciser. This yields a testable prediction of the model: in the latter part of the sample, state uncertainty is lower which means that the adjustment of the term structure should also be smaller. I calibrate the model to match the term structure in the early part of the sample and use the late part of the sample as a test of the proposed mechanism.

By proposing an alternative explanation for the term premium puzzle, this paper adds to recent research trying to explain it with consumption based models. Two recent papers make contributions stressing mechanisms that are different and partly complementary to the one stressed here. Piazzesi and Schneider (2006) argue that surprise inflation predicts low consumption growth and that the payoff of long term nominal bonds is particularly sensitive to inflation shocks. This generates a positive nominal term premium. A common feature of their model and the one presented here is that investors care about the timing of resolution of uncertainty. In their model it is introduced with Epstein-Zin preferences. As they

¹See e.g. Detemple (1986), Dothan and Feldman (1986), David (1997), Veronesi (1999, 2000), Brennan and Xia (2001), Lettau et al. (2006), and many more.

show, Epstein-Zin preferences predicts a real term discount, so the term premium they find is nominal one. In the model of Wachter (2006), surprise inflation is also a predictor of low consumption growth. In addition, she modifies the external habit model of Campbell and Cochrane (1999) to generate countercyclical real interest rates. This causes the holding returns also for long run real bonds to be procyclical, so investors will require a higher real rate to hold them.

Throughout the paper, I assume investors have perfect information on the underlying structure of the economy. That is, they have rational expectations. This allows me to distill out the pure effect of state uncertainty aversion. The obvious cost is that effects of parameter uncertainty and learning are lost.

The paper is organized as follows: The model economy and the filtering problem of the investors is presented and discussed in section 2. Section 3 derives equilibrium interest rates in this economy. Section 4, estimates the parameters of the endowment process on quarterly US consumption growth rates. In section 5, investors' preference parameters are calibrated to match the average US term structure in the early part of the sample and the predictions for the whole sample are discussed. Section 6 concludes.

2 An exchange economy with a hidden growth state

2.1 Endowment process

Real consumption C evolves according to the stochastic differential equation

$$dC = (g + \mu)C \, dt + \sqrt{v(t)}C \, dB_c, \quad (1a)$$

where B_c is a standard Brownian motion. Investors do not directly observe $g(t)$ but know the parameters of the process it follows. The dynamics of g is governed by the Ornstein-Uhlenbeck process

$$dg = -kg \, dt + \sqrt{w(t)} \, dB_g,$$

where B_g is a standard Brownian motion which is independent of B_c . We denote the precision of dC as a signal on g by h :

$$h = \frac{1}{\sqrt{v}} \quad (1b)$$

The variances $w(t)$ and $v(t)$ follow a joint 2 state Markov chain with a constant transition intensity λ from both states:

$$\lim_{h \rightarrow 0} \frac{\Pr(v(t+h) = v_i \mid v(t) = v_j)}{h} = \begin{cases} 1 - \lambda & i = j \\ \lambda & i \neq j \end{cases}$$

For identification, let $v_1 \geq v_2$, so that state 1 is the high volatility state.

2.2 Dynamics of agents' beliefs (innovations representation)

2.2.1 The Kalman-Busy filter

Let \mathcal{F}_t denote investors information set at time t . It is the σ -algebra generated by the whole history of consumption growth rates up to time t . We denote the expectation of $g(t)$ conditional on \mathcal{F}_t by $m(t)$. Assuming the prior for $g(0)$ is Gaussian $N(m(0), \gamma(0))$, then, by theorem 12.1 in Lipster and Shiryaev (2000), the conditional distribution of the current growth rate remains Gaussian over time. Its mean $m(t)$ and variance $\gamma(t)$ are determined by

$$dm = -km \, dt + \gamma h \, d\tilde{B} \quad (2a)$$

$$d\gamma = w - 2k\gamma - \frac{\gamma^2}{v} \quad (2b)$$

where

$$d\tilde{B} = h \left(\frac{dC}{C} - (m + \mu) \right)$$

As long as the volatility state does not switch, γ converges faster than geometrically to the positive root of (2b) at $d\gamma = 0$:

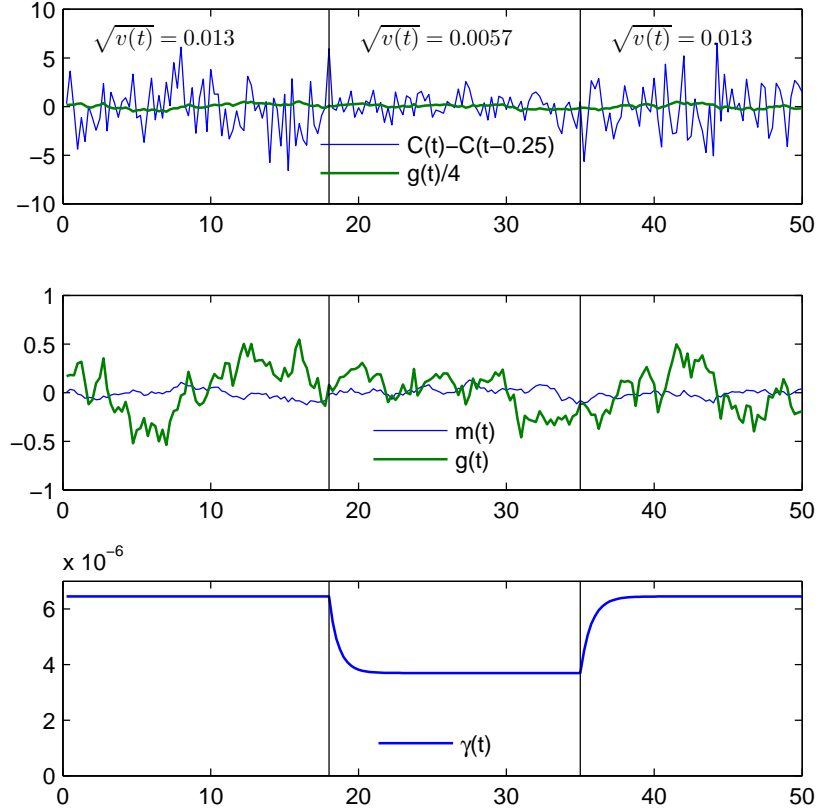
$$\gamma_{ss} = \sqrt{k^2 v^2 + vw} - kv \quad (3)$$

The steady state γ is increasing in v , which reflects that the higher the variance v , the noisier a signal consumption growth rates are on the underlying trend growth rate of the economy. The uncertainty about the trend growth rate is also higher when the underlying trend fluctuates more. The trend is more volatile, the more persistent it is (the lower k) and the more volatile its innovation terms are (the higher w). Notice that equation (3) implies that γ_{ss} is bounded above by the larger of v and w :

$$\gamma_{ss} = \sqrt{k^2 v^2 + vw} - kv \leq \sqrt{k^2 v^2} + \sqrt{vw} - kv = \sqrt{vw} \leq \max(v, w). \quad (4)$$

Figure 1: Illustration of filtering

At the points of time marked by vertical lines, there is a regime switch in the volatility state. In the upper panel, the dotted line gives the quarterly consumption growth rate ($C(t + 0.25) - C(t)$) and the solid line gives the (unobserved) mean growth rate conditional on the hidden growth state of the economy. In the low volatility state, realized consumption growth rates track the hidden state more closely. This pulls $m(t)$ closer to $g(t)$ (middle panel) and reduces $\gamma(t)$ (lower panel).



Following a period of initial convergence after the system is initialized, $g(t)$ will fluctuate between the steady state γ 's of the two volatility states. In section 3, this feature of $\gamma(t)$ is used to compute tight error bounds for the interest rate equation.

Figure 1 illustrates the workings of the Kalman-Busy filter on a simulated time series. The economy is initially in a high volatility state, after 9 years, it switches to the low volatility state. After 17 years it makes an-

other transition to the low volatility state. The top panel plots quarterly consumption growth rates against the underlying growth state of the economy. In the high volatility state, the underlying growth state is relatively well hidden by the shocks to the consumption process so the estimate of the growth state does not catch up very well with changes to the hidden state, but stays close to the long run mean (middle panel). After the switch to the low volatility state, it is easier to estimate the hidden state, and the estimate tracks the hidden value more closely. The bottom panel plots the variance of the estimate ($\gamma(t)$), as it moves with the volatility regime.

2.3 Preferences

The way aversion to state uncertainty is introduced follows Hansen and Sargent (2007). Without state uncertainty, the representative investor is a log-utility maximizer, yielding the value function:

$$\begin{aligned} V(c(t), g(t)) &= E \left[\rho \int_0^\infty e^{-\rho s} c(t+s) ds \mid c(t), g(t) \right] \\ &= c(t) + \frac{1}{\rho} \mu + \frac{1}{\rho + k} g(t) \end{aligned} \quad (5)$$

Where ρ denotes the time discount rate and $c(t) = \log C(t)$. The time discount rate has to be strictly positive to ensure finite utility. Investors dislike any uncertainty about $g(t)$. When $g(t)$ is not observed the investor's value function is found by applying a risk sensitivity operator T^2 . This operator is designed to make a valuation of future consumption that is cautious with respect to the distribution of $g(t)$ conditional on $m(t)$ and $\gamma(t)$:

$$\begin{aligned} \tilde{V}(c, m, \gamma, v) &= T^2[V(c, m, \gamma, v) \mid \gamma, v] \\ &= -\theta \log \int \exp \left(-\frac{V(c, g)}{\theta} \right) \phi(g \mid m, \gamma) dg \\ &= V(c, m, v) - \frac{1}{\theta} \frac{1}{(k + \rho)^2} \frac{\gamma}{2} \end{aligned} \quad (6)$$

Where $\phi(g|m, \gamma)$ is the current posterior beliefs of the investors about the hidden state of the economy. (Which is a normal with mean m and variance γ .) A mathematically equivalent representation of the T^2 -operator is

$$\tilde{V}(c, m, \gamma) = \min_{h(g) \geq 0, \int h(g) \phi(g|m, \gamma) = 1} \int (V(c, g) + \theta h(g)) h(g) \phi(g \mid m, \gamma) dg \quad (7)$$

This representation highlights that the T^2 operator can be interpreted as a robustness correction for uncertainty about the distribution of $g(t)$. Hansen and Sargent (2007) show that the worst case distribution for $h(g)\phi(g | m, \gamma)$ can be represented as a normal with variance γ and mean

$$\tilde{m} = m - \frac{1}{\theta} \frac{1}{k + \rho} \frac{\gamma}{2}.$$

The conditional worst case distribution for the growth state in future periods follows directly from the distribution assumptions. It is normal with mean

$$\tilde{E}[g(t + \Delta) | \mathcal{F}_t] = E[g(t + \Delta) | \mathcal{F}_t] - e^{-k\Delta} \frac{1}{\theta} \frac{1}{k + \rho} \frac{\gamma}{2} \quad (8)$$

and the same variance as the original conditional distribution. Here and in what follows, a tilde signifies that the moments are those for the worst case distribution and not the standard conditional distribution. The last term in equation (8) is strictly negative but decays at a geometric rate with Δ . The worst case distribution for $g(t + \Delta)$ is thus shifted downwards from the conditional one, but the downward shift goes to zero at the rate k as Δ increases.

The shift of the worst case density for $g(t)$ relative to the true one is proportional to the inverse of the robustness parameter θ . As $\theta \rightarrow \infty$, the aversion to state uncertainty dissipates and we are back to the standard expected utility case.

The effect of the discount rate ρ and the persistence parameter k are quite intuitive. Discounting blunts some of the effect of state uncertainty: The higher the discount rate, the less weight is given to future consumption relative to present, so utility will be less affected by changes to the growth rate.

A small k generates very persistent growth rates, so even small changes to the current growth rate compounds to large changes in future consumption. This increases the disutility of any uncertainty about the current state of the economy. In addition to this direct effect, k also enters in the differential equation governing γ . As $k \rightarrow 0$, the mean reversion in the consumption growth rate is shut down. This eliminates one of the channels through which state uncertainty dissipates over time.

3 The term structure of interest rates

Proposition 1 (Expectations). *Claims to discount bonds are priced using the conditional worst case probability distribution for the growth rate.*

Proof. Let $\tilde{V}(c, m, \gamma, \epsilon)$ be the value function of the representative investor who has invested a small amount ϵ in a risk free bond. Because bonds are in zero net supply, the representative investor will value bonds according to the limiting worst case distribution of his consumption stream as ϵ goes to zero. But as ϵ goes to zero, the consumption stream when holding ϵ of the bond approaches that of the endowment good. \square

The time t interest rate for a real discount bond that matures at time $t + \Delta$, denoted by $r(t, \Delta)$, is implicitly given by the Euler equation

$$e^{-r(t, \Delta)\Delta} = \tilde{E}_t \left[e^{-\rho\Delta} \frac{C(t)}{C(t + \Delta)} \right] \quad (9)$$

By the process assumptions, we know that

$$C(t + \Delta) = C(t) e^{\int_0^\Delta (\mu + g(t+s) - \frac{1}{2}v(s))ds + \int_0^\Delta \sqrt{v(s)}dB_c(s)}.$$

Using this equation to substitute for $C(t)/C(t + \Delta)$ in equation (9) and manipulating the expressions leads to the following result for the equilibrium rate on discount bonds:

Proposition 2 (Interest rates). *Under the process and utility specifications above, the time t interest rate on a discount bond maturing at time $t + \Delta$ is given by*

$$r(t, \Delta) = \rho + \mu + \frac{1}{k\Delta}(1 - e^{-k\Delta})\tilde{m}(t) - F(\Delta, \gamma, w, v, \lambda) \quad (10)$$

where

$$\begin{aligned} F(\Delta, \gamma, w, v, \lambda) = & \frac{1}{\Delta} \left(\frac{1}{k}(1 - e^{-k\Delta}) \right)^2 \frac{\gamma(t)}{2} \\ & + \frac{1}{\Delta} \log E_t \left[e^{\int_t^{t+\Delta} \int_s^\Delta e^{-ku} du} \sqrt{w(s)} dB_g(s) \right] \\ & + \frac{1}{\Delta} \log E_t \left[e^{\int_t^{t+\Delta} \frac{1}{2}v(s)ds + \int_t^{t+\Delta} \sqrt{v(s)} dB_c(s)} \right] \end{aligned}$$

Proof. See appendix \square

Due to the underlying log-utility function, the interest rate for maturity Δ is mainly determined by the sum of the time discount rate ρ and the expected average consumption growth rate between t and $t + \Delta$ under the worst case probability distribution, which is given by $\mu + \frac{1}{k\Delta}(1 - e^{-k\Delta})\tilde{m}(t)$. At very short horizons, the expected average growth rate will be close to $\mu + \tilde{m}(t)$ and it converges to μ as we increase the maturity Δ . The terms in $F(\Delta, \gamma, w, v, \lambda)$ are second-order variance adjustments.

3.1 Closed form solutions when volatility is constant

The case when the variance of the innovation terms are constant over time is convenient because it furnishes closed form expressions for $F(\Delta, \gamma, w, v, \lambda)$.

Proposition 3. *When w and v are constant over time, $F(\Delta, \gamma, w, v, \lambda)$ is given by*

$$F(\Delta, \gamma, w, v, \lambda) = -b_1(\Delta) \frac{\gamma(t)}{2} - b_2(\Delta) \frac{w}{2} - v \quad (11)$$

where

$$b_1(\Delta) \equiv \frac{1}{\Delta} \left(\frac{1}{k} (1 - e^{-k\Delta}) \right)^2$$

$$b_2(\Delta) \equiv \frac{1}{k^2} \left[1 - \frac{2}{k\Delta} (1 - e^{-k\Delta}) + \frac{1}{2k\Delta} (1 - e^{-2k\Delta}) \right].$$

Proof. Without stochastic volatility, $\int_t^{t+\Delta} \int_s^\Delta e^{-ku} du \sqrt{w} dB_g(s)$ is the time integral of a mean zero Ornstein-Uhlenbeck process starting at zero. This integral is a normal with known moments:²

$$\int_t^{t+\Delta} \int_s^\Delta e^{-ku} du \sqrt{w} dB_g(s) \sim \mathcal{N}(0, \Delta \cdot b_2(\Delta) \cdot w)$$

The integral involving the innovations to B_c is also normally distributed:

$$\int_t^{t+\Delta} \frac{1}{2} v ds - \int_t^{t+\Delta} \sqrt{v} dB_c(s) \sim \mathcal{N}\left(\frac{\Delta v}{2}, \Delta v\right)$$

Proposition 3 follows directly from applying the formula for means of log-normal variables to the expectations in $F(\Delta, \gamma, w, v, \lambda)$. \square

The average term premium that the model generates is of particular interest. Without volatility regimes, the unconditional (steady state) expectation of equation (10) is

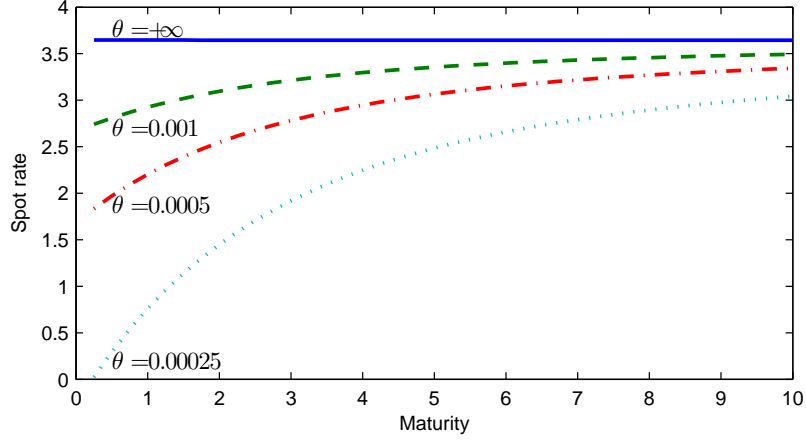
$$E[r(t, \Delta)] = \rho + \mu - \frac{1}{\Delta k} (1 - e^{-k\Delta}) \frac{1}{2\theta(k + \rho)} \gamma_{ss} - F(\Delta, \gamma_{ss}, w, v, \lambda) \quad (12)$$

For parameters of the magnitude estimate in the empirical part of this paper, $F(\Delta, \gamma, w, v, \lambda)$ is going to be in the order of a few basis points. This means that ability of the model to produce a quantitative significant term premium relies on the impact on the third term in the equation above. It's magnitude

²See e.g. Gillespie (1996), equation (2.8)

Figure 2: State uncertainty aversion and average term structure

This figure plots the unconditional term structure for varying levels of state uncertainty aversion without stochastic volatility for the following parameter values $\mu = 0.02$, $k = 0.65$, $\rho = 0.016$, $\sqrt{v} = 0.013$, and $\sqrt{w} = 0.0034$.



is proportional to (1) the amount of uncertainty about the growth state that is left after filtering γ , (2) on the impact of the current growth state on utility $(\rho+k)$, and (3) on how tolerant investors are to such uncertainty through the parameter θ . Halving θ doubles the adjustment of the term structure that comes from the uncertainty aversion at all maturities. This is illustrated in figure 2.

3.2 Error bounds when volatility is stochastic

The solution with constant volatility is instructive but would not be very useful if it varied very much from the solution in the case with stochastic volatility that one ultimately wants to tackle. In fact, for a given γ and \tilde{m}_t , the two interest rates are very close, as demonstrated by the following proposition

Proposition 4 (Error bounds). *Let $\underline{r}(t, \Delta)$ be the time t interest rate that would prevail if investors knew that for $s \in [0, \Delta]$, $v(t+s) = v_1$ and $w(t+s) = \max(w_1, w_2)$. Conversely, let $\bar{r}(t, \Delta)$ be the time t interest rate that would prevail if investors knew that for $s \in [0, \Delta]$, $v(t+s) = v_2$ and $w(t) = \min(w_1, w_2)$. Everything else equal then, $r(t, \Delta)$, the interest rate*

with stochastic volatility obeys

$$\underline{r}(t, \Delta) \leq r(t, \Delta) \leq \bar{r}(t, \Delta)$$

Proof. See appendix □

The coefficients on v in the constant volatility case is 1 while the coefficient on w is bounded above by $1/k^2$. This means that assuming that there will be no regime switches gives a good approximation of the real interest rate implied by the model. In particular,

$$r(t, \Delta) - \underline{r}(t, \Delta) \leq (1 + 1/k^2) \max(v, w) \quad (13)$$

For the calibration used later, the right hand side of (13) will be less than a few basis points. Because regime switches are such rare events, the actual approximation error will be less still.

Instead of relying on this approximation, one can approximate the the expectations arbitrarily well numerically. One possible method based on Dupuis and Kushner (2001) is described in appendix B. The results reported in the next sections rely on this numerical approximation, but the results would have been virtually identical using the closed form approximation and computing interest rates assuming no regime switching.

4 Estimation

4.1 Data

I use the following data to apply the model to the US economy:

- Consumption: Real per capita consumption of non-durables and services (NDS). The series was constructed from the available nominal and real chain weighted series ‘personal consumption expenditures’ and ‘durables consumption’ using the Divisia index approximation to the Fisher formula (see Whelan, 2002.) Source: NIPA, Table 7.1, line 12. Sample period: Q4:1953-Q4:2006.
- Nominal interest rates: End of the month three month risk free rate (Fama risk-free rate); as well as synthetic yields on 1, 2, 3, 4, and 5 year discount bonds from the Fama-Bliss dataset. All rates are from the CRSP database. Sample period: January 1964 - December 2006.
- Inflation: Change in the quarterly level of the NDS deflator. Sample period: Jan 1959-Jan 2006.

Quarterly NIPA data are available from 1947, but I restrict the sample to start from Q1:1953 to avoid issues related to the postwar reconstruction and the change in monetary policy after the 1951 Treasury-Federal Reserve Accord. The synthetic interest rate series are available from CRSP from June 1953, but their quality is poor in the early part of the sample because there are not enough non-callable, fully taxable bonds of different maturities outstanding. I follow Fama and Bliss (1987) who introduced the series and use January 1964 as the first period of the sample.

4.2 Discrete-time analog

I use Gibbs sampling to estimate the following discrete-time version of the model on the quarterly US data:

$$x_{t+1} = \kappa x_t + \sigma_x(S_{t+1})\epsilon_{t+1} \quad (14a)$$

$$c_{t+1} = c_t + \mu_c + x_t + \sigma_c(S_{t+1})\epsilon_{t+1} \quad (14b)$$

$$\Pr(S_{t+1} = i \mid S_t = j) = \begin{cases} \lambda & , \text{if } i \neq j \\ (1 - \lambda) & , \text{if } i = j \end{cases} \quad (14c)$$

$$\sigma_c(1) > \sigma_c(2) \quad (14d)$$

x_t is the underlying state of the consumption growth rates and c_t is real per capita log consumption. The estimation is a slightly complicated by the fact that the volatility state is hidden, but this can be dealt with by using lightly informative priors for the variance parameters.³ For μ and κ , I do not use prior information. For σ_c , the prior in each state is an Inverse Gamma distribution with a scale parameter of 5 and shape parameters of 5×0.005^2 and 5×0.0025^2 in states 1 and 2, respectively. For σ_x , the prior is also an Inverse Gamma with a scale parameter of 5, but I use the same shape parameter (5×0.0005^2) in both states. Finally, the prior for $\tilde{\lambda}$ is beta(1, 50). (Implying a prior mean of 0.0196 and a standard deviation 0.0192.) The individual steps of the Gibbs sampler are described in detail in appendix C.

As emphasized by Hansen (2007), the consumption data do not contain enough information to let us pin down the process parameters very well. This ambiguity leads to disperse distributions for the posterior distributions shown with the blue bars in figure 3. The most crucial model parameter is

³By construction, there will be draws for the time series of the volatility state where only one of the states appear. The priors prevent the algorithm from collapsing in these cases.

the persistence parameter κ . In a second stage, I fix κ to 0.885, the average of the estimated median and mode of the marginal density. I then re-estimate the model with κ fixed to this value and use the mean value as the estimate for each of the other parameters. Since 0.885 is higher than the mean value for κ , the marginal distributions conditional on $\kappa = 0.885$ (the transparent bars in the plot) are slightly shifted relative to the unconditional prior distributions. In particular, the posterior distributions for σ_c are shifted slightly to the left, reflecting that more consumption growth variation is accounted for by the growth rate process. The posterior distributions for σ_x are shifted to the right.

The parameter estimates from the Gibbs sampler are given in table 1. Except for the persistence parameter κ , the most important parameters are the variance terms because they determine the level of the steady state uncertainty in the two volatility regimes. The estimates point out that the main difference between the two regimes is the direct volatility of consumption. The standard deviation σ_c is more than twice as high in the high volatility state as in the low volatility state. The estimation indicates that also the volatility of the trend growth rate is lower in the low volatility state. The estimated transition probability of around 1.75 percent per quarter reflects that the Gibbs sampler on average picks up three regime switches (see below). Evidence from other macroeconomic series indicate that there was only one regime switch, so this number might be too high. As argued above, the influence of $\tilde{\lambda}$ on the yield curve is minimal, so this is not a big cause for concern.

Figure 4 shows the smoothed estimate of the hidden growth state plus the long term trend together with the raw consumption growth rates.

Subtracting the trend in figure 4 from the raw consumption growth rate data gives the detrended consumption innovations in the upper panel of figure 5. The plot attests to the well documented moderation in the US macroeconomic time series. (See e.g. Kim and Nelson, 1999; Blanchard and Simon, 2001; Stock and Watson, 2003.) In the latter part of the sample, the growth rates fluctuate significantly less than in the early part of the sample. The dashed lines in the plots mark the two standard deviation bands for the two volatility regimes that the estimator picks up. As one can see from the lower panel of the plot, the time series for this consumption aggregate indicates that the endowment process entered a low volatility regime sometime around the mid eighties, then briefly lapsed back to the high volatility regime around the short lived recession of 1990-1991, before settling in the low volatility regime for the rest of the sample.

The revealed chronology should be seen in the light of the evidence from

Figure 3: Posterior Distributions

This figure shows posterior distributions of parameter draws from the Gibbs sampler. The colored bars give the distribution of draws from the unconstrained estimation; The transparent bars with black borders given the distribution when κ is fixed to 0.885.

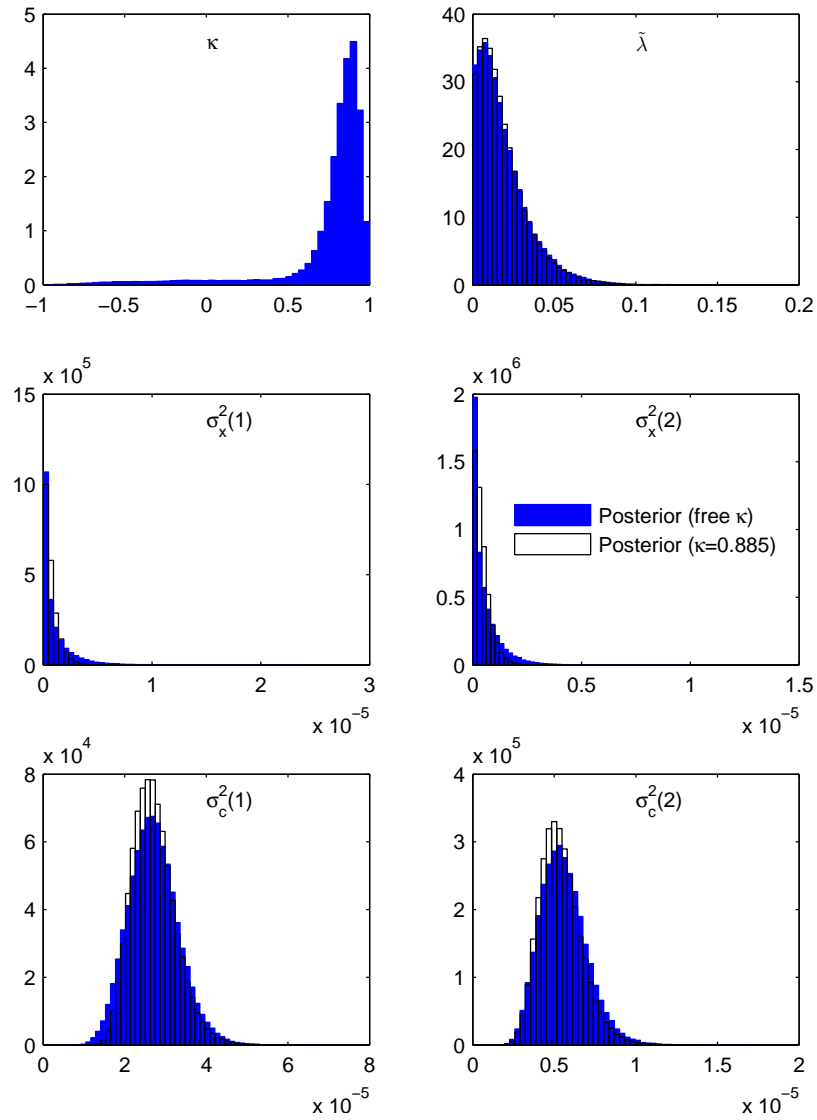


Table 1: Estimation results from the Gibbs sampler

This table gives the first stage estimation results for the growth rate process for personal consumption expenditures. For this stage of the estimation, the variance of the noise term in the observation equation is assumed to be constant. The standard deviations reported in parenthesis are those of the draws from the posterior distribution.

<i>Posterior Distributions</i>						
	Unconditional			Conditional on $\kappa = 0.885$		
	Mean	SD	MD	Mean	SD	MD
κ	0.7540	0.3034	0.8459			
μ_c	0.0051	0.0006	0.0051	0.0051	0.0005	0.0051
$\tilde{\lambda}$	0.0189	0.0158	0.0149	0.0182	0.0153	0.0144
$\sigma_x^2(1)$	0.0010 ²	0.0012 ²	0.0007 ²	0.0010 ²	0.0010 ²	0.0009 ²
$\sigma_x^2(2)$	0.0008 ²	0.0008 ²	0.0006 ²	0.0008 ²	0.0007 ²	0.0007 ²
$\sigma_c^2(1)$	0.0052 ²	0.0025 ²	0.0052 ²	0.0051 ²	0.0023 ²	0.0051 ²
$\sigma_c^2(2)$	0.0024 ²	0.0012 ²	0.0023 ²	0.0023 ²	0.0012 ²	0.0023 ²

Figure 4: Estimated growth state

This figure plots the time-series of consumption growth rates together with the smoothed estimate of the growth state ($x_t + \mu_c$). Values are in annualized percents.

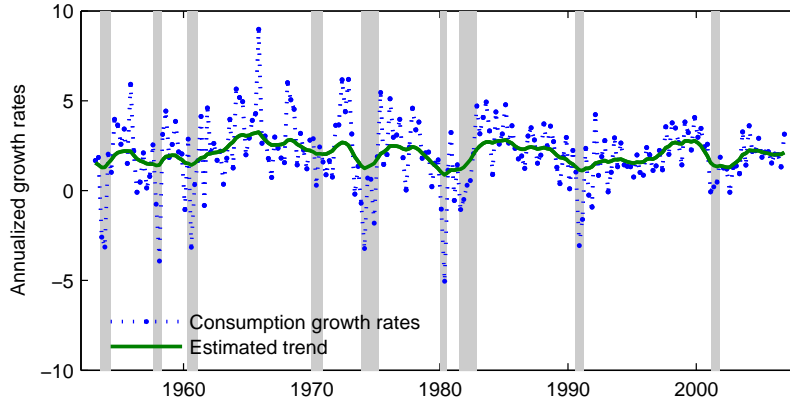
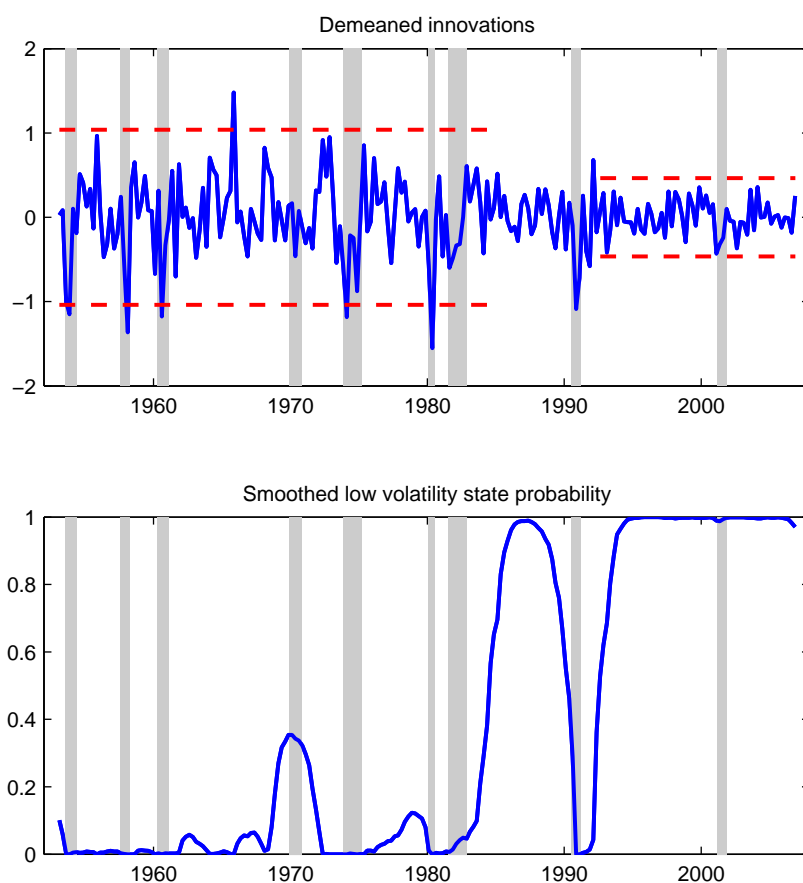


Figure 5: Demeaned consumption innovations and estimated volatility states

The top panel shows the quarterly growth rates of the consumption aggregate demeaned by the smoothed estimate of the trend from figure 4. The lower panel shows the smoothed probabilities of being in the low volatility state.

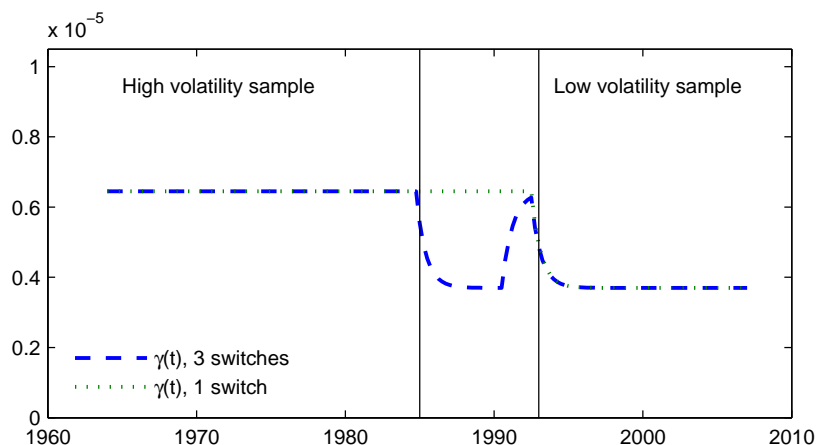


other time series. Unreported results using the same estimation procedure, but with *Personal Consumption Expenditures* used as a consumption aggregate, produces only one regime switch: The endowment process remains in the high volatility regime until the end of 1992, then switches to the low volatility state. (Lettau et al., 2006 observes the same pattern with slightly different process assumptions.)

A coherent series for $\gamma(t)$ is ideal for testing the model. As it turns out,

Figure 6: Volatility states and state uncertainty.

This figure shows the time series of state uncertainty ($\gamma(t)$) under two scenarios for the succession of volatility states. The dotted line shows $\gamma(t)$ in the case of a single regime switch taking place at the start of 1993. The dashed line shows $\gamma(t)$ in the case where there is an initial switch to the low volatility state in 1985, but then a short lived laps back to the high volatility state from 1990-1992. For the two sample periods we will use, the dashed and dotted lines practically overlap.



the series for $\gamma(t)$ predicted by the model up to 1984 and after 1993 is almost identical regardless of whether the only regime switch is at the end of 1992, or the switch to the low volatility state in the mid eighties is also bona fide. This is shown in figure 6. The dotted line ($\gamma(t)$ with three regime switches) overlap the dashed line ($\gamma(t)$ with just a single switch) in both the early and the late sample.

4.3 Time aggregation

Translating the continuous time estimate of the model into discrete time raises time aggregation issues. For the transition density, I multiply the estimated quarterly probability by 4 to arrive at the continuous time transition intensity λ . Likewise, I take the annualized trend consumption growth rate to be $\mu = 4 \times \mu_c$. The main difficulty is that quarterly measured consumption maps into average consumption over the whole quarter. This means that x_t represent the average of the integrated growth state between times

in the interval $t - 1$ to t and one quarter before. Likewise, the detrended consumption growth rate is taken to be the average compounded effect of the innovation terms in the consumption process over the same interval, i.e.

$$x_t = \int_{t-\Delta}^t \int_{s-\Delta}^s g(u) du ds$$

$$\hat{c}_t = \int_{t-\Delta}^t \int_{s-\Delta}^s \sqrt{v(u)} dB_c(u) ds$$

where the time interval Δ is quarter or 0.25 years. If v is constant over the whole interval $[t - 2\Delta, t]$, the last integral is known to have variance $2/3 \times \Delta \times v$ (see Working, 1960; Breeden et al., 1989). Consequently, I take the variance in each state to be $3/(2\Delta)$ times that estimated for the discrete model.

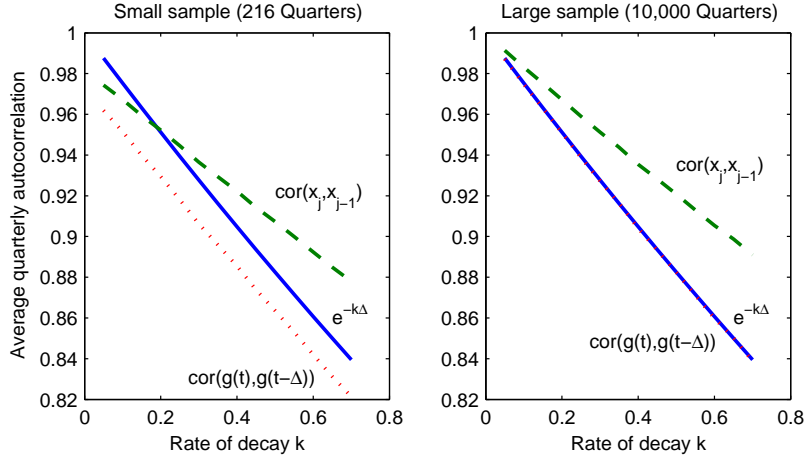
While $g(t)$ is mean zero, normally distributed, and so is its time integral, I am not aware of any closed form expression for the correlation of the integral between two successive periods. Without regime switches the unconditional variance of $g(t)$ is proportional to w . Thus we can factor out w from the integral, and x_t will be proportional to w , too. Now, to link κ to k , I generate 10,000 paths for the process x_t with length 54 years for various parameters k . Each path, is integrated over the 216 quarters it contains and sample correlation coefficient are computed for successive quarters. The average correlation coefficient over the 10,000 paths links the estimate of κ to the underlying k . Thus collected pairs (k, κ) are shown as the dashed line in the left hand side panel of figure 7. For comparison $e^{-0.25k}$, the true quarterly correlation coefficient of $g(t)$ is given by the solid line and the estimated correlation coefficient of the quarterly sampled $g(t)$ is given by the dotted line. When the decay rate is very low, the estimated correlation coefficient is less than $e^{-0.25k}$,⁴ but for higher decay rates, the relation between the estimate for the time aggregated data and the underlying model changes sign. On average, the estimated correlation coefficient of $\kappa = 0.885$ maps into a decay parameter k of 0.65. This is the value used to calibrate the model.

The rest of the figure attempts to disentangle this finding. It is well known that the normal estimate for the AR(1) coefficient is downward biased in small samples. In the left hand panel, this generates the shift of the dotted line with respect to the solid one. The distance between the dashed and the dotted line then gives the effect of time aggregation. The right hand

⁴This is what Bansal et al. (2006) find when aggregating a monthly model with a high persistence parameter to a yearly frequency.

Figure 7: Time aggregation of $g(t)$

This figures shows the effect of time aggregation of $g(t)$ on the measured correlation coefficient.



panel shows simulation results for a large sample with 10,000. Comparing the two panels, the small sample bias appears to be independent of the time-aggregation effect: In the large sample, the estimated correlation coefficient for $g(t)$ is close to the true one, while the distance between the average estimated correlation coefficients of $g(t)$ and $x(t)$ is roughly the same as in the short samples.

For a given k , the estimated σ_x^2 should be proportional to the true underlying w from the continuous time counterpart. To find the proportionality factor, I simulated 100,000 sample paths with $k = 0.65$ and $w = 1$. The average estimated σ_x^2 was 0.35. This proportionality factor was confirmed for other variances.

4.4 Inflation process

Lacking a long historical series of the real term structure of interest rates, it is necessary to impute it from the nominal bond prices. For parsimony, I assume that inflation is independent of any of the real state variables in the economy. Under this assumption, inflation neither Granger-cause consumption nor is Granger caused by it. This is sufficient for the Fisher equation to hold, so that log nominal interest rates at any horizon are simply the sum of the real interest rate for that horizon and the expected inflation

Table 2: Estimated ARMA(1,1) process for inflation

This table contains the maximum likelihood estimates of the ARMA(1,1) process for annualized inflation (π_t):

$$(\pi_t - \mu_\pi) = \phi_\pi(\pi_{t-1} - \mu_\pi) + w_t + \theta_\pi w_{t-1} \quad w_t \sim \mathcal{N}(0, \sigma_\pi^2)$$

The maximization is on the exact likelihood function. The standard errors reported in parenthesis are the estimated asymptotic ones from the numerical derivatives of the Hessian.

μ_π	ϕ_π	θ_π	σ_π
3.73191 (1.96366)	0.94414 (0.02548)	-0.39095 (0.06908)	1.83281 (0.08859)

rate.

Amongst the class of ARMA(p,q) processes, both the Akaike and the Schwartz' information criterion pick the ARMA(1,1) as the representation that fits the historical inflation series best.⁵ Maximizing the exact likelihood function from the Kalman filter yields the parameter estimates reported in table 2. Figure 8 shows the historical annualized changes to the NDS deflator along that forecasted one quarter before based on the estimated inflation process.

4.5 The historical US term structure

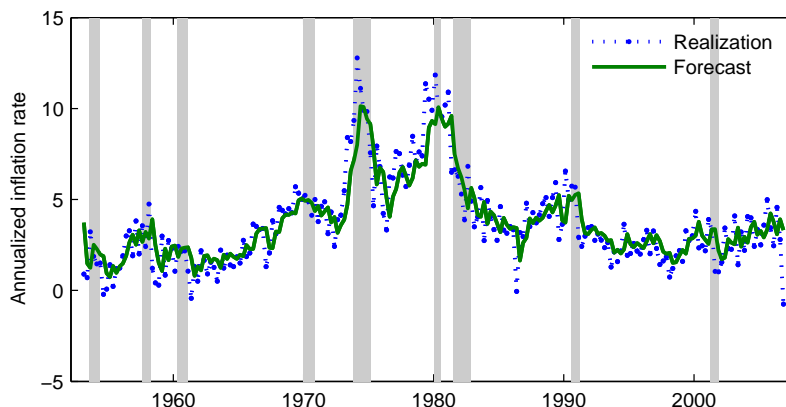
Assuming that the inflation process estimated above is a good description of investors' subjective beliefs about the data generating process, it is possible to generate time-series of real discount rates at various maturities. This assumption is obviously a stark simplification, especially since the time period analyzed potentially covers multiple monetary regimes and time spans where long term inflation expectations were not well anchored (e.g. Clarida et al. (2000) Cogley and Sargent (2005) and Sims and Zha (2006)). However, the inflation rates are so persistent in the data that, as long as we impose some stationarity, the true expectations might be reasonably well described by the estimated ARMA(1,1) process.

Table 3 gives key statistics for nominal rates, real rates, and holding returns for three intervals. The top part of the table gives averages for

⁵Both criteria reach minimum values on the grid $(p, q) \in \{0, 1, \dots, 10\}^2$ at $(p, q) = (1, 1)$.

Figure 8: Historical and predicted inflation rates

This figure shows the realized and forecasted inflation rates for the NDS deflator. The predicted inflation rate is the one quarter ahead forecast of the ARMA(1,1) process reported in table 2. The innovation terms to the process are estimated with the Kalman filter.



the whole sample period considered. Both the nominal and the real rates exhibit a clear term-structure profile: Interest rates grow with maturity, but at a decaying rate. Average 1 year yields are about 43 basis points higher than the average 3 month rate, while 5 year yields are on average only about 9 basis points higher than 4 year yields. Holding returns are generally computed as the realized return from holding the synthetic bond for 1 year. The exception is the 3 month bond, where it is computed as the one year realized return from rolling over 3 month treasury bills. The same profile can be discerned also in the holding returns, but, because returns are realizations and not expectations, it is more noisy.

The rest of the table looks at two subsamples with different volatilities. The first subsample, 1963-1985, covers a period with high volatility. Because of the uncertainty related to the classification of the volatility regime for 1985-1992, I start the low volatility sample in 1992.

When comparing the early sample with the full sample, two differences stand out. First, the nominal term structure is less upward sloping in the early sample. Within the model, this is due to higher average inflation in this period. Because inflation is mean reverting, inflation is expected to decline when current inflation is high. For a given real term structure, this shifts the long nominal rates down relative to short nominal rates. Second, the real

Table 3: Historical interest rates

This table reports means and standard deviations of discount rates, ex-ante expected real discount rates, and one year holding returns for different time periods for the synthetic bonds considered. The Newey-West standard errors with 4 lags are given in parenthesis. 3M rates correspond to the CRSP risk-free rate, while the other rates are yields on the synthetic discount bonds from the Fama-Bliss data set. (Also taken from CRSP.)

Maturity:	3M	1Y	2Y	3Y	4Y	5Y
<i>Sample: 1964-2006</i> (whole period)						
Nominal rates	5.88 (0.45)	6.31 (0.45)	6.52 (0.44)	6.69 (0.43)	6.82 (0.42)	6.90 (0.41)
Real rates	1.71 (0.32)	2.18 (0.33)	2.43 (0.34)	2.63 (0.34)	2.79 (0.34)	2.90 (0.34)
Holding returns (1Y)	5.85 (0.45)	6.27 (0.45)	6.67 (0.56)	6.93 (0.68)	7.11 (0.82)	7.09 (0.94)
<i>Sample: 1964-1984</i> (high volatility)						
Nominal rates	7.03 (0.67)	7.48 (0.65)	7.58 (0.66)	7.65 (0.65)	7.71 (0.65)	7.74 (0.65)
Real rates	1.76 (0.51)	2.34 (0.51)	2.58 (0.53)	2.78 (0.54)	2.94 (0.55)	3.07 (0.55)
Holding returns (1Y)	6.85 (0.67)	7.13 (0.66)	6.93 (0.83)	6.63 (0.96)	6.30 (1.14)	5.88 (1.29)
<i>Sample: 1993-2006</i> (low volatility)						
Nominal rates	3.85 (0.45)	4.23 (0.47)	4.50 (0.44)	4.73 (0.40)	4.91 (0.37)	5.02 (0.34)
Real rates	1.16 (0.50)	1.45 (0.51)	1.62 (0.48)	1.76 (0.44)	1.87 (0.41)	1.92 (0.38)
Holding returns (1Y)	3.78 (0.44)	4.15 (0.47)	4.75 (0.62)	5.30 (0.85)	5.76 (1.02)	6.00 (1.17)

term structure in this subsample is slightly more upward sloping than for the whole period. This is just as expected, because higher uncertainty drives down the short end of the term structure. On a lesser note, the holding returns for long run nominal bonds are higher in this subsample. This can be traced back to surprise inflation.

Based on the discussion above, the real term structure should flatten out somewhat when the economy moves to the low volatility regime. This is confirmed in the data. Also, the steepening of the nominal term structure compared to the real one is just the reversal of the effect in the early sample. The one feature which the model cannot account for is the apparent downward level shift of the whole curve.

5 Model Implications

5.1 Volatility states and term premium

As discussed above, the term structure model proposed in this paper has a strong testable prediction on how the slope of the yield curve should react to the Great Moderation in the US time-series. Following the switch to the low volatility regime, investors will gradually gain more accurate information on the state of the economy. As the uncertainty dissipates, so does the downward adjustment of the inferred current growth state too. This flattens the yield curve and reduces the term premium.

The model parameters θ and ρ are calibrated to minimize the mean squared distance between the average real discount curve in the high volatility part of the sample (1963-1984) and that predicted by the model. The minimizing parameter combination has $\rho = 0.016$ and $\theta = 0.00035$.

Table 4 compares the predicted yield curve with the empirical one for the high volatility sample. The predicted yield curve matches neatly not only the slope but also the curvature of the empirical one. This is only possible because the estimated persistence parameter ($k = 0.65$) makes the influence of uncertainty decay at the appropriate rate. A lower level of k (more persistence) would decrease the curvature, while a higher level of k (less persistence) would result in more curvature.

We are now ready to confront the model predictions for the entire sample with the data. Figure 9 collates predicted with empirical yield curves for both subsamples from table 3. The left-hand side panel merely reproduces the information in table 4. The fit is very good, which reflects both that that $k = 0.65$ is a good parameter for the the decay rate of the growth state and that the model is calibrated to match the average yield curve in

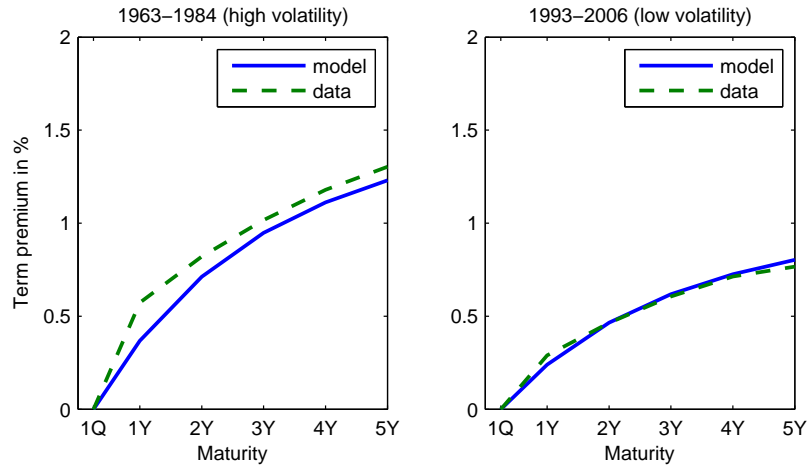
Table 4: Model calibration

Predicted versus empirical average real term structure for the period 1963-1984 for $k = 0.65$, $\mu = 0.02$, $\sqrt{w_1} = 0.0029$, $\sqrt{v_1} = 0.013$, $\rho = 0.016$, and $\theta = 0.00035$.

	3M	1Y	2Y	3Y	4Y	5Y
Data	1.76	2.34	2.58	2.78	2.94	3.07
Model	1.83	2.20	2.55	2.78	2.95	3.07
Difference	0.07	-0.13	-0.04	0.00	0.00	-0.00

Figure 9: Subsample real term premia

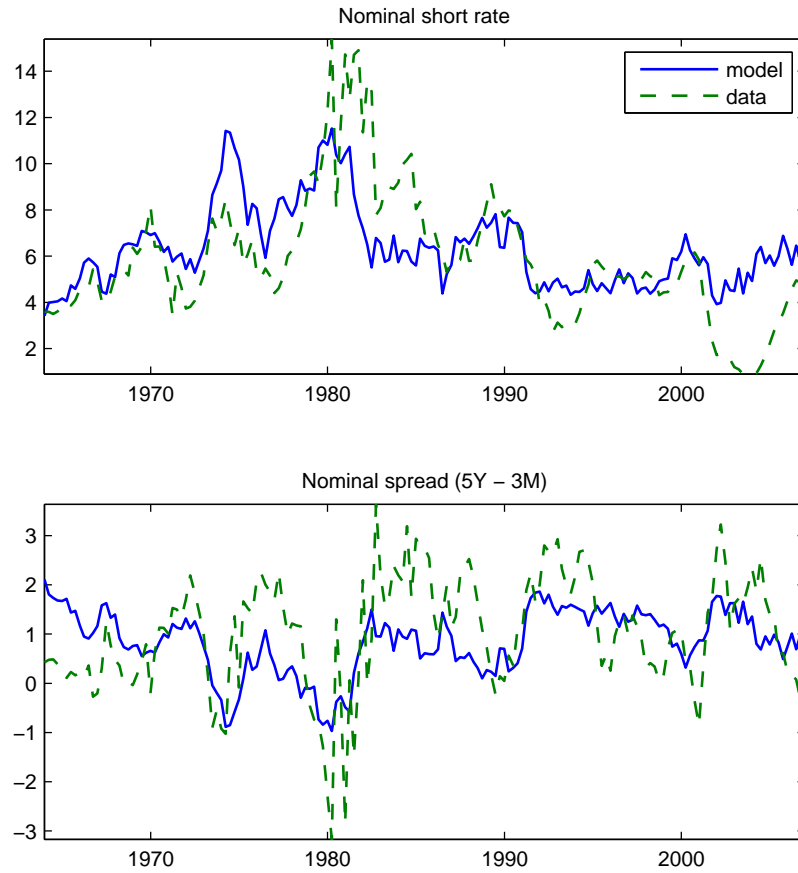
This figure shows average term premia predicted by the model for each of the subsamples together with the empirical averages.



this sample. The right-hand side panel, which shows the fit of the model in the low volatility sample, is more interesting. The average yield curve in the post 1992 sample is much flatter, but with roughly the same curvature. Since lower volatility translates in lower state uncertainty, we expect the model to predict a similar effect. Still, it is striking that the magnitude of the effect is so close to empirical one. The only aspect of the average term structure in the low volatility sample that the model does not reproduce is the level of the yield curve. In the data, the average real rates are about

Figure 10: Predicted vs empirical nominal rates

The lower panel plots the nominal short rate (3M) predicted by the model along with the empirical one, while the lower panel plots the predicted and empirical nominal spread (5Y rate minus 3M rate).



1 percent lower than what the model predicts. As shown below, this shortcoming stems from the much discussed slump in global interest rates following the deflation of equity prices at the turn of the century.

The two panels of figure 10 plot the predicted time series for the nominal short rate and the spread between the short rate and the yield on 5 year discount bonds against their empirical counterparts. For the intermediate period 1985-1992, it is assumed, as the Wonham filter indicates, that the

economy entered the low volatility state at the end of 1984, lapsed back into the high volatility state in 1990 before escaping it again in the end of 1992. The model matches the empirical short rate quite well. The correlation coefficient between the predicted short rate and the empirical one is 0.64, which is higher than the 0.50 coefficient Wachter (2006) reports for her model. Piazzesi and Schneider (2006) provide the same plot, but do not report the correlation coefficient. The main shortcomings of the model is that it cannot reproduce quantitatively the sharp jumps in the nominal interest rate around the tightening of monetary policy under Federal Reserve chairman Volcker and that the predicted nominal rates in the seventies are too high relative to the empirical ones. This is probably because the model does not take into account changes in monetary policy. The model also cannot account for the slump in the short rate after 2001. This is however not unique to this model. In the words of John Taylor (2007) “During the period from 2003 to 2006 the federal funds rate was well below what experience during the previous two decades of good economic macroeconomic performance—the Great Moderation—would have predicted.” As figure 10 shows, the inability of the model to match the average level of the yield curve during the low volatility part of the sample is entirely due to its inability to match this slump.

The lower panel of the figure shows that the model also matches the empirical nominal spread well. The correlation coefficient between the predicted series and the empirical one is 0.43. This is lower than for the short rate, but because the spread is the difference between two rates, it is not clear that the numbers are directly comparable. Again the biggest disparity between the model’s predictions and the data are around the turbulent period for monetary policy in the late seventies and early eighties. Apart from this subperiod, the fit of the model is good.

6 Concluding remarks

This paper shows that an otherwise standard asset pricing model where investors dislike uncertainty about the current level of a mean reverting component of trend consumption growth rates generates a positive real term premium. If investors are sufficiently averse to state uncertainty, the model can match the empirical term premium in the US data.

The Great Moderation provides a natural laboratory for testing the model mechanism. When the variance of consumption growth rates drop, uncertainty about the trend growth rate is also lowered. This reduces the

shift of the yield curve and the term premium. When the estimated reduction in state uncertainty is fed into the model, it predicts accurately the flattening of the yield curve since 1993.

Some cautionary remarks are necessary. First, the model treats inflation expectations in a somewhat pedestrian way. The consensus view is that monetary policy has changed over the time period. An extended model would account for investors' learning about monetary policy and forming their inflation expectations accordingly. Especially around the time of the Great Inflation this could improve the fit of the model. Piazzesi and Schneider (2006) give one example of how this can be done. Second, the model endows investors with complete knowledge about the structural parameters of the model. Again, a more complete model would incorporate parameter uncertainty and robustness considerations. Third, θ , which is a deep preference parameter, has been calibrated to fit the model to the data. Some guidance on the plausibility of the calibrated value would be welcome.

A Proofs

Proof of proposition 2. Substituting for $C(t)/C(t+\Delta)$ in equation (9) yields:

$$\begin{aligned} e^{-r(t,\Delta)\Delta} &= \tilde{E}_t \left[e^{-(\rho+\mu)\Delta - \int_t^{t+\Delta} (g(s) - \frac{1}{2}v(s))ds - \int_t^{t+\Delta} \sqrt{v(s)}dB_c(s)} \right] \\ &= e^{-(\rho+\mu)\Delta} \tilde{E}_t \left[e^{-\int_t^{t+\Delta} g(t+s)ds} \right] E_t \left[e^{\frac{1}{2}\int_t^{t+\Delta} v(s)ds - \int_t^{t+\Delta} \sqrt{v(s)}dB_c(s)} \right] \end{aligned} \quad (15)$$

We can decompose the integral of future growth states in the following manner:

$$\begin{aligned} \int_t^{t+\Delta} g(s)ds &= \int_t^{t+\Delta} e^{-ks}g(t)ds + \int_t^{t+\Delta} \int_s^\Delta e^{-ku}du \sqrt{w(s)}dB_g(s) \\ &= \frac{1}{k}(1 - e^{-k\Delta})g(t) + \int_0^\Delta \frac{1}{\kappa} \left(1 - e^{-k(\Delta-s)}\right) \sqrt{w(s)}dB_g(t+s) \end{aligned} \quad (16)$$

With respect to the risk adjusted distribution, $g(t)$ is normally distributed with mean $\tilde{m}(t)$ and variance $\gamma(t)$; $g(t)$ is also independent of the increments to $dB_g(t+s)$ for $s > 0$. The stochastic integral on the right hand side is also normally distributed with mean zero.⁶ Conditional on \mathcal{F}_t we hence have

$$\begin{aligned} E_t[e^{-\int_0^\Delta g(t+s)ds}] &= \tilde{E}_t[e^{-\frac{1}{k}(1-e^{-k\Delta})g(t)}] \cdot E_t[e^{\int_t^{t+\Delta} \frac{1}{\kappa}(1-e^{-k(\Delta-s)}) \sqrt{w(s)}dB_g(s)}] \\ &= e^{-\frac{1}{k}(1-e^{-k\Delta})\tilde{m}(t) + \frac{1}{k^2}(1-e^{-k\Delta})^2 \frac{\gamma(t)}{2}} \cdot E_t[e^{\int_t^{t+\Delta} \frac{1}{\kappa}(1-e^{-k(\Delta-s)}) \sqrt{w(s)}dB_g(s)}] \end{aligned}$$

Substituting this expression in equation (15) gives equation (10). \square

Proof of proposition 4. Let $\underline{w} = \min(w_1, w_2)$ and $\bar{w} = \max(w_1, w_2)$ To prove the proposition, we need to show that

(i)

$$b_1(\Delta) \frac{\underline{w}}{2} \leq \frac{1}{\Delta} \log E_t \left[e^{\int_t^{t+\Delta} \int_t^s e^{-k(u-t)} \sqrt{w(u)}dB_g(u)ds} \right] \leq b_1(\Delta) \frac{\bar{w}}{2}$$

(ii)

$$v_2 \leq \frac{1}{\Delta} \log E_t \left[e^{\int_t^{t+\Delta} \frac{1}{2}v(s)ds + \int_t^{t+\Delta} \sqrt{v(s)}dB_c(s)} \right] \leq v_1$$

Notation: Let $\mathbf{t} = [t_1 \ t_2 \ \dots \ t_n]$ denote the times at which the volatility regime switches between t and $t + \Delta$, let $t_0 = t$ and $t + \Delta = t_{n+1}$,

⁶See e.g. Gillespie (1996)

and $\Delta = [\Delta_1 \ \Delta_2 \ \dots \ \Delta_{n+1}]$ the length of the time intervals $(t_1 - t_0)$, $(t_2 - t_1)$, \dots , $(t_{n+1} - t_n)$. Finally, let W_j and V_j be the variance of the innovation terms of the growth rate process and the innovation terms to the observation equation on the interval (t_{j-1}, t_j) , respectively. In addition, let $X(t_j) = g(t) - E_{t_{j-1}}[g(t_j)]$, $Y_j = \int_{t_{j-1}}^{t_j} X(t)dt$, and $N_j = \int_{t_{j-1}}^{t_j} \frac{1}{2}v(s)ds + \int_{t_j}^{t_j} \sqrt{v(s)} dB_c(s)$

Part (i): Since the volatility on every sub interval is constant by construction, we know from above that

$$Y_j \sim \mathcal{N}(0, \Delta \ b_2(\Delta)W_j)$$

Now, it is easy to verify that

$$\int_t^{t+\Delta} \int_t^s e^{-k(u-t)} \sqrt{w(u)} dB_g(u) ds = \sum_{j=1}^{n+1} \left(X(t_j)(1 - e^{-k(t+\delta-t_j)})/k + Y_j \right)$$

Now, let $M_j = (X(t_j)(1 - e^{-k(t+\delta-t_j)})/k + Y_j)$. M_j is the sum of two random normal variables and hence also a random normal variable. Moreover, M_i is independent of M_j for $i \neq j$. Using this, it follows that each of the elements in the summation are independent random normal variables.

$$E[e^{\int_t^{t+\Delta} \int_t^s e^{-k(u-t)} \sqrt{w(u)} dB_g(u) ds}] = \prod_{j=1}^{n+1} E[e^{M_j}]$$

The variance of M_j is increasing in W_j . Since W_j is either \bar{w} or \underline{w} , it follows that

$$E[e^{M_j} \mid W_j = \underline{w}] \leq E[e^{M_j}] \leq E[e^{M_j} \mid W_j = \bar{w}]$$

Hence

$$\prod_{j=1}^{n+1} E[e^{M_j} \mid W_j = \underline{w}] \leq \prod_{j=1}^{n+1} E[e^{M_j}] \leq \prod_{j=1}^{n+1} E[e^{M_j} \mid W_j = \bar{w}]$$

The left hand side of the inequality equals $E[e^{\int_t^{t+\Delta} \int_t^s e^{-k(u-t)} \sqrt{\underline{w}} dB_g(u) ds}] = e^{\Delta \cdot b_2(\Delta) \underline{w}}$. While the right hand side of the inequality is equal to $e^{\Delta \cdot b_2 \bar{w}}$. Since the inequality is true for any realization of the random vector \mathbf{t} , it is always true. Taking logs and dividing by $-1/\Delta$ yields part (i).

Part (ii): Again, by construction the variance in each partition of $(t, t + \Delta)$ is constant. The compounded innovations to observation equation over the subinterval (t_{j-1}, t_j) , N_j is a normal with

$$N_j \sim \mathcal{N}\left(\frac{V_j \Delta_j}{2}, V_j \Delta_j\right)$$

It follows that

$$e^{v_2 \Delta_j} \leq e^{V_j \Delta_j} = E[e^{N_j}] \leq e^{v_1 \Delta_j}$$

Now, since

$$\int_t^{t+\Delta} \frac{1}{2} v(s) ds + \int_t^{t+\Delta} \sqrt{v(s)} dB_c(s) = \sum_{j=1}^{n+1} N_j$$

and, for a given t , N_i and N_j are independent for $i \neq j$

$$e^{v_2 \Delta} \leq e^{\sum_j V_j \Delta_j} = E_t \left[e^{-\int_t^{t+\Delta} \frac{1}{2} v(s) ds + \int_t^{t+\Delta} \sqrt{v(s)} dB_c(s)} \right] \leq e^{v_1 \Delta} \quad (17)$$

As above, because the inequality is true for any realization of the random vector \mathbf{t} , it is always true. Taking logs and dividing by $-1/\Delta$ yields part (ii). \square

B Numerical approximation of variance terms

This appendix describes the discretization approach used to approximate the expectations of the variance terms from the continuous time model. We divide the relevant maturity Δ into n subintervals of equal length $\delta = \Delta/n$. Let t_i denote the starting point of the i th interval. The volatility state is only allowed to change at the end of each interval. The probability of a switch in the volatility state between two subintervals is given by λ/δ . $\hat{s}^{t+\Delta}$ gives one possible succession of volatility states over the discretized Markov chain. The probability of this path realizing is denoted by $\pi(\hat{s}^{t+\Delta})$. The first element of $\hat{s}_1^{t+\Delta}$ is given by the current volatility state.

It is straightforward to deal with the influence of stochastic volatility on the diffusion part, because

$$\log E \left[e^{\int_0^\delta \frac{1}{2} v(\hat{s}_i^{t+\Delta}) du + \int_0^\delta \sqrt{v(\hat{s}_i^{t+\Delta})} dB(u)} \right] = \delta v(\hat{s}_i^{t+\Delta}).$$

More care is needed with the innovations to the state equation, since the impact of early innovations on time $t + \Delta$ consumption is greater than late

innovations. The influence of innovations in the i th interval on consumption at $t + \Delta$ is given by

$$\int_{t_i}^{t_{i+1}} \left(\int_s^{t+\Delta} e^{-ku} du \right) \sqrt{w(\hat{s}_i^{t+\Delta})} dB_g(s) = \int_{t_i}^{t_{i+1}} \frac{1}{k} (1 - e^{-k(t+\Delta-t_i)}) \sqrt{w(\hat{s}_i^{t+\Delta})} dB_g(s)$$

This stochastic integral is mean zero normally distributed with variance

$$\alpha(\Delta, i) w(\hat{s}_i^{t+\Delta}) = \frac{1}{k^2} \left(\delta - \frac{2}{k} (e^{-k(t+\Delta-t_{i+1})} - e^{-k(t+\Delta-t_i)}) + \frac{1}{2k} (e^{-2k(t+\Delta-t_{i+1})} - e^{-2k(t+\Delta-t_i)}) \right) w(\hat{s}_i^{t+\Delta}).$$

Finally, we approximate the expectations of the stochastic integrals by

$$\begin{aligned} \log E_t \left[e^{\int_t^{t+\Delta} \int_s^\Delta e^{-ku} du \sqrt{w(s)} dB_g(s)} \right] &\approx \sum_{\hat{s}^{t+\Delta}} \pi(\hat{s}^{t+\Delta}) \left(\sum_{i=1}^n \alpha(\Delta, i) w(\hat{s}_i^{t+\Delta}) \right) \\ \log E_t \left[e^{\int_t^{t+\Delta} \frac{1}{2} v(s) ds + \int_t^{t+\Delta} \sqrt{v(s)} dB_c(s)} \right] &\approx \sum_{\hat{s}^{t+\Delta}} \pi(\hat{s}^{t+\Delta}) \left(\sum_{i=1}^n \delta v(\hat{s}_i^{t+\Delta}) \right) \end{aligned}$$

C Gibbs sampler

The Gibbs sampler takes a set of prior distributions as well as initial values for the parameters and cycles through draws of each of the unknown variables from their densities conditional on the draws for the other parameters. After each draw, the draw replaces the old value of that parameter. To wipe out the effect of the initial parameter values, I discard the 10,000 first of a total of 100,000 swipes of the sampler. Each swipe thereafter is stored. The individual steps of the sampler are schematically described below.

for $j = 1$ to 100,000 **do**

1) *Draw a series of growth states:*⁷ Conditional on the current draw of the other parameters and the volatility state, run the Kalman filter to obtain a series of conditional distributions for x_t , $f(x_t | c^t)$. Make a draw for \hat{x}_T . Then, moving backward.

for $t = T - 1$ to 1 **do**

Use the last draw to compute the conditional distribution $f(x_t | c^t, \hat{x}_{t+1})$ and make a draw for \hat{x}_t .

⁷To save space, I do not include formulas for sampling the hidden states here. These can be found in Kim and Nelson (1999), chapters 8-10.

end for

2) *Draw a series of volatility states:* Conditional on the current draw of the other parameters and the growth state, run the Wonham filter (see e.g. Hamilton, 1994, chapter 22) to obtain a series of conditional distributions for S_t , $f(S_t | c^t)$. Make a draw for \hat{S}_T . Then, moving backward

for $t = T - 1$ to 1 **do**

 Compute the conditional distribution $f(S_t | c^t, \hat{S}_{t+1})$ and make a draw for \hat{S}_t .

end for

3) *Draw a new mean μ_c :* Compute the terms $\hat{y}_t = c_{t+1} - c_t - \hat{x}_t$. Conditional on c^t, \hat{x}^t , and σ_c .

$$\mu_c | \hat{\sigma}_c^2, c^t, \hat{x}^t \sim \mathcal{N} \left(\frac{1}{T} \sum \hat{y}_t, \frac{1}{T} \sum \hat{\sigma}_c^2(\hat{S}_t) \right).$$

4) *Draw a new persistence parameter κ :* κ is conditionally normal with mean and variance given by

$$\kappa | \hat{\sigma}_c^2, c^t, \hat{x}^t \sim \mathcal{N} \left(\frac{1}{T-1} \sum \hat{x}_t \hat{x}_{t-1}, \sum \frac{\hat{\sigma}_x^2(\hat{S}_t)}{\hat{x}_t^2} \right).$$

5) *Draw $\sigma_c^2(1)$:* Since the prior distribution for $\sigma_c(1)$ is an Inverse Gamma, the posterior distribution will also be an Inverse Gamma. The shape parameter of the distribution is the sum of the shape parameter of the prior and the number of observations from state 1. The scale parameter is the sum of the scale parameter of the prior and the sum of squared detrended consumption innovations from state 1.

repeat

 6) *Draw $\sigma_c^2(2)$:* as in step 5.⁸

until $\hat{\sigma}_c^2(2) < \hat{\sigma}_c^2(1)$

 7) *Draw $\sigma_x^2(1)$:* as in step 5.

 8) *Draw $\sigma_x^2(2)$:* as in step 5.

if $j > 10,000$ **then**

 Save current draw for parameters and states.

end if

end for

The estimated parameter values correspond to the mean of the draws of each variable.

⁸This loops implements rejection sampling to ensure that stat 1 is identified as the high volatility state.

References

- Backus, David K., Allan W. Gregory, and Stanley E. Zin**, “Risk premiums in the term structure : Evidence from artificial economies,” *Journal of Monetary Economics*, November 1989, *24* (3), 371–399.
- Bansal, Ravi, Dana Kiku, and Amir Yaron**, “Risks for the long run: Estimation and Inference,” Working paper, Duke University October 2006.
- Blanchard, Olivier and John Simon**, “The Long and Large Decline in U.S. Output Volatility,” *Brookings Papers on Economic Activity*, 2001, *2001* (1), 135–164.
- Breeden, Douglas T, Michael R Gibbons, and Robert H Litzenberger**, “Empirical Tests of the Consumption-Oriented CAPM,” *Journal of Finance*, June 1989, *44* (2), 231–62.
- Brennan, Michael J. and Yihong Xia**, “Stock price volatility and equity premium,” *Journal of Monetary Economics*, April 2001, *47* (2), 249–283.
- Campbell, John Y. and John H. Cochrane**, “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior,” *Journal of Political Economy*, April 1999, *107* (2), 205–251.
- Clarida, Richard, Jordi Gal, and Mark Gertler**, “Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory,” *The Quarterly Journal of Economics*, February 2000, *115* (1), 147–180.
- Cogley, Timothy and Thomas J. Sargent**, “The conquest of US inflation: Learning and robustness to model uncertainty,” *Review of Economic Dynamics*, April 2005, *8* (2), 528–563.
- David, Alexander**, “Fluctuating Confidence in Stock Markets: Implications for Returns and Volatility,” *Journal of Financial and Quantitative Analysis*, December 1997, *32* (4), 427–462.
- Detemple, Jerome B**, “Asset Pricing in a Production Economy with Incomplete Information,” *Journal of Finance*, June 1986, *41* (2), 383–91.
- Dothan, Michael U and David Feldman**, “Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy,” *Journal of Finance*, June 1986, *41* (2), 369–82.

- Dupuis, Paul G. and Harold J Kushner**, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer, 2001.
- Fama, Eugene F. and Robert R. Bliss**, “The Information in Long-Maturity Forward Rates,” *American Economic Review*, September 1987, 77 (4), 680–692.
- Gillespie, Daniel T**, “Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral,” *Physical Review E*, August 1996, 73 (2), 2084–2091.
- Hamilton, James D.**, *Time Series Analysis*, Princeton, New Jersey: Princeton University Press, 1994.
- Hansen, Lars Peter**, “Beliefs, Doubts and Learning: Valuing Macroeconomic Risk,” *American Economic Review*, May 2007, 97, 1–30.
- **and Thomas J. Sargent**, “Fragile beliefs and the price of model uncertainty,” Working Paper, New York University and University of Chicago March 2007.
- Kim, Chang-Jin and Charles R. Nelson**, “Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle,” *Review of Economics and Statistics*, November 1999, 81 (4), 608–616.
- Lettau, Martin, Sydney Ludvigson, and Jessica Wachter**, “The Declining Equity Premium: What Role Does Macroeconomic Risk Play?,” *Review of Financial studies*, 2006, *Forthcoming*.
- Lipster, Robert S. and Albert N. Shiryaev**, *Statistics of Random Processes, Vol I and II*, New York: Springer-Verlag, 2000.
- Piazzesi, Monika and Martin Schneider**, “Equilibrium Yield Curves,” *NBER Macroeconomics Annual*, 2006, 21 (1), 389–442.
- Sims, Christopher A. and Tao Zha**, “Were There Regime Switches in U.S. Monetary Policy?,” *American Economic Review*, March 2006, 96 (1), 54–81.
- Stock, James H. and Mark W. Watson**, “Has the business cycle changed?,” *Proceedings, Federal Reserve Bank of Kansas City*, 2003, pp. 9–56.

- Taylor, John B.**, “Housing and Monetary Policy,” Working paper, Stanford University September 2007.
- Veronesi, Pietro**, “Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model,” *Review of Financial Studies*, Winter 1999, *12* (5), 975–1007.
- , “How Does Information Quality Affect Stock Returns?,” *Journal of Finance*, 04 2000, *55* (2), 807–837.
- Wachter, Jessica A.**, “A consumption-based model of the term structure of interest rates,” *Journal of Financial Economics*, February 2006, *79* (2), 365–399.
- Whelan, Karl**, “A Guide to U.S. Chain Aggregated NIPA Data,” *Review of Income and Wealth*, June 2002, *48* (2), 217–33.
- Working, Holbrook**, “Note on the Correlation of First Differences of Averages in a Random Chain,” *Econometrica*, October 1960, *28* (4), 916–918.