BOND RISK PREMIA IN CONSUMPTION-BASED MODELS

Drew D. Creal
Jing Cynthia Wu

Bond Risk Premia in Consumption-based Models
Drew D. Creal and Jing Cynthia Wu
NBER Working Paper No. 22183
April 2016
JEL No. C11,E31,E43,E52,G12

ABSTRACT

Workhorse Gaussian affine term structure models (ATSMs) attribute time-varying bond risk premia entirely to changing prices of risk, while structural models with recursive preferences credit it completely to stochastic volatility. We reconcile these competing channels by introducing a novel form of external habit into an otherwise standard model with recursive preferences. The new model has an ATSM representation with analytical bond prices making it empirically tractable. We find that time variation in bond term premia is predominantly driven by the price of risk, especially, the price of expected inflation risk that co-moves with expected inflation itself.

Drew D. Creal
University of Chicago
Booth School of Business
5807 South Woodlawn Ave
Chicago, IL 60637
dcreal@chicagobooth.edu

Jing Cynthia Wu
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
cynthia.wu@chicagobooth.edu

1 Introduction

Term premia, risk premia in the bond market, are a key object of interest for central banks. They influence how monetary policy implementation via the short term interest rate gets transmitted into the real economy through borrowing costs at longer maturities, and ultimately determines its effectiveness. Policy speeches – made by the current Fed Chair Yellen(2014) and past Chairs Bernanke(2006) and Greenspan(2005), for example – call for the importance of understanding how and why term premia fluctuate. Our paper aims to answer these questions by proposing a new consumption based asset pricing model.

Central banks around the world rely on reduced form Gaussian affine term structure models (ATSM) to produce estimates of term premia for policy discussion, because of the tractability and reliability resulting from the affine structure. This class of models generate variability in term premia through a time-varying price of risk that is a function of the conditional mean of yields; see, e.g. Duffee(2002), Wright(2011) and Bauer, Rudebusch, and Wu(2012). Conversely, structural consumption-based asset pricing models studying bond risk premia often use recursive preferences, Bansal and Shaliastovich(2013) for example. In these models, time-variation in term premia are driven only by stochastic volatility of consumption growth and inflation, meaning that the levels of these macro variables play no role. This conclusion is at odds with the empirical evidence from reduced form ATSMs.

We reconcile the two literatures by building a structural model with both time-varying prices and quantities of risk. We introduce a time-varying price of risk through a new formulation of habit formation. Just as in Abel(1999), the agent's period utility relies on the ratio of consumption to habit. Our habit specification depends on current and past consumption and inflation, with the latter accounting for inflation non-neutrality as argued by Piazzesi and Schneider(2007) and Bansal and Shaliastovich(2013). The novelty is that the state dependent risk sensitivity function – how habit growth loads on economic shocks – is chosen to make bond prices analytical and retain an affine structure. The tractability gained from the affine structure allows us to empirically disentangle the roles that habit,

recursive preferences, and stochastic volatility have on term premia. We introduce time-varying quantities of risk through stochastic volatility, similar to the long run risk literature. The difference is that the volatility process in our paper is guaranteed to remain positive, unlike most of the literature.

We evaluate the empirical performance of our model by first quantifying the information contained in observed inflation and consumption data about macroeconomic state variables using Markov chain Monte Carlo and particle filters. We then ask how well these macroeconomic state variables do in terms of matching the bond yield data by least squares. We show that empirically our model can adequately capture the time-variation of term premia. It fits other key moments of Treasury bonds as well: it has an upward slope for the yield curve; and it also mimics the time series dynamics of the average yield and slope of the yield curve well.

Next, we turn to the key question that motivates our research: is a time-varying price or quantity of risk the key driver for the time variation in term premia? We answer this question by shutting down one channel at a time. First, we shut down the price of risk channel. This model is similar to a long run risk model with stochastic volatility studied in the literature. We find that although such a model produces time variation in term premia, the implied term premia are implausible: they are economically insignificant, and have the wrong sign.

On the other hand, a model with habit but not stochastic volatility mimics the time variation of term premia produced by the reduced form Gaussian ATSM and our benchmark model. Overall, our empirical evidence attributes the time variation in the term premia primarily to a time-varying price of risk through habit.

We further examine whether it is inflation or consumption that drives this price of risk, and we find the crucial component is the price of the expected inflation risk that comoves with the expected inflation itself. This is consistent with inflation non-neutrality as argued by Piazzesi and Schneider(2007) and Bansal and Shaliastovich(2013) in a structural framework.

The important contribution inflation makes to bond price dynamics are also highlighted in the ATSM literature, see Ang and Piazzesi(2003) and Rudebusch and Wu(2008).

Introducing habit also has important implications for the unconditional slope of the yield curve, which has been the main focus for the majority of the literature. Another counterfactual implication of the long run risk model with stochastic volatility is a downward sloping yield curve, once the macro latent factors are pinned down by the observed macro data. Adding habit formation reverts the situation completely, and implies an unconditional slope just like what we see in the data.

Habit formation grants an economic interpretation to a long standing intuition in the ATSM literature: In a structural model without habit, the key parameters that determine the autocorrelation of consumption growth and inflation are the same as the parameters controlling the slope of the yield curve. This strong link between the two parameters has undesirable implications. For example, in our sample, the time series properties of the data dictate a downward sloping yield curve. Conversely, Gaussian ATSMs separate these two parameters. This property of the model is the main reason for their success in fitting the data because it is the mechanism that allows the price of risk to be time-varying. Habit formation provides an economic motivation for this separation in a structural model.

This separation alone provides the same implications for bond slope and term premia even when other preference parameters of the model – including the time discount factor, intertemporal elasticity of substitution and risk aversion – vary substantially over regions with different economic interpretations. We demonstrate this point using two local maxima with these three structural parameters taking economically different values yet with the same implications for bonds.

The empirical examination of the asset pricing implications discussed above requires solving for the stochastic discount factor.[1] However, such a solution does not always exist

---

[1]Specifically, we implement a solution method developed by Bansal and Yaron(2004) that is widely used in the macroeconomics and finance literatures; see, e.g. Bollerslev, Tauchen, and Zhou(2009), Bansal, Kiku, and Yaron(2012) and Schorfheide, Song, and Yaron(2014). Rudebusch and Swanson(2012) and Caldara, Fernández-Villaverde, Rubio-Ramírez, and Yao(2012) describe other solution methods.

for regions of the parameter space that researchers have traditionally found plausible. This is an issue for all models in the literature that use recursive preferences with a few exceptions, and not specific to our model. We provide conditions on the model's parameters guaranteeing the existence of a solution. The general rule is that agents cannot be too patient. A direct implication is that an extremely patient agent faces strong restrictions in their risk aversion and intertemporal elasticity of substitution. These conditions partition the parameter space and make it cumbersome for econometricians implementing either an optimization-based frequentist estimator or Bayesian Markov chain Monte Carlo algorithm.

This paper continues as follows. Subsection 1.1 discusses the literature. We introduce the new model with habit formation and recursive preferences in Section 2, and discusses its properties in Section 3. Section 4 describes empirical strategy for estimation and the consequent estimates. Section 5 examines the model's implication for bonds. In Section 6, we discuss the conditions for a solution to exist. The paper concludes in Section 7.

1.1 Relationship to the literature

Our specification for habit is motivated by Abel(1999) and Campbell and Cochrane(1999). Campbell and Cochrane(1999) introduce a risk sensitivity function that generates time-varying risk premia that are functions of the past history of shocks to consumption growth. Our model uses ratio habits as in Abel(1999) and allows the habit to be influenced by a risk sensitivity function as in Campbell and Cochrane(1999), albeit with a different functional form. The combination of the two allow the representative agent's period utility and their marginal rate of substitution to depend on the state of the economy. Wachter(2006) studies bonds with Campbell and Cochrane(1999) style habit formation. Dew-Becker(2014) introduces a multiplicative habit to study bonds in a DSGE model. Our primary contribution to the habit literature is that we develop a model that has an analytical bond prices, making this class of models tractable.

A large literature in macroeconomics and finance uses recursive preferences as developed

by Kreps and Porteus(1978), Epstein and Zin(1989), and Weil(1989). These preferences separate intertemporal substitution from risk aversion, which are directly linked under power utility. An endowment economy with recursive preferences and affine dynamics of the state variables is particularly attractive because it generates (approximate) closed-form solutions for bond and equity prices. A prime example is the long-run risk model of Bansal and Yaron(2004). Several authors have studied the yield curve using this framework. Piazzesi and Schneider(2007) study an economy with recursive preferences when the elasticity of intertemporal substitution equals one and shocks to consumption growth and inflation are homoskedastic; see also, e.g. Tallarini(2000), and Hansen, Heaton, and Li(2008). Bansal and Shaliastovich(2013) evaluate a model where expected inflation and consumption growth are slow moving and have stochastic volatility. The difference between our work and these papers is the existence of habit formation, which is critical for introducing a time-varying price of risk.

Albuquerque, Eichenbaum, and Rebelo(2014) and Schorfheide, Song, and Yaron(2014) analyze models where the period utility has a time varying preference. We show how this formulation relates to habit formation. In their specifications, shocks to preferences introduce additional latent factors that can influence asset prices. But, these shocks do not introduce time-varying risk premia, which is the key focus of our paper.

Another class of consumption-based models that have recently drawn attention for their ability to explain asset pricing anomalies are models with rare consumption disasters; see, e.g. Barro(2006), Gabaix(2012), and Wachter(2013). Wachter(2013) considers a model with recursive preferences with a time-varying probability of a jump in consumption growth. When applied to the yield curve, this model will generate time-varying term premia that are driven by the time-varying intensities of jumps to consumption growth and/or inflation. This mechanism can be seen as an alternative to stochastic volatility for generating a time-varying quantity of risk. Note that the model has a constant price of risk.

2   Model

In this section, we build a structural model that encompasses the two competing channels that drive time variation in term premia: time-varying price and quantity of risk. We do so by introducing external habit into recursive preferences in a tractable framework. The novelty is two fold. First, unlike a standard model with recursive preferences, such a model generates a time-varying market price of risk. Moreover, different from the popular habit formation models in the literature, our model gains its tractability by retaining an affine form.

2.1   Agent's problem

We consider an endowment economy, where the representative agent optimizes over his lifetime utility

$$V_t \;=\; \max_{C_t} \left[(1-\beta) \left(\frac{C_t}{H_t} \right)^{1-\eta} + \beta \left\{ \mathrm{E}_t \left[V_{t+1}^{1-\gamma} \right] \right\}^{\frac{1-\eta}{1-\gamma}} \right]^{\frac{1}{1-\eta}} \tag{1}$$

with respect to consumption C_t. H_t is the level of the external habit. For the same amount of consumption, a higher habit makes the agent less happy. The same as Abel(1999), we assume that agent's utility depends on the ratio between consumption and habit. For example, doubling both the consumption and habit does not change the agent's utility. β is the time discount factor, γ measures risk aversion, and $\psi = \frac{1}{\eta}$ is the elasticity of intertemporal substitution when there is no uncertainty.

Agents maximize utility (1) subject to the budget constraint

$$W_{t+1} \;=\; (W_t - C_t) \, R_{c,t+1}, \tag{2}$$

where W_t is wealth and $R_{c,t+1}$ is the gross return on the consumption asset between t and $t+1$.

The first order condition of the agent's problem implies that the log stochastic discount factor (SDF) is

$$m_{t+1} = \vartheta \ln(\beta) + \vartheta \Delta v_{t+1} - \eta \vartheta \Delta c_{t+1} + (\vartheta - 1) r_{c,t+1}, \tag{3}$$

where $\vartheta \equiv \frac{1-\gamma}{1-\eta}$, $\Delta c_{t+1} = \ln(C_{t+1}) - \ln(C_t)$ is consumption growth, and $r_{c,t+1} = \ln(R_{c,t+1})$ is the continuously compounded return. These terms are standard in models with recursive preferences.

The new term $\Delta v_{t+1} \equiv (\eta - 1)(\ln H_{t+1} - \ln H_t)$ measures the stochastic growth rate of habit, and it is unique to our new model. It is this term that enables enough variation in the pricing kernel to capture the time-varying risk premium in bond prices through a time-varying price of risk. For a derivation, see Appendix A. Nominal assets are priced using the nominal pricing kernel

$$m_{t+1}^{\$} = m_{t+1} - \pi_{t+1}, \tag{4}$$

where inflation is $\pi_{t+1} = \ln(\Pi_{t+1}) - \ln(\Pi_t)$ and Π_t is the nominal price level.

2.2 Dynamics

The state of the economy is summarized by a $G \times 1$ vector g_t, which includes consumption growth Δc_t and inflation π_t

$$\Delta c_t = Z_c' g_t, \tag{5}$$

$$\pi_t = Z_\pi' g_t, \tag{6}$$

where Z_c and Z_π are $G \times 1$ selection vectors containing only zeros and ones. The state vector follows a heteroskedastic vector autoregressive process, summarized in companion form as

$$
\begin{aligned}
g_{t+1} &= \mu_g + \Phi_g g_t + \Phi_{gh} h_t + \Sigma_{gh}\varepsilon_{h,t+1} + \Sigma_{g,t}\varepsilon_{g,t+1} \qquad \varepsilon_{g,t+1} \sim \mathrm{N}(0, I) \qquad (7) \\
\Sigma_{g,t}\Sigma'_{g,t} &= \Sigma_{0,g}\Sigma'_{0,g} + \sum_{i=1}^{H}\Sigma_{i,g}\Sigma'_{i,g}h_{it} \\
h_{t+1} &\sim \mathrm{NCG}(\nu_h, \Phi_h, \Sigma_h) \qquad\qquad\qquad\qquad\qquad\qquad\qquad (8) \\
\varepsilon_{h,t+1} &= h_{t+1} - \mathrm{E}_t[h_{t+1}|h_t]
\end{aligned}
$$

where h_t is a $H \times 1$ vector following a non-central gamma (NCG) process as in Creal and Wu(2015b). This process guarantees the non-negativity of volatility. This contrasts with the Gaussian process that is prevalent in most of the literature, for example, see Bansal and Yaron(2004) and Bansal and Shaliastovich(2013).

This is an affine process that is the exact discrete time equivalent of a multivariate Cox, Ingersoll, and Ross(1985) process. The conditional mean is $\mathrm{E}_t[h_{t+1}|h_t] = \Sigma_h\nu_h + \Phi_h h_t$ meaning that Φ_h controls the autocovariance of h_{t+1} and $\Sigma_h\nu_h$ is the drift. Σ_h is a matrix of scale parameters and ν_h are a vector of shape parameters. The vector $\varepsilon_{h,t+1}$ are mean zero, heteroskedastic shocks to volatility and Σ_{gh} measures the covariance between Gaussian and non-Gaussian shocks, i.e. the volatility feedback effect. Further details on properties of the model can be found in Appendix B.[2]

The general process (7) - (8) nests popular models in the literature. Consider a model with long-run risks for consumption growth as in Bansal and Yaron(2004) and a time-varying

---

[2]The timing of how volatility scales the shocks in discrete-time models such as (7) is an outstanding issue in financial econometrics. Both timings $\Sigma_{g,t}\varepsilon_{g,t+1}$ and $\Sigma_{g,t+1}\varepsilon_{g,t+1}$ lead to log-SDF's that are linear in the state variables and produce affine bond prices. Our analysis focuses on the timing in (7), which is common in finance. However, the alternative timing $\Sigma_{g,t+1}\varepsilon_{g,t+1}$ means the representative agent faces greater short term uncertainty. Volatility has an immediate and potentially more significant impact on the short rate (and consequently all yields) even when there is no volatility feedback effect.

trend in inflation similar to Stock and Watson(2007). The model is

$$\pi_{t+1} \;\; = \;\; \bar{\pi}_t + \varepsilon_{\pi_1, t+1} \qquad\qquad\qquad \varepsilon_{\pi_1, t+1} \sim \mathrm{N}\left(0, h_{t, \pi_1}\right) \qquad (9)$$

$$\Delta c_{t+1} \;\; = \;\; \bar{c}_t + \varepsilon_{c_1, t+1} \qquad\qquad\qquad \varepsilon_{c_1, t+1} \sim \mathrm{N}\left(0, h_{t, c_1}\right) \qquad (10)$$

$$\bar{\pi}_{t+1} \;\; = \;\; \mu_\pi + \phi_\pi \bar{\pi}_t + \phi_{\pi, c} \bar{c}_t + \varepsilon_{\pi_2, t+1} \qquad \varepsilon_{\pi_2, t+1} \sim \mathrm{N}\left(0, h_{t, \pi_2}\right) \qquad (11)$$

$$\bar{c}_{t+1} \;\; = \;\; \mu_c + \phi_{c, \pi} \bar{\pi}_t + \phi_c \bar{c}_t + \sigma_{c, \pi} \varepsilon_{\pi_2, t+1} + \varepsilon_{c_2, t+1} \qquad \varepsilon_{c_2, t+1} \sim \mathrm{N}\left(0, h_{t, c_2}\right) \qquad (12)$$

where $\bar{c}_t$ is the expected consumption growth rate and $\bar{\pi}_t$ is expected inflation. The shocks $\varepsilon_{c_1, t}$ and $\varepsilon_{\pi_1, t}$ are transitory, and determine the high-frequency movements in their respective series whereas $\varepsilon_{\pi_2, t}$ and $\varepsilon_{c_2, t}$ are shocks to their persistent components. In our model, shocks to expected inflation have a contemporaneous impact on expected consumption growth, and all shocks have stochastic volatility. The state vectors are $g_t = \left(\pi_t, \Delta c_t, \bar{\pi}_t, \bar{c}_t\right)'$ and $h_t = \left(h_{t, \pi_1}, h_{t, c_1}, h_{t, \pi_2}, h_{t, c_2}\right)'$. See Appendix B.2 for more details.

Although our empirical implementation will focus on this popular long run risk specification, note that our general framework in (7) - (8) can also nest a vector autoregressive moving average model for consumption growth and inflation, like the one studied in Wachter(2006).

2.3 Habit

The empirical conclusion in the GATSM literature is that risk premia are driven by the levels of the state variables through the price of risk. Motivated by this empirical finding, we build this channel in our model through the time varying growth rate of habit. We discipline our model using this empirical fact, and only allow habit growth to depend on the levels of macroeconomic variables. Then we will empirically test whether such a parsimonious specification generates enough variation in the term premium from our structural model to capture the data. The habit growth is

$$\Delta v_{t+1} \;\; = \;\; \Lambda_1\left(g_t\right) + \Lambda_2\left(g_t\right)' \varepsilon_{g, t+1}, \qquad\qquad (13)$$

which depends on the levels of past inflation and consumption as well as their shocks in the current period. The dependence on inflation is motivated by the real effect of inflation as argued by Piazzesi and Schneider(2007) and Bansal and Shaliastovich(2013).

The key term, the risk sensitivity function $\Lambda_2(g_t)$, works through a similar mechanism as in Campbell and Cochrane(1999) and Wachter(2006) to generate a time-varying price of risk, and hence risk premium. The difference is that our choice of functional form allows the model to stay within the affine family and have analytical bond prices:

$$\Lambda_2(g_t) = -\eta\Sigma_{g,t}^{-1}(\lambda_0 + \lambda_g g_t), \tag{14}$$

and $\Lambda_1(g_t) = -\frac{\vartheta\eta^2}{2}(\lambda_0 + \lambda_g g_t)'\left(\Sigma_{g,t}\Sigma_{g,t}'\right)^{-1}(\lambda_0 + \lambda_g g_t)$. This specification can be readily extended to a more complex model where changes in habit Δv_{t+1} also depend on the volatilities and their shocks, see the earlier working paper version of Creal and Wu(2015a). This extension could potentially introduce more complex channels to explain movements in asset prices. Instead, we keep the model intentionally simple and examine how well the current channels can explain yields.

When $\gamma = \eta$, the model with recursive preferences reduces to power utility with habit. In this case, if one chooses

$$\Lambda_{1t} = (1 - \phi)(\bar{v} - v_t)$$
$$\Lambda_{2t} = \frac{1}{\bar{H}}\sqrt{\eta + 2(v_t - \bar{v})} + \eta\sigma_c,$$

$\Delta c_{t+1} = \bar{c} + \sigma_c\varepsilon_{c_1,t+1}$, and $\varepsilon_{g,t+1} = \varepsilon_{c_1,t+1}$, then the resulting SDF in (3) becomes the SDF of the habit formation model in Wachter(2006). Consequently, asset prices in the two models are observationally equivalent.

There are two main differences. First, our choice for the risk sensitivity functions yield an affine structure. This implies analytical bond prices which provide tractability. Second, we allow not only consumption, but inflation risk to be priced. It turns out the price of

expected inflation is the key driving factor.

2.4 Solving $r_{c,t+1}$

The SDF in (3) is a function of the return on the consumption asset $r_{c,t+1}$, which is generally regarded as unobserved in the data. We solve for it using the log-linearization technique of Campbell and Shiller(1989), applied by, for example, Bansal and Yaron(2004) and Bansal, Kiku, and Yaron(2012).[3] We express the return as a function of the price to consumption ratio

$$
\begin{aligned}
r_{c,t+1} & \equiv \ln\left(\frac{P_{t+1}+C_{t+1}}{P_t}\right) = \Delta c_{t+1} - pc_t + \ln\left(1+\exp\left(pc_{t+1}\right)\right) \\
& \approx \kappa_0 + \kappa_1 pc_{t+1} - pc_t + \Delta c_{t+1},
\end{aligned}
\tag{15}
$$

where P_{t+1} is the price of consumption goods, $pc_t = \ln\left(\frac{P_t}{C_t}\right)$ is the log price to consumption ratio. κ_0 and κ_1 are log-linearization constants that depend on the average price to consumption ratio $\bar{pc} = \mathrm{E}\left[pc_t\right]$. The derivation can be found in Appendix C.1.

As the real pricing kernel in (3) must also price the consumption good

$$
1 = \mathrm{E}_t\left[\exp\left(m_{t+1} + r_{c,t+1}\right)\right]
\tag{16}
$$

we can guess and verify a solution for

$$
pc_t = D_0 + D'_g g_t + D'_h h_t.
\tag{17}
$$

This is a fixed point problem: pc_t depends on κ_0, κ_1 through D_0, D_g, D_h, which in turn depend on $\bar{pc}$. We discuss the details of this fixed point problem in Appendix C, and the existence of a solution in Section 6.

---

[3]The solution method used by Campbell, Giglio, Polk, and Turley(2014) for their ICAPM model is similar, only they substitute out consumption instead of the return on the consumption asset.

Finally, (15) and (17) together express $r_{c,t+1}$ and hence the pricing kernel m_{t+1} as functions of the underlying state variables g_t and h_t.

2.5 Relation to preference shock

The habit H_t can be re-parameterized to take on the interpretation of time-varying preference in the recent macroeconomics literature, see Albuquerque, Eichenbaum, and Rebelo(2014) and Schorfheide, Song, and Yaron(2014). If we define $\Upsilon_t \equiv H_t^{\eta-1}$, then (1) becomes

$$V_t = \max_{C_t} \left[(1-\beta)\Upsilon_t C_t^{1-\eta} + \beta \left\{ \mathrm{E}_t \left[V_{t+1}^{1-\gamma} \right] \right\}^{\frac{1-\eta}{1-\gamma}} \right]^{\frac{1}{1-\eta}}, \tag{18}$$

where Υ_t is the time preference. To demonstrate the basic intuition why these two interpretations are equivalent, let us suppose $\eta < 1$ for now. When the habit H_t is higher, agents become less satisfied with the same amount of consumption. This is equivalent to saying that agents reduce their preference for consumption this period, hence a smaller Υ_t.

The preference shock can be defined as $\Delta v_{t+1} = \ln \Upsilon_{t+1} - \ln \Upsilon_t$. If we re-parameterize the process in (13) as $\Delta v_{t+1} = Z_v' g_{t+1}$ and allow Δv_{t+1} to be represented by a new latent factor, then this gets us the specification in the macroeconomics literature, for example, Albuquerque, Eichenbaum, and Rebelo(2014) and Schorfheide, Song, and Yaron(2014). Note, we can readily incorporate this ingredient into our model; see the original working paper version of Creal and Wu(2015a). Adding this feature is expected to improve the model's fit to the data because the role of a preference shock factor is similar to the latent factors that are present in reduced form Gaussian ATSMs. However, a key difference is that the latent preference shock factors do not introduce a time-varying market price of risk, which is one of the key ingredients of our paper.

3 Model properties

This section examines the different components contributing to time variation in bond risk premia. Specifically, we examine how the new specification of habit in the previous section translates into a time-varying price of risk, and hence time-varying risk premium. We then analyze habit's implication for bond prices.

3.1 Sources of risk premia

Using the solution method described, in Subsection 2.4, the nominal log-SDF in deviation from the mean form becomes

$$m_{t+1}^{\$} - \mathrm{E}_t\left[m_{t+1}^{\$}\right] \quad = \quad -\lambda_{g,t}^{\$,\prime}\Sigma_{g,t}\varepsilon_{g,t+1} - \lambda_h^{\$,\prime}\Sigma_{h,t}\tilde{\varepsilon}_{h,t+1} \tag{19}$$

where the vector of shocks to volatility $\tilde{\varepsilon}_{h,t+1} = \Sigma_{h,t}^{-1}\varepsilon_{h,t+1}$ have been standardized to have unit variance. Due to the risk sensitivity functions, shocks to the SDF are heteroskedastic with time-varying price of level risks $\lambda_{g,t}^{\$}$ being

$$
\begin{aligned}
\lambda_{g,t}^{\$} \quad = \quad & \gamma Z_c + Z_\pi && \leftarrow \text{power utility} \\
& +\kappa_1 \frac{\gamma - \eta}{1 - \eta} D_g && \leftarrow \text{recursive preferences} \\
& +\vartheta\eta\left(\Sigma_{g,t}\Sigma_{g,t}'\right)^{-1}\left(\lambda_0 + \lambda_g g_t\right). && \leftarrow \text{habit formation} \tag{20}
\end{aligned}
$$

The first two terms are inherited from power utility, the second line comes from recursive preferences, and the terms in the third line are due to habit.

The key term in (20) is $\lambda_g g_t$. Only when λ_g is non-zero does the model have a time-varying price of level risk and produce time-varying term premia that are functions of g_t. This channel remains the same even when we shut off the stochastic volatility in the dynamics (7) - (8). Conversely, if there were no habit, i.e. $\lambda_0 = 0, \lambda_g = 0$, the price of level risk is a constant as in the literature. A time-varying price of risk that co-moves with the levels of

macroeconomic variables and yields is a feature of Gaussian ATSMs that makes it successful empirically, and it is a feature that is absent in standard models with recursive preferences.

If there were no habit, i.e. if $\lambda_0 = 0$ and $\lambda_g = 0 \Rightarrow H_t = 1$, then the price of short run consumption risk is equal to the risk aversion coefficient γ, the price of short run inflation risk is 1 for the nominal pricing kernel, and 0 for the real pricing kernel. These are consistent with the literature.

Next, we decompose the prices of risk for the non-Gaussian shocks as

$$
\begin{aligned}
\lambda_h^{\$} \;=\; & \Sigma_{gh}' \left(\gamma Z_c + Z_\pi \right) && \leftarrow \text{power utility} \\
& + \kappa_1 \frac{(\gamma - \eta)}{(1 - \eta)} \left(\Sigma_{gh}' D_g + D_h \right) && \leftarrow \text{recursive preference}
\end{aligned}
$$

These terms have similar features and functional forms as those in (20). Power utility only has an impact on the price of volatility risk if there is a volatility feedback effect and $\Sigma_{gh} \neq 0$, while recursive preferences generates a price of risk even when $\Sigma_{gh} = 0$. To keep our model simple and tractable, as in the standard models with recursive preferences, our model does not have time-varying prices of volatility risk. Like the standard models, we also have time-varying quantities of risk through stochastic volatility.

3.2   Bond prices and term premium

The price of a zero-coupon nominal bond with maturity n at time t is the expected price of the same asset at time $t + 1$ discounted by the stochastic discount factor

$$
P_t^{\$,(n)} \;=\; \mathrm{E}_t \left[\exp \left(m_{t+1}^{\$} \right) P_{t+1}^{\$,(n-1)} \right]. \tag{21}
$$

Affine representation   Following Creal and Wu(2015b), nominal yields are an affine function of both the Gaussian state vector and volatility

$$
y_t^{\$,(n)} \;=\; a_n^{\$} + b_{n,g}^{\$,\prime} g_t + b_{n,h}^{\$,\prime} h_t, \tag{22}
$$

15

where the key coefficient $b_{n,g}^\$ = -\frac{1}{n}\bar{b}_{n,g}^\$$, and $\bar{b}_{n,g}^\$$ follows this recursion

$$\bar{b}_{n,g}^\$ = \left(\Phi_g - \eta\vartheta\lambda_g\right)'\bar{b}_{n-1,g}^\$ + \bar{b}_{1,g}^\$. \tag{23}$$

This and the rest of the bond-loadings detailed in Appendix D are similar to those found in ATSMs.

One key insight from the Gaussian ATSM literature is $\Phi_g^{\mathbb{Q}^\$} \equiv \Phi_g - \eta\vartheta\lambda_g \neq \Phi_g$. The basic intuition is this separates the autoregressive coefficient driving the physical dynamics Φ_g from the coefficient driving the cross section slope of the yield curve $\Phi_g^{\mathbb{Q}^\$}$. It is this feature that enables Gaussian ATSM to fit the data and generate enough variation in term premia. We micro found this key channel through habit. Another contribution of our paper is to map a consumption-based model with habits and recursive preferences into an affine framework.

Consumption-inflation representation   To gain more economic insight, we can express yields as functions of expected consumption growth and expected inflation. Assume $\Sigma_{gh} = 0$ for the sake of intuition, the short term nominal interest rate is

$$
\begin{aligned}
r_t^\$ &\equiv -\log\left(P_t^{\$,(1)}\right) \\
&= -\ln\left(\beta\right) + \eta\mathrm{E}_t\left[\Delta c_{t+1}\right] + \mathrm{E}_t\left[\pi_{t+1}\right] \\
&\quad -\eta\vartheta(\eta Z_c + Z_\pi)'(\lambda_0 + \lambda_g g_t) \\
&\qquad +\text{Jensen's ineq.}
\end{aligned}
$$

It captures the time discount through β, expected consumption growth, and expected inflation. The third line adjusts for risk compensation, and is newly introduced in our model through the risk sensitivity function in habit specification.

If consumption growth increases, then the short term interest rate increases between t and $t + 1$. This is due to the agent's motive to smooth intertemporally. With relatively higher expected consumption tomorrow, the representative agent borrows today against

tomorrow's consumption goods to smooth consumption. This pushes the interest rate higher. If consumption growth is expected to be negative, this motivates the agent to save. They are willing to do so even at a lower nominal interest rate. If inflation is expected to increase, then the nominal interest rate is higher.

Term premium The nominal term premium is defined as the difference between the model implied yield $y_t^{\$,(n)}$ and the average of expected future short rates over the same period

$$tp_t^{\$,(n)} \quad = \quad y_t^{\$,(n)} - \frac{1}{n}\mathrm{E}_t \left[r_t^\$ + r_{t+1}^\$ + \ldots, +r_{t+n-1}^\$ \right]. \tag{24}$$

The term premium has a simple portfolio interpretation. An investor can buy an n-period bond and hold it until maturity or he can purchase a sequence of 1 period bonds, repeatedly rolling them over for n periods. The term premium measures the additional compensation a risk averse agent needs to choose one option over another. Under the expectations hypothesis, term premia are constant. For Gaussian models with homoskedastic shocks, this coincides with setting $\lambda_g = 0$ and eliminating time-variation in the risk sensitivity functions.

4   Estimation

Stacking (22) in order for N different maturities $n_1, n_2, ..., n_N$ and adding a vector of pricing errors e_t, the observation equations for yields are

$$y_t^\$ = A^\$ + B_g^\$ g_t + B_h^\$ h_t + e_t, \quad e_t \sim \text{i.i.d.} \left(0, \Omega\right) \tag{25}$$

where $y_t^\$ = \left(y_t^{\$,(n_1)}, y_t^{\$,(n_2)}, \ldots, y_t^{\$,(n_N)} \right)$, $A^\$ = (a_{n_1}^\$, \ldots, a_{n_N}^\$)'$, $B_g^\$ = (b_{g,n_1}^{\$,\prime}, ..., b_{g,n_N}^{\$,\prime})'$, and $B_h^\$ = (b_{h,n_1}^{\$,\prime}, ..., b_{h,n_N}^{\$,\prime})'$.

Let $\theta^{\mathbb{Q}} = (\beta, \gamma, \psi, \theta^\lambda)$ denote the structural parameters that enter the bond loadings, where θ^λ are the habit parameters. We estimate $\theta^{\mathbb{Q}}$ by minimizing the average of the squared

pricing errors e_t as

$$\min \frac{1}{NT} \sum_{t=1}^{T} e_t' \Omega^{-1} e_t. \tag{26}$$

Computing the pricing error e_t involves estimating the latent state variables g_t and h_t, which capture aspects of inflation and consumption and their volatilities we do not directly observe from the data.

To quantify the latent variables and capture uncertainty around them, we resort to their joint posterior distributions with corresponding parameters in (7) - (8) given the observed macroeconomic data $p(g_t, h_t, \theta^{\mathbb{P}}|m_{1:T})$. This distribution is approximated by a Particle Gibbs sampler detailed in Subsection 4.2.

We then minimize (26) with respect to $\theta^{\mathbb{Q}}$ where the pricing error is computed with $e_t = y_t^{\$} - A^{\$} - B_g^{\$}\hat{g}_t - B_h^{\$}\hat{h}_t$, and $\hat{g}_t, \hat{h}_t, \hat{\theta}^{\mathbb{P}}$ are evaluated at their posterior means. This procedure is similar to Piazzesi and Schneider(2007), who estimated a stochastic process for consumption and inflation and calibrated the structural parameters. The difference is that we estimate these parameters by minimizing the pricing errors. This procedure has the benefit of ensuring that the latent factors (expected inflation, expected consumption growth, and their stochastic volatilities) maintain their intended economic interpretation because it only uses macroeconomic data to extract them. Given the macroeconomic factors, we then ask the question: how much variation in asset prices can we explain with the structural model? See details in Subsection 4.3. We then comment on the estimation method popular in the literature in Subsection 4.4.

4.1 Data and restrictions

Data The data we use are standard in the literature. Our measure of monthly real per capita consumption growth is constructed from nominal non-durables and services data downloaded from the NIPA tables at the U.S. Bureau of Economic Analysis. We deflate

each of these series by their respective price indices, add them together, and divide by the civilian population. The population series and monthly U.S. CPI inflation are downloaded from the Federal Reserve Bank of St. Louis. Yields are the Fama and Bliss(1987) zero coupon bond data available from the Center for Research in Securities Prices (CRSP) with maturities of (3, 12, 24, 36, 48, 60) months. The data spans from February 1959 through June 2014 for a total of $T = 665$ observations.

Parameter restrictions For the dynamics of consumption and inflation, $\Sigma_{0,g}$ is imposed to be 0 for identification.[4] We also impose Φ_h, Σ_h to be diagonal, and $\Phi_{gh} = 0$ and $\Sigma_{gh} = 0$ for simplicity.

If there is no habit, there are three structural parameters (β, ψ, γ) to fit the cross section of yields. In models with habit, to keep the model simple, we only introduce four new parameters into the matrix λ_g (the elements related to expected consumption, expected inflation, and their cross terms) while we set $\lambda_0 = 0$. There are a total of seven free structural parameters. We assume the variance for the pricing errors is the same across different maturities $\Omega = \omega^2 I$.

4.2 Estimation of macroeconomic factors

We quantify the distribution of the latent macroeconomic factors related to consumption growth and inflation by Bayesian methods. We use a particle Gibbs sampler which is an MCMC algorithm that uses a particle filter to draw from distributions that are intractable; see Creal and Wu(2015b) for an application on interest uncertainty, and Creal(2012) for a survey on particle filtering.

Given a prior distribution $p\left(\theta^{\mathbb{P}}\right)$ for the $\mathbb{P}$ parameters in (7) - (8), we sample from the joint posterior distribution

$$p\left(\theta^{\mathbb{P}}, g_{1:T}, h_{0:T}|m_{1:T}\right) \quad \propto \quad p\left(m_{1:T}|g_{1:T}, h_{0:T}, \theta^{\mathbb{P}}\right) p\left(g_{1:T}|h_{0:T}, \theta^{\mathbb{P}}\right) p\left(h_{0:T}|\theta^{\mathbb{P}}\right) p\left(\theta^{\mathbb{P}}\right) \quad (27)$$

---

[4]This assumption is lifted in the Gaussian model with $h_t = 0$.

Table 1: Estimates of time series parameters

μ_π	$\bar{\mu}_\pi$	ϕ_π	$\phi_{\pi,c}$	$-$	$\bar{\mu}_{h,\pi_2}$	ϕ_{h,π_2}	σ_{h,π_2}	$\bar{\mu}_{h,\pi_1}$	ϕ_{h,π_1}	σ_{h,π_1}
$-$	0.0033	0.975	0		$1.97e^{-7}$	0.975	$1.70e^{-11}$	$4.04e^{-6}$	0.975	$6.23e^{-9}$
	$(0.708e^{-3})$	(0.10)	(0.05)		$(9.20e^{-7})$	(0.009)	$(2.51e^{-11})$	$(5.08e^{-6})$	(0.009)	$(5.35e^{-10})$
μ_c	$\bar{\mu}_c$	$\phi_{c,\pi}$	ϕ_c	$\sigma_{h,c}$	$\bar{\mu}_{h,c_2}$	ϕ_{h,c_2}	σ_{h,c_2}	$\bar{\mu}_{h,c_1}$	ϕ_{h,c_1}	σ_{h,c_1}
$-$	0.0015	0	0.90	0.00	$3.77e^{-8}$	0.975	$1.30e^{-12}$	$1.08e^{-5}$	0.975	$1.02e^{-8}$
	$(0.750e^{-3})$	(0.05)	(0.10)	(3.5)	$(8.32e^{-8})$	(0.009)	$(1.73e^{-12})$	$(6.57e^{-6})$	(0.009)	$(1.57e^{-9})$
μ_π	$\bar{\mu}_\pi$	ϕ_π	$\phi_{\pi,c}$	$-$	$\bar{\mu}_{h,\pi_2}$	ϕ_{h,π_2}	σ_{h,π_2}	$\bar{\mu}_{h,\pi_1}$	ϕ_{h,π_1}	σ_{h,π_1}
$-0.096e^{-4}$	0.0031	0.978	0.057		$0.269e^{-6}$	0.984	$0.317e^{-9}$	$0.333e^{-5}$	0.989	$0.724e^{-8}$
$0.734e^{-4}$	$(0.497e^{-3})$	(0.014)	(0.033)		$(0.091e^{-6})$	(0.007)	$(0.120e^{-9})$	$(0.077e^{-5})$	(0.004)	$(0.0729e^{-8})$
μ_c	$\bar{\mu}_c$	$\phi_{c,\pi}$	ϕ_c	$\sigma_{c,\pi}$	$\bar{\mu}_{h,c_2}$	ϕ_{h,c_2}	σ_{h,c_2}	$\bar{\mu}_{h,c_1}$	ϕ_{h,c_1}	σ_{h,c_1}
$0.856e^{-4}$	0.0014	-0.002	0.941	-0.394	$0.664e^{-6}$	0.980	$0.340e^{-10}$	$0.891e^{-5}$	0.992	$0.137e^{-7}$
$(0.592e^{-4})$	$(0.278e^{-3})$	(0.010)	(0.029)	(0.190)	$(0.372e^{-6})$	(0.009)	$(0.169e^{-10})$	$(0.216e^{-5})$	(0.003)	$(0.225e^{-8})$

Prior (top) and posterior (bottom) mean and standard deviation (in parentheses) of our benchmark model in (9)-(12).

where $m_t = (\Delta c_t, \pi_t)$ and $x_{t:t+k} = (x_t, \ldots, x_{t+k})$. Starting with an initial value for the parameters $\theta^{\mathbb{P},(0)}$, the particle Gibbs sampler draws from this distribution by iterating for $j = 1, \ldots, M$ between the two full conditional distributions

$$
(g_{1:T}, h_{0:T})^{(j)} \sim p\left(g_{1:T}, h_{0:T} | m_{1:T}, \theta^{\mathbb{P},(j-1)}\right) \tag{28}
$$

$$
\theta^{\mathbb{P},(j)} \sim p\left(\theta^{\mathbb{P}} | m_{1:T}, g_{1:T}^{(j)}, h_{0:T}^{(j)}\right) \tag{29}
$$

This produces a Markov chain whose stationary distribution is the posterior (27). The models for consumption growth and inflation (9)-(12) are non-linear, non-Gaussian state space models. In these models, the full conditional distribution of the latent state variables given the data and model's parameters (28) is not easy to sample. The particle Gibbs sampler overcomes this limitation by using a particle filter to jointly sample paths of the state variables $(g_{1:T}, h_{0:T})$ in large blocks. Consequently, it improves the mixing of the MCMC algorithm and the efficiency with which the Markov chain explores the parameter space. Further details of the algorithm can be found in Appendix F.1, see also Creal and Tsay(2015) for a longer discussion.

Using the particle Gibbs sampler, we estimate the long-run risk model of consumption and inflation in (9)-(12). Posterior means and standard deviations for the parameters of the model are in Table 1. Filtered (in red) and smoothed (in blue) estimates of the latent state

Figure 1: Estimated factors from the long run risk model.
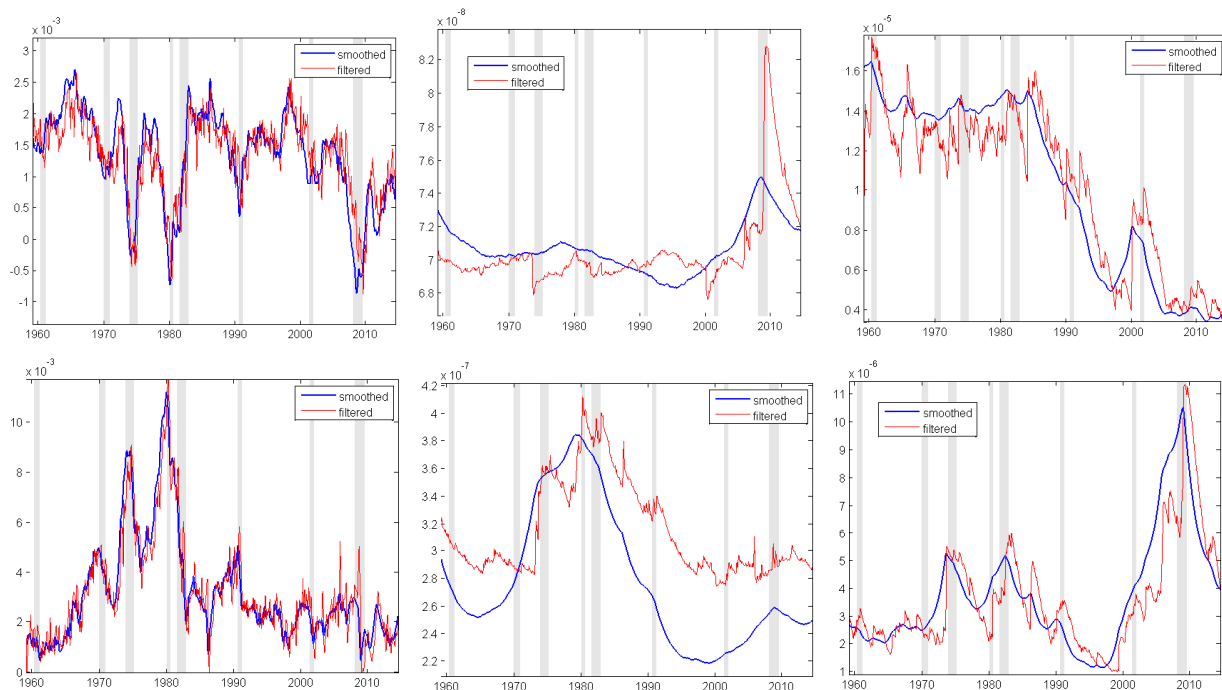
Filtered (red) and smoothed (blue) estimates of the factors from the long-run risk model. Top row is consumption growth and bottom row is inflation. Top left to right: long-run risk $\bar{c}_t$, variance of long-run risk h_{t,c_2}, variance of high-frequency component h_{t,c_1}. Bottom left to right: inflation trend $\bar{\pi}_t$, variance of expected inflation h_{t,π_2}, variance of high-frequency component h_{t,π_1}.

variables are plotted in Figure 1.

There is considerable variation in the long-run risk factor $\bar{c}_t$ of consumption growth (top left). It shows a noticeable decline during each recession, with the largest decline during the Great Recession. The pattern replicates the long run risk in the literature. While the volatility of the long run growth rate (top middle) is economically small and does not vary much, the stochastic volatility of the high frequency component (top right) is larger with more variation. The volatility of trend inflation increases during the mid-1970's, peaks during the early 1980's, and declines gradually until the mid 1980s and keeps at a low level afterwards. The estimates of stochastic volatility from this model of inflation are similar to those found by Stock and Watson(2007) and Creal(2012).

4.3 Cross sectional regression

Given the estimated parameters $\hat{\theta}^{\mathbb{P}}$ and the factors $\hat{g}_t$ and $\hat{h}_t$ from Subsection 4.2, we estimate the structural parameters $\theta^{\mathbb{Q}} = (\beta, \gamma, \psi, \theta^{\lambda})$ through non-linear least squares. We compute the pricing error by

$$e_t = y_t^{\$} - A^{\$}\left(\theta^{\mathbb{Q}}, \hat{\theta}^{\mathbb{P}}\right) - B_g^{\$}\left(\theta^{\mathbb{Q}}, \hat{\theta}^{\mathbb{P}}\right)\hat{g}_t - B_h^{\$}\left(\theta^{\mathbb{Q}}, \hat{\theta}^{\mathbb{P}}\right)\hat{h}_t, \tag{30}$$

and then minimize the objective function in (26) with respect to $\theta^{\mathbb{Q}}$.

The structural parameter estimates are in Table 2. With the restrictions in Subsection 4.1, $\theta^{\lambda} = \lambda_g$. Note that estimating $(\beta, \gamma, \psi, \lambda_g)$ is equivalent to estimating $(\beta, \gamma, \psi, \Phi_g^{\mathbb{Q}^{\$}})$. We implement the latter as we have a better prior knowledge of the scale of $\Phi_g^{\mathbb{Q}^{\$}}$. The left panel reports the global maximum. The time discount factor β is 0.9998, the intertemporal rate of substitution ψ is 1.02, and the risk aversion parameter γ is about 7. The risk-neutral autoregressive matrix $\Phi_g^{\mathbb{Q}^{\$}}$ is much more persistent than its time series counterpart Φ_g in Table 1, with both eigenvalues around 0.995. This high persistence generates the upward sloping yield curve discussed in Subsection 5.3. In the right panel, we report a local maximum. This maximum also generates an upward sloping yield curve as the global, see Table 4, and its implied term premium is qualitatively similar to the global maximum. However, the structural parameters take economically different values.

There exists a long standing debate about the qualitative feature of the structural parameters, for example, whether $\psi > 1$. The long run risk literature (Bansal and Yaron(2004)) argues that values of $\psi > 1$ should be the case. On the other hand, Campbell(2011) argues the opposite to be consistent with the aggregate evidence. Although the estimate in the global maximum is consistent with the former $\psi > 1$, our local maximum features $\psi < 1$. β is smaller than 1 in the global, and is bigger than 1 in the local. γ varies from 7 in the global to about 2 in the local.

The differences in these preference parameters do not change the model's implication for

22

Table 2: Structural parameter estimates

		global		local	
Preference	ψ	1.02		0.70	
		(0.03)		(0.04)	
	β	0.9998		1.003	
		(0.0000)		(0.000)	
	γ	6.75		1.73	
		(2.02)		(0.16)	
Habit	$\Phi_g^{\mathbb{Q}\$}$				
		0.993	0.018	0.994	-0.015
		(0.002)	(0.007)	(0.003)	(0.005)
		0.000	0.997	-0.005	0.996
		(0.001)	(0.000)	(0.002)	(0.000)
	λ_g	$1e^{-3}\times$			
		0.05	-0.12	-0.007	0.030
		0.00	0.18	0.001	-0.023

Structural parameters estimates for our benchmark model. Left: the global estimates. Right: one local maximum. Standard errors are Newey and West(1987).

bond prices. The key that is consistent across the two maxima is how persistent $\Phi_g^{\mathbb{Q}\$}$ is. It is this feature that enables both maxima to capture the feature of the data, generating upward sloping yield curve, and realistic term premia. This feature is also consistent with the findings in the Gaussian ATSM literature.

4.4 Alternative methods in the literature

A popular alternative approach adopted, for example, by Bansal and Shaliastovich(2013) among others, is to assume a subset of yields are measured without error and invert the latent macroeconomic factors as linear combinations of yields. Specifically, a researcher could assume that the pricing errors e_t in (25) are zero for N_1 yields: $y_t^{o,\$} = A^{o,\$} + B_g^{o,\$}g_t + B_h^{o,\$}h_t$. Then, the N_1 latent factors among g_t and h_t can be estimated by inverting this relationship making the estimated factors linear functions of $y_t^{o,\$}$. Note that N_1 does not have to be equal to $G+H$. For example, Bansal and Shaliastovich(2013) choose $N_1 = 3$ and estimate expected consumption growth and the two volatilities for expected components as linear combinations

of yields. By construction, the latent macroeconomic factors are completely made of yields. The factors can lose their intended interpretation as 'macroeconomic' factors.

Table 3 illustrates this point by computing the R^2's of regressing implied macroeconomic factors on yields. The first column uses our estimated macro factors described in Subsection 4.2 as the dependent variables, and yields only account for less than 50% of their variation for most variables. Put differently, macroeconomic variables and volatilities have their own rich dynamics which are not spanned by the yields. This is consistent with the ATSM literature, see Collin-Dufresne, Goldstein, and Jones(2008) and Creal and Wu(2015b). If we use the alternative procedure that inverts all the latent factors from yields, the R^2 are 100% by construction (the second column). That is because macro factors are linear combination of yields, and they look similar to the level and slope factors one would typically estimate from a reduced-form GATSM.

The loss of economic interpretation resulting from the inversion method does not change completely when we allow measurement errors on all the yields. More precisely, we use the particle Gibbs sampler as discussed in Subsection 4.2 to estimate a long run risk model with stochastic volatility but not habit when yields and macroeconomic variables are observed simultaneously. In this case, we assume that all the yields are priced with errors.[5] We report the R^2's from this model in column 3. The R^2's are still extremely high, especially, the numbers are still close to 100% for the level factors in the first two rows. This is because the estimated latent factors load predominantly on the yields as we observe an entire cross section of yields, and they display smaller idiosyncratic variance than macroeconomic data. The forecasting errors for the macroeconomic data from this model are significantly larger than our estimates, to the degree that the estimated latent factors do not resemble macroeconomic variables themselves. The R^2 for the expected growth volatility is 72% percent, much higher than its counterpart in the first column. The only exception is the volatility of expected inflation, only 36% of which is explained by yields. The explanation is intuitive: as expected

---

[5]We impose $\eta = 1$ so the model has an analytical solution. This model generalizes the model in Piazzesi and Schneider(2007) to include stochastic volatility.

Table 3: R^2 regressing macro factors on yields

	our estimates	inversion	
		w/o p.e	w/ p.e.
expected inflation	57%	100%	98%
expeted growth	31%	100%	96%
expected inflation vol	48%	100%	36%
expected growth vol	31%	100%	72%

R^2s from regressing macro factors on 6 observed yields. First column: our estimates of latent macro factors using only information from macro data as detailed in Subsection 4.2. Second column: we assume some of the yields are priced without error and the latent factors are estimated from yields by inverting the pricing equation. Third column: we estimate a long run risk model, with SV and without habit. The factors are jointly estimated from both macroecomomic variables and yields, where the latter have pricing errors. Row 1-4: expected inflation, expected consumption growth and their volatilities.

inflation is basically a linear combination of yields with 98% of its variation explained, therefore, much of its volatility can be explained by yield volatility rather than yields. The R^2 of a regression of this factor on yield volatility is 75%, when we use a simple GARCH model to estimate the regressors.

The inversion approach is not without merit. As the estimated factors are close to the level and slope factors associated with reduced-form ATSMs, it can fit the cross section of yields well as expected.[6] Therefore, from an econometric perspective, it lowers the prediction errors overall by trading off an improved fit for yields while sacrificing the fit of macroeconomic variables. However, from an economic point of view, it is questionable whether these objects should still be labeled as macroeconomic factors, as they do not resemble the true dynamics of the observed macroeconomic data.

Instead, our goal is to take the macroeconomic factors as given, and ask how much variation in asset prices the structural model can explain. To do this, we impose a strong discipline on the estimation procedure such that the macroeconomic factors are estimated using only information from macroeconomic data. This makes it harder for the model to fit the yield data. But, we view retaining the factors' intended economic interpretation as an appealing feature of our approach.

---

[6]Bansal, Gallant, and Tauchen(2007) demonstrate a similar point studying equity returns and dividends.

5 Bond term premium and habit

Bond term premia are a crucial input for central banks to implement monetary policy and the key object of interest for our paper. In this section, we empirically study whether the model proposed in Section 2 adequately captures the time-variation of term premia. Then, we decompose this time variation into the alternative channels that contribute to it. Empirically, we find the predominant channel is the time-varying habit that drives the variation in the price of risk. Specifically, the key term is the price of expected inflation risk, which loads on expected inflation itself. We then study the property of habit and its role for the consumption based models to fit the slope of the yield curve, the key moment of the term structure discussed in the literature.

5.1 Term premium and its sources

We plot the 1 year (in blue) and 5 year (in red) term premia from our model in the top left panel of Figure 2. The long term (5-year) term premium displays more variation than the medium term (1-year) term premium. The 5-year term premium was low (less than 1%) at the beginning of our sample. It increased through the 1960s and 70s, peaked in the early 1980s at about 4%, and then it trended down. It became negative during the Great Recession. This can be attributed to a flight to quality or the Fed's large-scale asset purchases. For comparison, we plot the output from a three factor reduced form Gaussian ATSM in the top right panel, which serves as a benchmark for many policy discussions (for implementation details, see for example, Hamilton and Wu(2012) and Creal and Wu(2015b)). Both the size and time variation of our estimates resemble the reduced form ATSM estimates.

With both time-varying price and quantity of risk built in, our model does an adequate job of capturing the pattern of term premia exhibited in the data. The question is then which channel contributes more? The literature provides two opposite answers: the reduced form Gaussian ATSM attributes the time-varying term premia completely to a time-varying price

Figure 2: Estimated term premia from alternative models
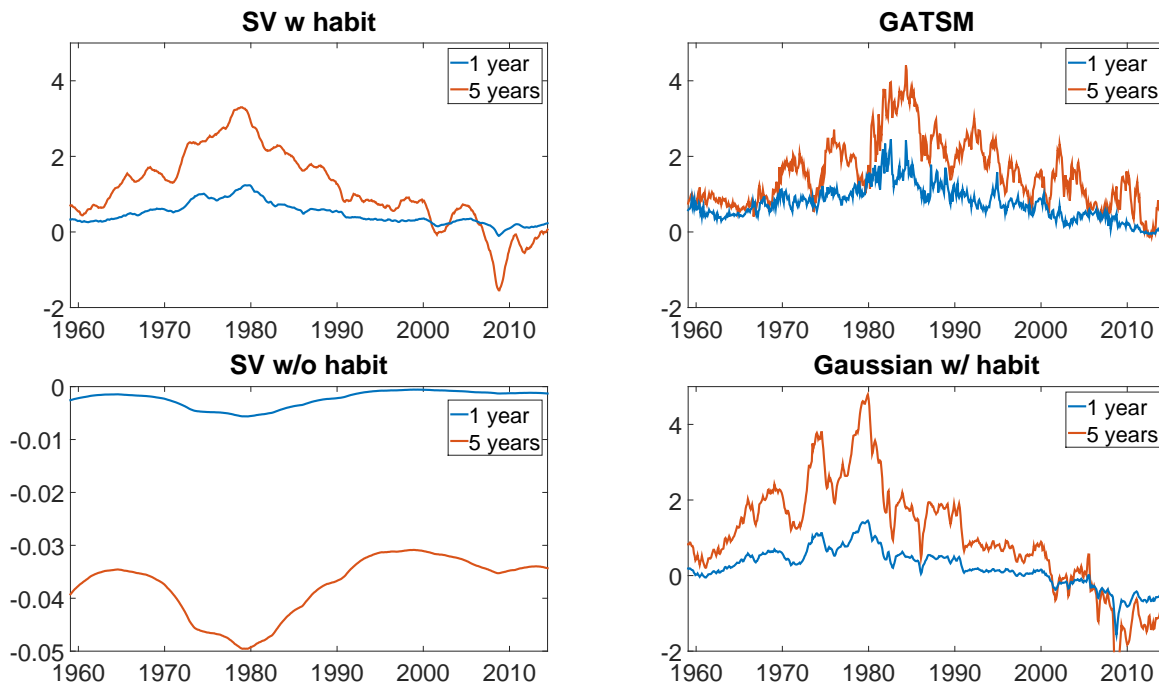
Estimated 1 and 5 year term premia from alternative models. Top left: SV model with habit; Top right: reduced-form 3 factor Gaussian ATSM. bottom left: SV model with no habit; Bottom right: Gaussian model with habit; Y-axis: interest rates measured in annualized percentage points.

of risk; while the literature on recursive preferences, especially the long run risk literature, attributes it completely to a time-varying quantity of risk. Our unifying framework equips us with a more comprehensive view to answer this question. We do so by studying how much time variation there would be if we shut down one channel at a time.

First, we shut down the price of risk channel by setting $\lambda_g = 0$, or equivalently $\Phi_g^{\mathbb{Q}^{\$}} = \Phi_g$. This model is similar to those in the long run risk literature, Bansal and Shaliastovich(2013), for example. The difference is that we model the volatility process with a non-central Gamma process guaranteeing its non-negativity, whereas the literature models it with a Gaussian process. We re-optimize the objective function subject to the constraint $\lambda_g = 0$, and plot the implied term premium in the bottom left panel of Figure 2. Without habit, although non-constant, the term premia are economically insignificant: the one year term premium is essentially zero over time, and the five year term premium peaks at about -5 basis points,

orders of magnitude smaller than the estimates in the top panels. Moreover, the term premia generated by this model are negative, which is the wrong sign.[7] All these are counter-intuitive and implausible. Hence, only time-varying quantity of risk is not sufficient to account for variation in term premium.

The conclusion seems to be at odds with the long run risk literature. As discussed in detail in Subsection 4.4, the difference is a consequence of different estimation techniques. Our approach disciplines the macroeconomic factors, such as expected inflation and consumption, to fit the observed macroeconomic data. Consequently, the estimated factors retain their economic interpretations whereas the literature obtains these macroeconomic factors as linear combinations of bond yields, potentially compromising their interpretation, see a more detailed discussion in Subsection 4.4.

Next, we shut down the time variation in the quantity of risk channel by setting $h_t = 0$ in (7) - (8), but still allow habit $\lambda_g \neq 0$. Then the factor dynamics follow a Gaussian VAR. The resulting term premia from re-optimizing this restricted model are depicted in the bottom right panel. Interestingly, both the size and time variation of the term premia resemble the estimates in the top panels. Hence, the price of risk alone generates the amount of variation of term premia as we observe from the reduced form estimates.

We have established that a time-varying price of risk through habit is a channel that can explain almost all of the variation in the bond term premia. We then further ask: is the price of inflation risk or consumption risk time varying? What drives the variation in this price? First, we only allow the price of expected inflation risk to vary over time, and also restrict it to comove with the expected inflation itself. We implement this by imposing the following restrictions on our estimates in Table 2: all components in λ_g are zero but the $\lambda_{\bar{\pi},\pi}$, or equivalently $\Phi_g^{\mathbb{Q}^\$} = \Phi_g$ for all but one component $\phi_{\bar{\pi},\pi}^{\mathbb{Q}^\$} \neq \phi_{\bar{\pi},\pi}$. This is plotted in the red line in Figure 3. As a comparison, we plot our benchmark estimate in blue, which is

---

[7]Our results are not specific to our estimates for the structural parameters. If we calibrate the structural parameters using the values from Bansal and Yaron(2004) for (β, γ, η), the model still produces the same pattern.

Figure 3: Term premia
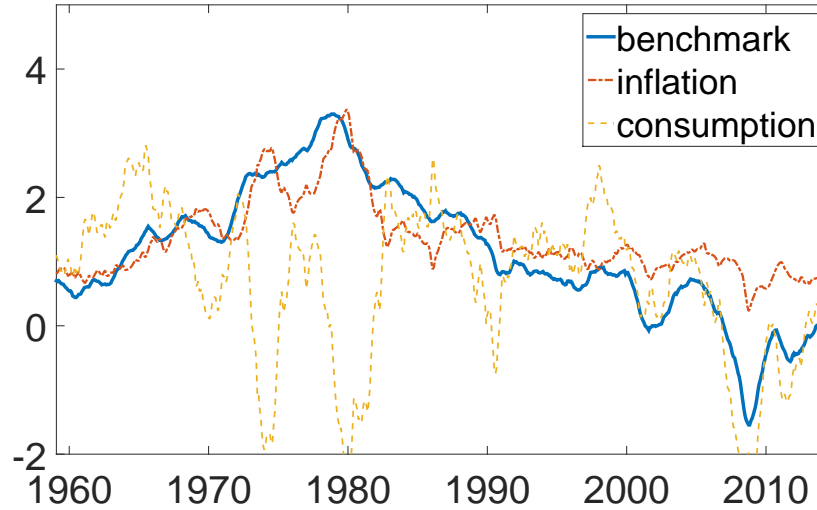
Estimated 5 year term premia. Blue: benchmark. Red: λ_g is all set to 0 but the component relating to price of expected inflation risk loading on itself. Yellow: λ_g is all set to 0 but the component relating to price of expected consumption risk loading on itself.

identical to the red line in the top left panel of Figure 2. We can see the variation in blue is primarily coming from the price of expected inflation risk loading on itself, and this one single component allows us to capture the predominant variation in the term premium. As a contrast, we plot in yellow estimates of the term premia when we only allow the price of expected consumption risk to be non-zero, and to vary with itself. Although displaying as much variation, it does not resemble the key economic feature in the term premium. For example, the term premium was lower in the 1960s, and peaked in the early 1980s in the benchmark model. The yellow line showed an opposite pattern: it was high in the middle of 1960s, and became negative during the 1970-1980s when the term premium was generally considered to be extremely high. This is counter-intuitive.

The overall picture is that the key driver for the time variation in term premium of the Treasury bond is a time-varying price of expected inflation, and the time variation comoves with the expected inflation itself.

Figure 4: Habit and habit growth
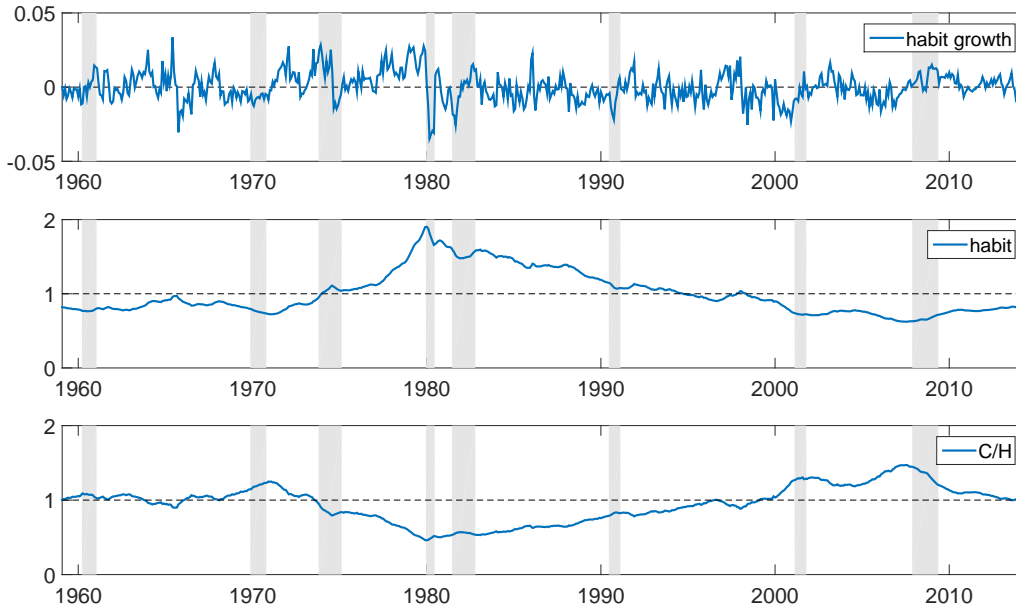
Top panel: demeaned $\Delta v_t/(\eta - 1)$, black dashed line is 0. Middle panel: H_t accumulated by the demeaned $\Delta v_t/(\eta - 1)$, and then initialized to have mean 1, black line is 1. Bottom panel: consumption to habit ratio C_t/H_t, initializing at 1. Shades: NBER recessions.

5.2 Habit

In the previous subsection, we established that the time varying price of risk through habit is the main channel that drives the dynamics of bond term premia. We now turn to study the property of habit. Figure 4 plots the historical dynamics of habit and its growth rate. Habit growth is $\ln H_{t+1} - \ln H_t = \Delta v_{t+1}/(\eta - 1)$. We plot the demeaned version in the top panel. Note, the habit growth is heteroskedastic, and the time-varying conditional variance is determined by the risk sensitivity function. The habit growth rate fluctuates around its mean. It peaked twice between 1970 and 1980. Then it dropped to its lowest point during the first recession in 1980s. Since then it was mostly negative, and only went back to persistently positive during the Great Recession.

The middle panel plots the habit H_t itself, which intuitively summarizes how agents shift their desire for consumption goods over time. We initialize it such that the mean is 1.

The habit above 1 means that agents require more consumption to achieve the same utility; when it is below 1, agents require less consumption goods. The bottom panel plots the consumption-habit ratio C_t/H_t, where we take out the deterministic trend for both series[8], and initialize the consumption and habit at the same level, hence the ratio starts at 1.

The habit started at 0.8, and it was kept below 1 until the onset of the second recession in the 1970s. For this period, a lower habit than average allows agents to be satisfied with less consumption goods. The overall effect on agents' utility reflected in the consumption-habit ratio is above 1, meaning consumption is higher than habit, and agents are happier. Then the habit kept increasing until the first recession in the early 1980s reaching a maximum of about 1.9. At that point, 90% more consumption than average is required to maintain the same satisfaction. Similarly, consumption dropped to about 46% of habit, making agents least satisfied. Since then, habit displayed a downward trend for the second half of the sample. It bottomed at 0.6 during the Great Recession. The consumption to habit ratio went the opposite direction and peaked right before the Great Recession at about 1.5, meaning consumption is 50% more than habit, making agents happier. Finally, habit leveled out at about 0.8 since, which is the same level as the beginning of the sample. At the same time, the consumption to habit ratio returned to 1.

The dynamics of habit is persistent, and it is this slow moving nature of the habit that generates enough variation in (3), and hence capture the risk premium in asset prices. This is consistent with the intuition in the literature on habit, see Campbell and Cochrane(1999) and Wachter(2006) for examples.

5.3   Slope of the yield curve

This section assesses the role habit plays in capturing a key moment of the yield curve, its unconditional slope. Table 4 shows the cross section of the yield curve, averaged over time. The top row represents the data. A well established feature of the cross section of the yield

---

[8]The deterministic trend of consumption to habit ratio is mathematically indistinguishable from β.

Table 4: Unconditional yield curves

|		3	12	24	36	48	60	average	slope
data		4.94	5.33	5.54	5.72	5.88	5.98	5.57	1.04
SV w/ habit	global	4.91	5.27	5.63	5.85	5.92	5.84	5.57	0.93
	local	4.95	5.20	5.49	5.74	5.95	6.13	5.58	1.18
Gaussian w/ habbit		5.08	5.25	5.47	5.69	5.89	6.09	5.58	1.01
SV w/o habit		5.64	5.63	5.61	5.59	5.57	5.56	5.60	-0.08

Average yields in annualized percentage points across time in the data (first row), our benchmark model with both stochastic volatility and habit (second and third row), model without stochastic volatility (fourth row), and model without habit (last row) for maturities of 3 - 60 months. Each column corresponds to one maturity. The last two columns are the average level of yields across all 6 maturities, and the slope is defined as the difference between the 60 month and 3 month yields.

curve is that it slopes upwards, and the difference between the long end and the short end (i.e. slope) is 1.04%. The second row captures our model, it mimics the first row closely, and implies a slope of 0.93%. The third row is the local maximum of our model in Table 2. It shows a similar slope of 1.18%. Row 4 allows habit but shuts down the stochastic volatility channel, and it paints the same picture as our benchmark model.

In contrast, if we do not allow habit to play a roll by imposing $\lambda_g = 0$, then the 5th row in Table 4 shows a counterfactual downward slope, and the number is -0.08%. The basic intuition is: the stochastic volatility models seem to be flexible with 4 Gaussian factors and 4 volatility factors. However, the main constraint is that there are only 3 structural parameters (β, γ, ψ) to control the cross section, and all of them mainly enter the intercept term $a_n^\$$ in (22). Without habit $\lambda_g = 0$, the autoregressive coefficient determining the time series dynamics Φ_g controls the slope of the yield curve at the same time reflected in the key term $\bar{b}_{n,g}^\$$ in (23). In the Gaussian ATSM literature, Duffee(2002) has shown that the separation between the two is important for capturing key features of the data. With habit, we are able to achieve this separation, and allow the risk neutral parameter $\Phi_g^{\mathbb{Q}^\$} \equiv \Phi_g - \eta\vartheta\lambda_g \neq \Phi_g$ to fit the cross section of the yield curve, and not be constrained by the time series dynamics.

We have demonstrated that our model fits the cross section of the yield curve well where

Figure 5: Level and slope
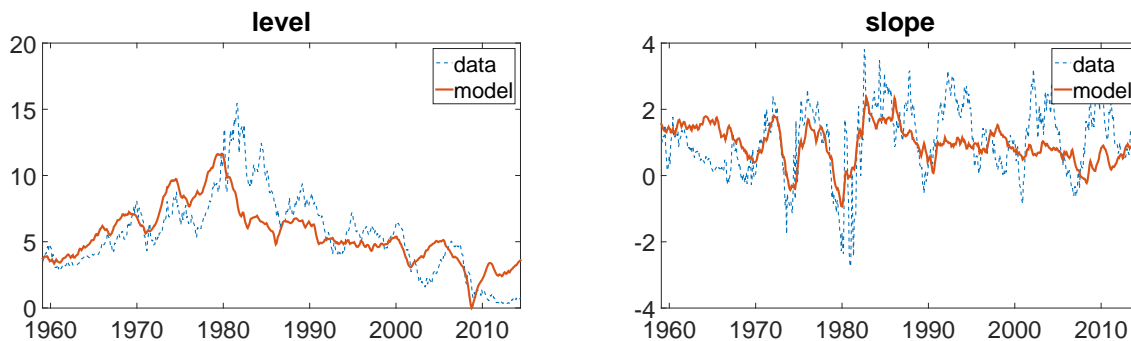
Left panel: level defined as average of yields across all maturities. Right panel: slope defined as the 5 year - 3 month yield. Blue dashed line: data. Red line: model. Y-axis: annulized percentage points

habit is a key ingredient. The more challenging task is to see if it fits the time series as well. We showed its success in the dynamics of term premia in Subsection 5.1. Although term premia have been studied elaborately in the ATSM literature, it is fundamentally an unobserved object. Now we turn to our model's ability to fit the observable moments in the data: level and slope of the yield curve.

In the left panel of figure 5, we plot the level of the yield curve over time, defined as the average of yields across all maturities in our sample. The red line is our model, and the blue dashed line is the data. In the right panel, we plot the slope of the yield curve defined as 5 year yield minus 3 month yield. Our model implied level and slope trace the data well, considering our model uses macro factors rather than latent yield factors.

6   Model properties

Empirical examination of the asset pricing implications of recursive preferences requires solving for the SDF. We base our analysis on the approximation method of Campbell and Shiller(1989), used by Bansal, Kiku, and Yaron(2012) and Schorfheide, Song, and Yaron(2014), among many others. With this approach, we need to solve for the return on the consumption asset $r_{c,t+1}$ as a function of the underlying state of the economy as in

Subsection 2.4. Whether such a solution exists mathematically amounts to a fixed point problem. In this section, we characterize the conditions that lead to a valid solution for the Euler equation and asset prices.[9]

We partition the vector of all parameters of the model $\theta = (\beta, \psi, \gamma, \theta^{\mathbb{P}}, \theta^\lambda)$ into the preference parameters (β, ψ, γ), the parameters governing the physical dynamics $\theta^{\mathbb{P}}$, and the parameters controlling habit θ^λ. We condition our analysis on $(\theta^{\mathbb{P}}, \theta^\lambda)$ and characterize the restrictions on the parameter space for the more intuitive parameters (β, ψ, γ).

6.1 General case

Laid out briefly in Subsection 2.4, the fixed point problem can be rephrased more explicitly as follows. Define the function $f(\bar{p}c, \theta)$ as

$$f(\bar{p}c; \theta) = D_0\left(\bar{p}c, \theta\right) + D_g\left(\bar{p}c, \theta\right)' \bar{g} + D_h\left(\bar{p}c, \theta\right)' \bar{h}. \tag{31}$$

For a given value of θ, a solution to the fixed point problem is obtained when $f(\bar{p}c; \theta) = \bar{p}c$. Such a solution does not always exist. Instead, the parameters must lie in a restricted space that ensures a solution.

Before we discuss the solution for the fixed point problem to exist, the parameters need to satisfy some conditions that are specific to models with stochastic volatility.[10]

Assumption 1 *The parameters $\theta \in \Theta^r$ must satisfy that for any real $\bar{p}c$,*

 1. *the loadings $D_h(\bar{p}c, \theta)$ are real,*

 2. *the expectation in (16) exists for $D_h\left(\bar{p}c, \theta\right)$.*

The first part of the assumption is used to guarantee that $f(\bar{p}c, \theta)$ is real. It amounts to a real solution for a system of H quadratic equations in H unknowns, i.e. their respective

---

[9] Hansen and Scheinkman(2012) also discuss conditions that guarantee a solution to the representative agent's problem under recursive preferences.

[10]Models with rare consumption disasters with time-varying jump intensities following a Cox, Ingersoll, and Ross(1985) process will require similar conditions.

discriminant must be positive.[11] Second, the guess and verify technique used to solve the coefficients in (17) requires the expectation in (16) to exist. This expectation does not always exist when stochastic volatility follows a multivariate Cox, Ingersoll, and Ross(1985) process. The second part of Assumption 1 guarantees the existence of the integral. These conditions are discussed in more detail in Appendix E.

Given these conditions, the following proposition provides a general condition that guarantees a solution to the representative agent's problem.

Proposition 1 *Given Assumption 1, there is a value $\bar{\beta}(\psi, \gamma, \theta^{\mathbb{P}}, \theta^{\lambda})$ such that if $\beta < \bar{\beta}$, then there exists a real solution for the fixed point problem.*

Proof: See Appendix E.1.

We use the proposition to characterize the joint restrictions that exist among all the parameters. Given the dynamics of the economy in $\theta^{\mathbb{P}}$ and the parameters determining habit in θ^{λ}, agents' risk appetite γ, and the intertemporal elasticity of substitution ψ, the representative agent needs to be sufficiently impatient (small β) in order for a solution to exist. The nature of the fixed point problem requires that all three conditions be jointly satisfied.

6.2 Special cases

In this section, we provide more intuition by discussing a special case where the dynamics are Gaussian by imposing $h_t = 0$ in (7) - (8). In this case, we are no longer constrained by Assumption 1. We can provide stronger conditions that apply to any $\beta \leq 1$, i.e., it reduces to relationships between γ and ψ. The following corollary also characterizes the upper bound $\bar{\beta}$ as a monotonic function in γ.

Corollary 1   *1. If $Z_1^{\infty\prime}\mu_g^* \leq 0$ and $\beta \leq 1$, then $\frac{1-\gamma}{1-\psi} > 0$ guarantees the existence of a solution.*

---

[11] This condition is similar to an existence condition discussed by Campbell, Giglio, Polk, and Turley(2014) in their ICAPM model. They do not provide a condition guaranteeing a solution to the fixed point problem.

2. *If $\beta \leq 1$, then there is a value $\bar{\gamma}(\theta^{\mathbb{P}}, \theta^\lambda)$ such that $\frac{\bar{\gamma} - \gamma}{1 - \psi} > 0$ guarantees a solution.*

3. *For any ψ, $\bar{\beta}$ is monotonic in γ: for $\psi > 1$, then $\frac{d\bar{\beta}}{d\gamma} > 0$; for $\psi < 1$, then $\frac{d\bar{\beta}}{d\gamma} < 0$.*

Under the condition specified in part 1 of Corollary 1, a solution exists if $(\gamma > 1, \psi > 1)$ or $(\gamma < 1, \psi < 1)$. This divides the parameter space for (γ, ψ) into four quadrants, and only two of these four have a solution. Part 2 of Corollary 1 says that $(\gamma > \bar{\gamma}, \psi > 1)$ or $(\gamma < \bar{\gamma}, \psi < 1)$ guarantees a solution, regardless of how patient the agent is. Again, two out of the four quadrants have a solution, similar to part 1. The intuition is also similar. Although the cutoff for ψ is always 1, the difference is the boundary on γ now depends on the parameters $\theta^{\mathbb{P}}$ and θ^λ.

The separation of the parameter space into quadrants makes estimation more challenging. For example, if the optimum is within the upper-right region and we start from the lower left region, a numerical optimization algorithm or a Bayesian MCMC algorithm, can have a hard time getting through the tiny bottleneck and reaching the correct part of the parameter space. In practice, we observe these algorithms hitting the regions where no solution exists and often stopping. Estimation gets more complicated when the structural parameters interact with the remaining parameters of the model as the boundaries can shift creating strong dependencies among the model's parameters.

Corollary 1 part 3 states the relationship between the upper bound for β and γ. If $\psi < 1$, then an agent cannot have a high risk aversion and be patient at the same time. The more risk averse he is, the less patient he needs to be, vice versa. If $\psi > 1$, the opposite is true.

6.3   Numerical illustrations

Figure 6 provides numerical illustrations of Proposition 1 and Corollary 1. The top row takes a special case without stochastic volatility or habit. The upper left panel provides a demonstration for part 2 of Corollary 1. $\beta = 0.9998$ is taken from the global estimates of Table 2. A similar pattern holds for other values of $\beta \leq 1$ as well. Blue dots indicate

the existence of a solution, and red stars imply no solution. The green dashed lines mark the boundaries $\psi = 1$ and $\gamma = \bar{\gamma} = 146.5$. Consistent with part 2 of Corollary 1, when γ and ψ are both bigger than their corresponding boundaries (upper right quadrant) or both smaller than the boundaries (lower left quadrant), a solution exists. The top right panel illustrates part 3 of Corollary 1. We take $\psi = 0.7$ from the local of Table 2. As prescribed by the Corollary, when $\psi < 1$, we see a downward sloping line that separates the parameter space for (β, γ) into feasible (blue) and infeasible (red) regions. The larger the value of risk aversion γ gets, the smaller the discount rate β needs to be to remain in a region with a valid solution.

While the top panels brings a visualization for Corollary 1, the bottom panels demonstrate how restrictive the space looks in the benchmark setting. In the bottom left, we take the estimates from Table 1 and the global solution of Table 2. The upper-left and lower-right regions remain infeasible as before with similar intuition as Gaussian models. The difference is now the upper-right region becomes infeasible in addition to the earlier regions in order to satisfy Assumption 1. This emphasizes that in stochastic volatility models both the intertemporal elasticity of substitution and risk aversion need to be modest. We find although the lower left region is still feasible, the region is much smaller. For $\psi = 0.97$, γ cannot exceed 4.8. For $\psi = 0.52$, γ cannot exceed 6.9. The upper bound for γ in the Gaussian case marked by the green line is 146.5.

The implications are two-fold. First, much of the economics literature evaluates a model's success according to whether or not it can produce a small value for the risk aversion parameter γ. We need to interpret this result with caution. As we show, for stochastic volatility models, a small value of γ is required to satisfy the constraints of the model. Second, stochastic volatility models have much smaller feasible regions of the parameter space, and they are more likely to encounter numerical problems and boundaries.

The bottom right panel takes the estimates from Table 1 and the local solution of Table 2. It is similar to the upper right plot. Again the downward sloping line that divides the blue

Figure 6: Feasible and infeasible regions of the parameter space.
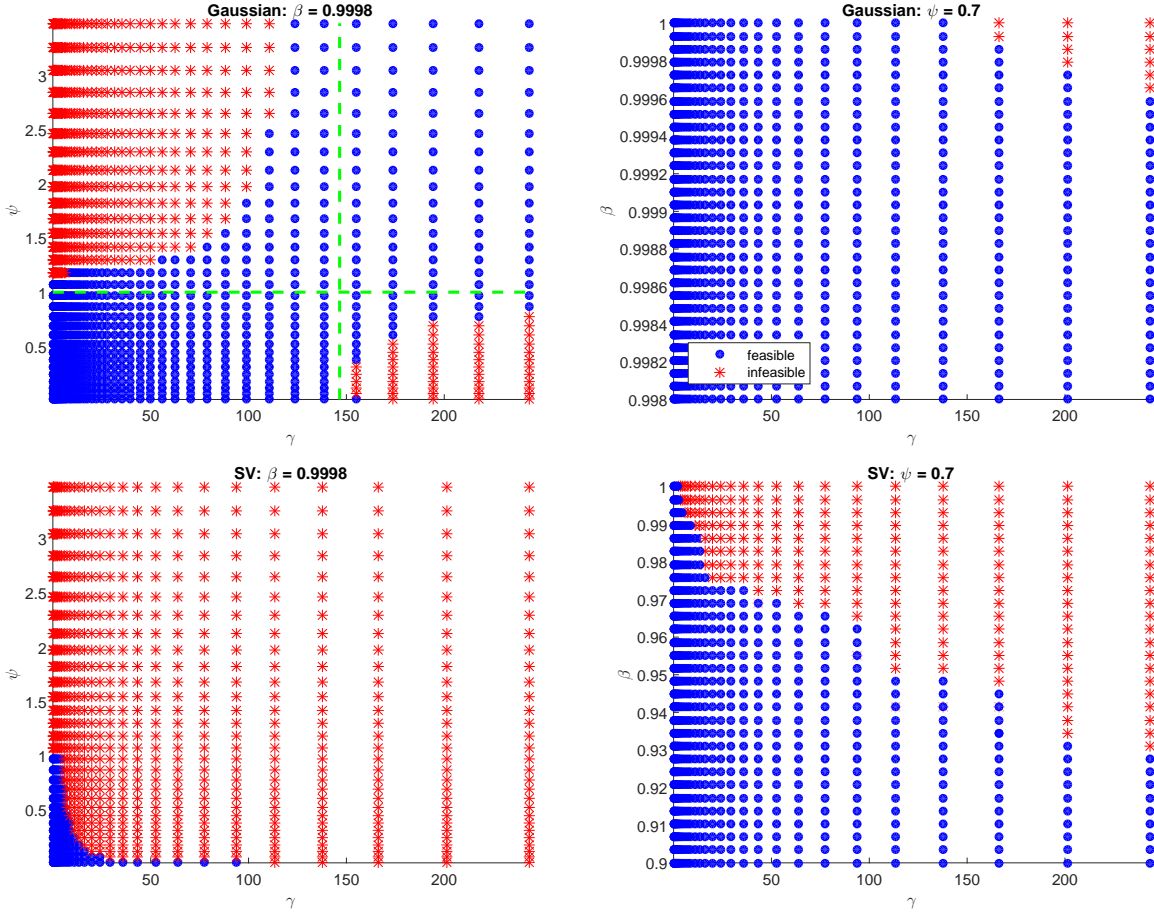
Feasible (blue dots) and infeasible (red stars) regions of the parameters space. Green dashed lines are the theoretical bounds derived in Corollary 1 part 2. The top row is a simplified model without stochastic volatility: $\theta^{\mathbb{P}}$ is taken from the estimates of this model, and $\lambda_g = 0$. The bottom row shows our benchmark model with stochastic volatility. Parameters $\theta^{\mathbb{P}}$ are taken from Table 1. θ^{λ} is taken from the global solution of Table 2 for the bottom left, and local solution for the bottom right. Left: parameter space for (γ, ψ) with $\beta = 0.9998$. Right: Parameter space for (γ, β) with $\psi = 0.7$.

and red regions indicates that with the intertemporal elasticity of substitution less than 1, an agent needs to be less patient as their risk aversion increases. The difference is that the feasible region again is much smaller. For example, for $\gamma = 244$, β can be as big as 0.9996 in the Gaussian model, but it will not be able to exceed 0.93 in the SV setting.

7 Conclusion

Two strands of related literature attribute the time variation in bond term premia to two different sources: Gaussian ATSMs credit time-varying prices of risk, whereas structural models with recursive preferences and long run risk attribute it to time-varying quantities of risk. We developed a consumption based model to capture both of these competing sources. We introduced time-varying prices of risk through an external habit that depends on current and past components of consumption growth and inflation. This generates a time-varying risk premia even when the shocks are homoskedastic. Our novel formulation of habit yields analytical bond prices, gaining tractability for this class of model. We introduce time varying quantity of risk through stochastic volatility, which follows a non-negative NCG process. We found that the time-varying price of expected inflation risk driven by expected inflation itself is the primary channel empirically. On the contrary, once habit is a component in the model, the presence of stochastic volatility does not alter the economic implication of the dynamics of term premia. Moreover, a stochastic volatility model without habit cannot match the upward sloping unconditional yield curve, the fundamental moment in the term structure. Adding habit solves this problem as well.

Empirical implementation of recursive preferences requires careful attention when solving for the stochastic discount factor. A solution does not exist for certain combinations of structural parameters. Our paper provided conditions that guaranteed the existence of a solution. We use these conditions to provide guidelines for empirical implementation.

Several authors have studied term structure models with recursive preferences in DSGE models, e.g. Rudebusch and Swanson(2008), Rudebusch and Swanson(2012), van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramírez(2012), and Dew-Becker(2014). How to introduce our technology of capturing realistic dynamics of term premia and other key aspects of bonds and other assets into a DSGE framework remains an open question, and logical next step for the literature.

References

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

Albuquerque, R., Martin Eichenbaum, and Sergio Rebelo (2014) "Valuation risk and asset pricing." Unpublished manuscript, Northwestern University, Department of Economics.

Andrieu, Christophe, Arnaud Doucet, and Roman Holenstein (2010) "Particle Markov chain Monte Carlo methods (with discussion)" *Journal of the Royal Statistical Society, Series B* 72, 1–33.

Ang, Andrew, and Monika Piazzesi (2003) "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables" *Journal of Monetary Economics* 50, 745–787.

Bansal, Ravi, A. Ronald Gallant, and George Tauchen (2007) "Rational pessimism, rational exuberance, and asset pricing models" *The Review of Economic Studies* 74, 1005–1033.

Bansal, Ravi, Dana Kiku, and Amir Yaron (2012) "An empirical evaluation of the long-run risks model for asset prices." *Critical Finance Review* 1, 1481–1509.

Bansal, Ravi, and Ivan Shaliastovich (2013) "A long-run risks explanation of predictability puzzles in bond and currency markets" *The Review of Financial Studies* 26, 1–33.

Bansal, Ravi, and Amir Yaron (2004) "Risks for the long run: a potential explanation of asset pricing puzzles" *The Journal of Finance* 59, 1481–1509.

Barro, Robert J. (2006) "Rare disasters and asset markets in the twentieth century." *Quarterly Journal of Economics* 121, 823–866.

Bauer, Michael D., Glenn D. Rudebusch, and Jing Cynthia Wu (2012) "Correcting estimation bias in dynamic term structure models." *Journal of Business and Economic*

Statistics 30, 454–467.

Bernanke, Ben S. (2006) "Reflections on the Yield Curve and Monetary Policy, March 20, 2006." Board of Governers of the Federal Reserve.

Bollerslev, Timothy, George Tauchen, and Hao Zhou (2009) "Expected stock returns and variance risk premia" *The Review of Financial Studies* 22, 4463–4492.

Caldara, Dario, Jesus Fernández-Villaverde, Juan Rubio-Ramírez, and Wen Yao (2012) "Computing DSGE models with recursive preferences and stochastic volatility." *Review of Economic Dynamics* 15, 188–206.

Campbell, John Y. (2011) "Asset Pricing I Class Notes, Consumption Based Asset Pricing." Harvard University, Department of Economics.

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, Stefano Giglio, Christopher Polk, and Bob Turley (2014) "An intertemporal CAPM with stochastic volatility." Working Paper, Booth School of Business, University of Chicago.

Campbell, John Y, and Robert J. Shiller (1989) "The dividend-price ratio and expectations of future dividends and discount factors." *The Review of Financial Studies* 1, 195–228.

Chen, Rong, and Jun S. Liu (2000) "Mixture Kalman filters." *Journal of the Royal Statistical Society, Series B* 62, 493–508.

Collin-Dufresne, Pierre, Robert S. Goldstein, and Charles Jones (2008) "Identification of maximal affine term structure models." *The Journal of Finance* 63, 743–795.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross (1985) "A theory of the term structure of interest rates" *Econometrica* 53, 385–407.

Creal, Drew D. (2012) "A survey of sequential Monte Carlo methods for economics and finance." *Econometric Reviews* 31, 245–296.

Creal, Drew D., and Ruey S. Tsay (2015) "High-dimensional dynamic stochastic copula models." *Journal of Econometrics* 189, 335–345.

Creal, Drew D., and Jing Cynthia Wu (2015a) "Bond risk premia in consumption based models." Working paper, University of Chicago, Booth School of Business.

Creal, Drew D., and Jing Cynthia Wu (2015b) "Estimation of affine term structure models with spanned or unspanned stochastic volatility." *Journal of Econometrics* 185, 60–81.

Dew-Becker, Ian (2014) "Bond pricing with a time-varying price of risk in an estimated medium-scale Bayesian DSGE model." *Journal of Money, Credit, and Banking* 46, 837888.

Duffee, Gregory R. (2002) "Term premia and interest rate forecasts in affine models" *The Journal of Finance* 57, 405–443.

Durbin, James, and Siem Jan Koopman (2002) "A simple and efficient simulation smoother for state space time series analysis." *Biometrika* 89, 603–616.

Durbin, James, and Siem Jan Koopman (2012) *Time Series Analysis by State Space Methods* Oxford University Press, Oxford, UK 2 edition.

Epstein, Lawrence, and Stanley Zin (1989) "Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework" *Econometrica* 57, 937–969.

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

Gabaix, Xavier (2012) "An exactly solved framework for ten puzzles in macro-finance." *Quarterly Journal of Economics* 127, 645–700.

Godsill, Simon J., Arnaud Doucet, and Michael West (2004) "Monte Carlo smoothing for nonlinear time series." *Journal of the American Statistical Association* 99, 156–168.

Gouriéroux, Christian, and Joann Jasiak (2006) "Autoregressive gamma processes." *Journal of Forecasting* 25, 129–152.

Greenspan, Alan (2005) "To the International Monetary Conference, Beijing, People's Republic of China (via satellite) June 6, 2005" Board of Governers of the Federal Reserve.

Hamilton, James D., and Jing Cynthia Wu (2012) "Identification and estimation of Gaussian affine term structure models." *Journal of Econometrics* 168, 315–331.

Hansen, Lars Peter, John C. Heaton, and Nan Li (2008) "Consumption strikes back? Measuring long-run risk" *Journal of Political Economy* 116, 260–302.

Hansen, Lars Peter, and José A. Scheinkman (2012) "Recursive utility in a Markov environment with stochastic growth." *Proceedings of the National Academy of Sciences* 109, 11967–72.

Kreps, David M., and Evan L. Porteus (1978) "Temporal resolution of uncertainty and dynamic choice theory." *Econometrica* 46, 185–200.

Lindsten, Fredrik, Michael I. Jordan, and Thomas B. Schön (2145–2184) "Particle Gibbs with ancestor sampling." *Journal of Machine Learning Research* 15.

Liu, Jun S., and Rong Chen (1998) "Sequential Monte Carlo computation for dynamic systems." *Journal of the American Statistical Association* 93, 1032–1044.

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

Piazzesi, Monika, and Martin Schneider (2007) "Equilibrium yield curves." in *NBER*

Macroeconomics Annual 2006, edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford MIT Press, Cambridge, MA pages 389–442.

Rudebusch, Glenn D., and Eric T. Swanson (2008) "Examining the bond premium puzzle with a DSGE model." *Journal of Monetary Economics* 55, 111–126.

Rudebusch, Glenn D., and Eric T. Swanson (2012) "The bond premium in a DSGE model with long-run real and nominal risks." *American Economic Journal: Macroeconomics* 4, 105–143.

Rudebusch, Glenn D., and Tao Wu (2008) "A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy" *The Economic Journal* 118, 906–926.

Schorfheide, Frank, Dongho Song, and Amir Yaron (2014) "Identifying long-run risks: a Bayesian mixed-frequency approach" Working paper, University of Pennsylvania, Department of Economics.

Stock, James H., and Mark W. Watson (2007) "Why has U.S. inflation become harder to forecast?" *Journal of Money, Credit, and Banking* 39, 3–34.

Tallarini, Thomas D. (2000) "Risk-sensitive real business cycles" *Journal of Monetary Economics* 45, 507–532.

van Binsbergen, Jules H., Jesús Fernández-Villaverde, Ralph S.J. Koijen, and Juan Rubio-Ramírez (2012) "The term structure of interest rates in a DSGE model with recursive preferences." *Journal of Monetary Economics* 59, 634–648.

Wachter, Jessica A. (2006) "A consumption-based model of the term structure of interest rates." *Journal of Financial Economics* 79, 365–399.

Wachter, Jessica A. (2013) "Can time-varying risk of rare disasters explain aggregate stock market volatility?" *The Journal of Finance* 68, 987–1035.

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

Whiteley, Nick (2010) "Discussion on Particle Markov chain Monte Carlo Methods." *Journal of the Royal Statistical Society, Series B* 72, 306–307.

Wright, Jonathan H. (2011) "Term premia and inflation uncertainty: empirical evidence from an international panel dataset." *American Economic Review* 101(4), 1514–1534.

Yellen, Janet (2014) "Transcript of Chair Yellens FOMC Press Conference, December 17, 2014." Board of Governers of the Federal Reserve.

Appendix A   Stochastic discount factor

This appendix provides the derivation for the stochastic discount factor for the agent's problem in (1) - (2). Guess that the solution is $V_t = \phi_t W_t$ for some coefficients ϕ_t, then the agent's problem becomes

$$\phi_t W_t = \max_{C_t} \left[(1-\beta) H_t^{\eta-1} C_t^{1-\eta} + \beta \left\{ \mathrm{E}_t \left[(\phi_{t+1} W_{t+1})^{1-\gamma} \right] \right\}^{\frac{1-\eta}{1-\gamma}} \right]^{\frac{1}{1-\eta}}.$$

Substitute in W_{t+1} from the constraint (2)

$$\phi_t^{1-\eta} = \max_{C_t} \left[(1-\beta) H_t^{\eta-1} \left(\frac{C_t}{W_t} \right)^{1-\eta} + \beta \left(1 - \frac{C_t}{W_t} \right)^{1-\eta} \left\{ \mathrm{E}_t \left[(\phi_{t+1} R_{c,t+1})^{1-\gamma} \right] \right\}^{\frac{1-\eta}{1-\gamma}} \right]. \tag{A.1}$$

Take the first-order condition w.r.t. C_t, we get

$$(1-\beta) H_t^{\eta-1} \left(\frac{C_t}{W_t} \right)^{-\eta} = \beta \left(1 - \frac{C_t}{W_t} \right)^{-\eta} \left\{ \mathrm{E}_t \left[(\phi_{t+1} R_{c,t+1})^{1-\gamma} \right] \right\}^{\frac{1-\eta}{1-\gamma}}. \tag{A.2}$$

Use the first order condition to substitute out the expectation term in (A.1) to solve ϕ_t,

$$\phi_t = (1-\beta)^{\frac{1}{1-\eta}} H_t^{-1} \left(\frac{C_t}{W_t} \right)^{-\frac{\eta}{1-\eta}}.$$

Substitute back to the FOC in (A.2), and and use the budget constraint to get the pricing equation

$$1 = \beta^{\vartheta} \mathrm{E}_t \left[\left(\frac{H_{t+1}^{\eta-1}}{H_t^{\eta-1}} \right)^{\vartheta} \left(\frac{C_{t+1}}{C_t} \right)^{-\eta\vartheta} R_{c,t+1}^{\vartheta} \right].$$

Therefore, the pricing kernel is

$$M_{t+1} = \beta^{\vartheta} \left(\frac{H_{t+1}^{\eta-1}}{H_t^{\eta-1}} \right)^{\vartheta} \left(\frac{C_{t+1}}{C_t} \right)^{-\eta\vartheta} R_{c,t+1}^{\vartheta-1},$$

and the log SDF is shown in (3).

Appendix B   Dynamics of the state vector

Appendix B.1   General Model

The dynamics of the Gaussian state vector g_t driving Δc_t and π_t are

$$g_{t+1} = \mu_g + \Phi_g g_t + \Phi_{gh} h_t + \Sigma_{gh} \varepsilon_{h,t+1} + \Sigma_{g,t} \varepsilon_{g,t+1}, \qquad \varepsilon_{g,t+1} \sim \mathrm{N}(0, I),$$

$$\Sigma_{g,t} \Sigma_{g,t}' = \Sigma_{0,g} \Sigma_{0,g}' + \sum_{i=1}^{H} \Sigma_{i,g} \Sigma_{i,g}' h_{it},$$

$$\varepsilon_{h,t+1} = h_{t+1} - \mathrm{E}_t[h_{t+1}|h_t],$$

where the volatility dynamics are a non-central gamma process. They can be written as a Gamma distribution and a Poisson distribution

$$h_{t+1} = \Sigma_h w_{t+1},$$
$$w_{i,t+1} \sim \mathrm{Gamma}(\nu_{h,i} + z_{i,t+1}, 1), \qquad i = 1, \ldots, H \tag{B.3}$$
$$z_{i,t+1} \sim \mathrm{Poisson}(e_i' \Sigma_h^{-1} \Phi_h \Sigma_h w_t), \qquad i = 1, \ldots, H. \tag{B.4}$$

This is a discrete-time, multivariate Cox, Ingersoll, and Ross(1985) process. To guarantee positivity and existence of h_t, the process requires $\Sigma_h > 0$, $\Sigma_h^{-1}\Phi_h\Sigma_h > 0$ and the Feller condition $\nu_{h,i} > 1$ for $i = 1,\ldots,H$. The conditional mean and variance of the process are

$$
\begin{aligned}
\mathrm{E}_t\left[h_{t+1}|h_t\right] &= \Sigma_h\nu_h + \Phi_h h_t, & \text{(B.5)}\\
V_t\left[h_{t+1}|h_t\right] &= \Sigma_{h,t}\Sigma'_{h,t}\\
&= \Sigma_h\mathrm{diag}\left(\nu_h\right)\Sigma'_h + \Sigma_h\mathrm{diag}\left(2\Sigma_h^{-1}\Phi_h h_t\right)\Sigma'_h, & \text{(B.6)}
\end{aligned}
$$

where Σ_h is a $H\times H$ matrix of scale parameters, Φ_h is a $H\times H$ matrix of autoregressive parameters and the intercept is equal to $\Sigma_h\nu_h$. The unconditional mean and variance of h_t are $\bar{\mu}_h = \left(I_H - \Phi_h\right)^{-1}\Sigma_h\nu_h$ and $\bar{\Sigma}_h\bar{\Sigma}'_h = \left(I_H - \Phi_h\right)^{-1}\Sigma_h\mathrm{diag}\left(\nu_h\right)\Sigma'_h\left(I_H - \Phi_h\right)^{-1,'}$. The unconditional mean of g_t is $\bar{\mu}_g = \left(I_G - \Phi_g\right)^{-1}\left(\mu_g + \Phi_{gh}\bar{\mu}_h\right)$. The transition density of h_t is

$$
\begin{aligned}
p\left(h_{t+1}|h_t,\nu_h,\Phi_h,\Sigma_h\right) &= |\Sigma_h^{-1}|\prod_{i=1}^{H}\left(\mathrm{e}'_i\Sigma_h^{-1}h_{t+1}\right)^{\frac{\nu_{h,i}-1}{2}}\left(\mathrm{e}'_i\Sigma_h^{-1}\Phi_h h_t\right)^{-\frac{\nu_{h,i}-1}{2}}\\
&\quad \exp\left(-\sum_{i=1}^{H}\mathrm{e}'_i\Sigma_h^{-1}h_{t+1} + \mathrm{e}'_i\Sigma_h^{-1}\Phi_h h_t\right)\\
&\quad I_{\nu_{h,i}-1}\left(2\sqrt{\left(\mathrm{e}'_i\Sigma_h^{-1}h_{t+1}\right)\left(\mathrm{e}'_i\Sigma_h^{-1}\Phi_h h_t\right)}\right) & \text{(B.7)}
\end{aligned}
$$

where $I_\nu\left(x\right)$ is the modified Bessel function. The Laplace transform needed to solve the model with recursive preferences and for pricing assets is

$$
\mathbb{E}_t\left[\exp\left(u'h_{t+1}\right)\right] = \exp\left(\sum_{i=1}^{H}\frac{\mathrm{e}'_i\Sigma'_h u}{1 - \mathrm{e}'_i\Sigma'_h u}\mathrm{e}'_i\Sigma_h^{-1}\Phi_h h_t - \sum_{i=1}^{H}\nu_{h,i}\log\left(1 - \mathrm{e}'_i\Sigma'_h u\right)\right),
$$

which exists only if $\mathrm{e}'_i\Sigma'_h u < 1$ for $i = 1,\ldots,H$. Further properties of the univariate process are developed by Gouriéroux and Jasiak(2006).

Appendix B.2 Long run risk

Non-Gaussian model The model with long-run risk to consumption growth and trend inflation maps into the general form as follows

$$
g_t = \begin{pmatrix}\pi_t\\\Delta c_t\\\bar{\pi}_t\\\bar{c}_t\end{pmatrix}\quad Z_c = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}\quad Z_\pi = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}\quad \Phi_g = \begin{pmatrix}0 & 0 & 1 & 0\\0 & 0 & 0 & 1\\0 & 0 & \phi_\pi & \phi_{\pi,c}\\0 & 0 & \phi_{c,\pi} & \phi_c\end{pmatrix}\quad \mu_g = \begin{pmatrix}0\\0\\\mu_\pi\\\mu_c\end{pmatrix}\quad \bar{\mu}_g = \begin{pmatrix}\bar{\mu}_\pi\\\bar{\mu}_c\\\bar{\mu}_\pi\\\bar{\mu}_c\end{pmatrix}
$$

$$
\Phi_{gh} = \begin{pmatrix}0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{pmatrix}\quad \Sigma_{gh} = \begin{pmatrix}0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{pmatrix}\quad \Sigma_{0,g} = \begin{pmatrix}0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{pmatrix}\quad \Sigma_{1,g} = \begin{pmatrix}\frac{1}{\sqrt{12000}} & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{pmatrix}
$$

$$
\Sigma_{2,g} = \begin{pmatrix}0 & 0 & 0 & 0\\0 & \frac{1}{\sqrt{12000}} & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{pmatrix}\quad \Sigma_{3,g} = \begin{pmatrix}0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & \frac{1}{\sqrt{12000}} & 0\\0 & 0 & \frac{\sigma_{c,\pi}}{\sqrt{12000}} & 0\end{pmatrix}\quad \Sigma_{4,g} = \begin{pmatrix}0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & \frac{1}{\sqrt{12000}}\end{pmatrix}
$$

We have scaled these matrices by $1/\sqrt{12000}$ so that the volatility factors h_t are roughly the same magnitude as the Gaussian factors g_t. For the volatility processes, the matrices are

$$
\bar{\mu}_h = \begin{pmatrix} \bar{\mu}_{h,\pi_1} \\ \bar{\mu}_{h,c_1} \\ \bar{\mu}_{h,\pi_2} \\ \bar{\mu}_{h,c_2} \end{pmatrix} \quad
\nu_h = \begin{pmatrix} \nu_{h,\pi_1} \\ \nu_{h,c_1} \\ \nu_{h,\pi_2} \\ \nu_{h,c_2} \end{pmatrix} \quad
\Phi_h = \begin{pmatrix} \phi_{\pi_1} & 0 & 0 & 0 \\ 0 & \phi_{c_1} & 0 & 0 \\ 0 & 0 & \phi_{\pi_2} & 0 \\ 0 & 0 & 0 & \phi_{c_2} \end{pmatrix} \quad
\Sigma_h = \begin{pmatrix} \sigma_{h,\pi_1} & 0 & 0 & 0 \\ 0 & \sigma_{h,c_1} & 0 & 0 \\ 0 & 0 & \sigma_{h,\pi_2} & 0 \\ 0 & 0 & 0 & \sigma_{h,c_2} \end{pmatrix}
$$

During estimation, we parameterize the model in terms of the unconditional mean of volatilities $\bar{\mu}_h$.

Gaussian model For the Gaussian model, we keep everything the same as above except for the scale matrix which is equal to

$$
\Sigma_{0,g} = \begin{pmatrix} \sigma_{\pi_1} & 0 & 0 & 0 \\ 0 & \sigma_{c_1} & 0 & 0 \\ 0 & 0 & \sigma_{\pi_2} & 0 \\ 0 & 0 & \sigma_{c,\pi} & \sigma_{c_2} \end{pmatrix},
$$

while $\Sigma_{i,g} = 0$ for $i > 0$, and $\bar{\mu}_h, \nu_h = 0, \Phi_h = 0, \Sigma_h = 0$.

Appendix C   Recursive preferences model solution

Appendix C.1   Solution for $r_{c,t+1}$

In order to simplify the expressions, we introduce the following notation

$$
\begin{aligned}
Z_1 &= (1-\eta) Z_c + \kappa_1 D_g \\
Z_2 &= -\gamma Z_c + (\vartheta - 1) \kappa_1 D_g \\
Z_3 &= \Sigma'_{gh} ((1-\eta) Z_c + \kappa_1 D_g) + \kappa_1 D_h \\
&= \Sigma'_{gh} Z_1 + \kappa_1 D_h \\
Z_4 &= \Sigma'_{gh} (-\gamma Z_c + (\vartheta - 1) \kappa_1 D_g) + (\vartheta - 1) \kappa_1 D_h \\
&= \Sigma'_{gh} Z_2 + (\vartheta - 1) \kappa_1 D_h
\end{aligned}
$$

where the vectors Z_c, Z_π are selection vectors and the vectors D_g and D_h are part of the price to consumption ratio $pc_t = D_0 + D'_g g_t + D'_h h_t$.

Step 1: Campbell-Shiller approximation

Let $pc_t = \ln\left(\frac{P_t}{C_t}\right)$ be the log price to consumption ratio. The return on the consumption asset is

$$
\begin{aligned}
r_{c,t+1} &\equiv \ln\left(\frac{P_{t+1} + C_{t+1}}{P_t}\right) = \ln(C_{t+1}) + \ln\left(\frac{P_{t+1} + C_{t+1}}{C_{t+1}}\right) - \ln(P_t) \\
&= \ln(C_{t+1}) - \ln(C_t) + \ln\left(1 + \frac{P_{t+1}}{C_{t+1}}\right) - \ln(P_t) + \ln(C_t) = \Delta c_{t+1} - pc_t + \ln(1 + \exp(pc_{t+1})).
\end{aligned}
$$

Take a first order Taylor expansion of the function $f(x) = \ln(1 + \exp(x))$ around $\bar{x}$.

$$
\begin{aligned}
r_{c,t+1} &\approx \Delta c_{t+1} - pc_t + \ln(1 + \exp(\bar{pc})) + \frac{\exp(\bar{pc})}{1 + \exp(\bar{pc})}(pc_{t+1} - \bar{pc}) \\
&= \kappa_0 + \kappa_1 pc_{t+1} - pc_t + \Delta c_{t+1}
\end{aligned} \tag{C.8}
$$

where $\kappa_0 = \ln(1 + \exp(\bar{pc})) - \kappa_1 \bar{pc}$ and $\kappa_1 = \frac{\exp(\bar{pc})}{1 + \exp(\bar{pc})}$.

Step 2: Solve for the price/consumption ratio

The real pricing kernel in (3) prices the consumption asset.

$$\begin{aligned}
1 &= \mathrm{E}_t\left[\exp\left(m_{t+1} + r_{c,t+1}\right)\right] = \mathrm{E}_t\left[\exp\left(\vartheta\ln\left(\beta\right) + \vartheta\Delta v_{t+1} - \eta\vartheta\Delta c_{t+1} + \vartheta r_{c,t+1}\right)\right] \\
&= \exp\left(\vartheta\ln\left(\beta\right) + \vartheta\kappa_0 - \vartheta pc_t\right)\mathrm{E}_t\left[\exp\left(\vartheta\Delta v_{t+1} + \vartheta\left(1-\eta\right)\Delta c_{t+1} + \vartheta\kappa_1 pc_{t+1}\right)\right] \quad\text{(C.9)}
\end{aligned}$$

where we have used (C.8). Conjecture a solution for the price to consumption ratio

$$pc_t = D_0 + D_g' g_t + D_h' h_t \quad\text{(C.10)}$$

for unknown coefficients D_0, D_g and D_h. Substitute the guess, (5) and (13) into (C.9)

$$\begin{aligned}
1 &= \exp\left(\vartheta\ln\left(\beta\right) + \vartheta\kappa_0 + \vartheta\kappa_1 D_0 - \vartheta pc_t + \vartheta\Lambda_1\left(g_t\right)\right) & \text{(C.11)} \\
&\quad \exp\left(\vartheta Z_1'\left(\mu_g + \Phi_g g_t + \Phi_{gh} h_t - \Sigma_{gh}\left(\Sigma_h \nu_h + \Phi_h h_t\right)\right)\right) & \text{(C.12)} \\
&\quad \mathrm{E}_t\left[\exp\left(\left(\vartheta\Lambda_2\left(g_t\right) + \vartheta\Sigma_{g,t}' Z_1\right)'\varepsilon_{g,t+1}\right)\right]\mathrm{E}_t\left[\exp\left(\vartheta Z_3' h_{t+1}\right)\right] & \text{(C.13)}
\end{aligned}$$

Calculate the expectations using the Laplace transform

$$\begin{aligned}
0 &= \vartheta\ln\left(\beta\right) + \vartheta\kappa_0 + \vartheta\kappa_1 D_0 - \vartheta pc_t + \vartheta\Lambda_1\left(g_t\right) \\
&\quad + \vartheta Z_1'\left(\mu_g + \Phi_g g_t + \Phi_{gh} h_t - \Sigma_{gh}\left(\Sigma_h \nu_h + \Phi_h h_t\right)\right) + \frac{\vartheta^2}{2}\left(\Lambda_2\left(g_t\right) + \Sigma_{g,t}' Z_1\right)'\left(\Lambda_2\left(g_t\right) + \Sigma_{g,t}' Z_1\right) \\
&\quad - \sum_{i=1}^{H}\nu_{h,i}\ln\left(1 - e_i'\Sigma_h'\vartheta Z_3\right) + \sum_{i=1}^{H}\frac{e_i'\Sigma_h'\vartheta Z_3}{1 - e_i'\Sigma_h'\vartheta Z_3}e_i'\Sigma_h^{-1}\Phi_h h_t
\end{aligned}$$

The solution exists if $e_i'\Sigma_h'\vartheta Z_3 < 1$ for $i = 1, \ldots, H$. Solve for pc_t by plugging in the risk sensitivity functions and cancel terms

$$\begin{aligned}
pc_t &= \ln\left(\beta\right) + \kappa_0 + \kappa_1 D_0 \\
&\quad + Z_1'\left(\mu_g + \Phi_g g_t + \Phi_{gh} h_t - \Sigma_{gh}\left(\Sigma_h \nu_h + \Phi_h h_t\right)\right) \\
&\quad - \eta\vartheta Z_1'\left(\lambda_0 + \lambda_g g_t\right) \\
&\quad + \frac{\vartheta}{2}Z_1'\Sigma_{g,t}\Sigma_{g,t}' Z_1 - \frac{1}{\vartheta}\sum_{i=1}^{H}\nu_{h,i}\log\left(1 - e_i'\Sigma_h'\vartheta Z_3\right) + \sum_{i=1}^{H}\frac{e_i'\Sigma_h' Z_3}{1 - \vartheta e_i'\Sigma_h' Z_3}e_i'\Sigma_h^{-1}\Phi_h h_t
\end{aligned}$$

We now solve for the coefficients. Both D_0 and D_g are analytical

$$\begin{aligned}
D_0 &= \frac{1}{\left(1 - \kappa_1\right)}\left[\ln\left(\beta\right) + \kappa_0 + Z_1'\left(\mu_g - \Sigma_{gh}\Sigma_h\nu_h - \eta\vartheta\lambda_0\right)\right. \\
&\qquad\qquad\left. - \frac{1}{\vartheta}\sum_{i=1}^{H}\nu_{h,i}\ln\left(1 - e_i'\Sigma_h'\vartheta Z_3\right) + \frac{\vartheta}{2}Z_1'\Sigma_{0,g}\Sigma_{0,g}' Z_1\right] \\
D_g &= \left(I_G - \kappa_1\left(\Phi_g - \eta\vartheta\lambda_g\right)'\right)^{-1}\left(\Phi_g - \eta\vartheta\lambda_g\right)'\left(1 - \eta\right)Z_c
\end{aligned}$$

A solution for D_g only exists when $\left(I_G - \kappa_1\left(\Phi_g - \eta\vartheta\lambda_g\right)\right)$ is invertible. The vector D_h is the solution to the system of equations

$$D_h = \left(\Phi_{gh} - \Sigma_{gh}\Phi_h\right)' Z_1 + \frac{\vartheta}{2}\left(\iota_H \otimes Z_1\right)'\tilde{\Sigma}_g\tilde{\Sigma}_g'\left(I_H \otimes Z_1\right) + \sum_{i=1}^{H}\frac{e_i'\Sigma_h' Z_3}{1 - \vartheta e_i'\Sigma_h' Z_3}\Phi_h'\Sigma_h^{-1,'}e_i \quad\text{(C.14)}$$

where $\tilde{\Sigma}_g\tilde{\Sigma}_g'$ is a $GH \times GH$ block diagonal matrix with $\Sigma_{i,g}\Sigma_{i,g}'$ along the diagonal. This cannot be solved in closed-form in the general case. However, if Σ_h and Φ_h are lower triangular, then it can be calculated in closed-form recursively for $i = 1, \ldots, H$. We discuss the analytical solution of this equation in more detail

Step 3: Solve for the fixed-point

During estimation, we determine the value of $\bar{p}c$ and the log-linearization constants κ_0 and κ_1 as a function of the model parameters by solving the fixed-point problem (averaging of (C.10))

$$0 = \bar{p}c - D_0(\bar{p}c) - D_g(\bar{p}c)'\bar{\mu}_g - D_h(\bar{p}c)'\bar{\mu}_h$$

where the coefficients D_0, D_g and D_h are functions of $\bar{p}c$ through κ_0 and κ_1. The parameters $\bar{\mu}_g$ and $\bar{\mu}_h$ are the unconditional means of g_t and h_t.

Step 4: Substitute the solution into the SDF

If the fixed point problem has a solution, then the return on the consumption asset is

$$r_{c,t+1} \approx \kappa_0 + \kappa_1(D_0 + D_g'g_{t+1} + D_h'h_{t+1}) - (D_0 + D_g'g_t + D_h'h_t) + \Delta c_{t+1}$$

by substituting (C.10) into (C.8). We can now write the log-SDF as a function of the r.v.'s $\varepsilon_{g,t+1}$ and h_{t+1} by substituting this, (5) and (13) into (3)

$$
\begin{aligned}
m_{t+1} = {} & \vartheta \ln(\beta) + (\vartheta - 1)(\kappa_0 - (1-\kappa_1)D_0) \\
& - (\vartheta - 1)D_g'g_t - (\vartheta - 1)D_h'h_t + \vartheta\Lambda_1(g_t) \\
& + Z_2'(\mu_g + \Phi_g g_t + \Phi_{gh}h_t - \Sigma_{gh}(\Sigma_h\nu_h + \Phi_h h_t)) + (\vartheta\Lambda_2(g_t) + \Sigma_{g,t}'Z_2)'\varepsilon_{g,t+1} + Z_4'h_{t+1}
\end{aligned}
$$

Appendix C.2   Analytical solution of D_h

The $H \times 1$ vector of loadings D_h are a system of H equations in H unknowns in (C.14). They can be solved analytically when both Φ_h and Σ_h are lower triangular by recursively solving one equation after another. We will consider the simpler case when they are both diagonal. Under this assumption, each equation is independent of one another and they simplify to

$$D_{h,i} = \bar{D}_i + \frac{(\bar{Z}_{3,i} + \kappa_1 D_{h,i})\Phi_{h,i}}{1 - \vartheta\Sigma_{h,i}(\bar{Z}_{3,i} + \kappa_1 D_{h,i})} \qquad\qquad i = 1, \ldots, H \qquad\qquad \text{(C.15)}$$

where $\Phi_{h,i}$ and $\Sigma_{h,i}$ are the i-th diagonal elements and $\bar{D}_i, \bar{Z}_{3,i}$ are the ith elements of the following quantities

$$
\begin{aligned}
\bar{D} &= (\Phi_{gh} - \Sigma_{gh}\Phi_h)'Z_1 + \frac{\vartheta}{2}(\iota_H \otimes Z_1)'\tilde{\Sigma}_g\tilde{\Sigma}_g'(I_H \otimes Z_1) \\
\bar{Z}_3 &= \Sigma_{gh}'Z_1
\end{aligned}
$$

Each loading (C.15) for $i = 1, \ldots, H$ is a quadratic equation

$$0 = \kappa_1\vartheta\Sigma_{h,i}D_{h,i}^2 + D_{h,i}(\kappa_1\Phi_{h,i} - \kappa_1\vartheta\Sigma_{h,i}\bar{D}_i - 1 + \vartheta\Sigma_{h,i}\bar{Z}_{3,i}) + \bar{D}_i(1 - \vartheta\Sigma_{h,i}\bar{Z}_{3,i}) + \bar{Z}_{3,i}\Phi_{h,i} \text{(C.16)}$$

The solutions are

$$
\begin{aligned}
D_{h,i} = {} & \frac{-(\kappa_1\Phi_{h,i} - \kappa_1\vartheta\Sigma_{h,i}\bar{D}_i - 1 + \vartheta\Sigma_{h,i}\bar{Z}_{3,i})}{2\kappa_1\vartheta\Sigma_{h,i}} \\
& \pm \frac{\sqrt{(\kappa_1\Phi_{h,i} - \kappa_1\vartheta\Sigma_{h,i}\bar{D}_i - 1 + \vartheta\Sigma_{h,i}\bar{Z}_{3,i})^2 - 4\kappa_1\vartheta\Sigma_{h,i}[\bar{D}_i(1 - \vartheta\Sigma_{h,i}\bar{Z}_{3,i}) + \bar{Z}_{3,i}\Phi_{h,i}]}}{2\kappa_1\vartheta\Sigma_{h,i}} \text{(C.17)}
\end{aligned}
$$

A real solution exists as long as the discriminant is greater than or equal to zero. If the discriminant is greater than zero, there are two solutions. Only one solution leads to a sensible value. This is the value with

a negative sign, see also Campbell, Giglio, Polk, and Turley(2014) for the ICAPM model.

Appendix D Bond prices

Define

$$Z_5 \;=\; Z_4 - \Sigma'_{gh} Z_\pi$$

in addition to Z_1 - Z_4 defined in Appendix C.1.

Appendix D.1 Real bonds

We will guess and verify that the solution for zero coupon bonds is $P_t^{(n)} \;=\; \exp\left(\bar{a}_n + \bar{b}'_{n,g} g_t + \bar{b}'_{n,h} h_t\right)$ for some unknown coefficients $\bar{a}_n$ and $\bar{b}_{n,g}$ and $\bar{b}_{n,h}$.

For a maturity $n = 1$, the payoff is guaranteed to be $P_{t+1}^{(0)} = 1$ in the next period, in which case $P_t^{(1)} = \mathrm{E}_t\left[M_{t+1}\right]$. Using standard techniques for affine bond pricing in discrete-time (see Creal and Wu(2015b)), we find that at maturity $n = 1$ the bond loadings are

$$
\begin{aligned}
\bar{a}_1 &= \ln(\beta) - \eta Z'_c \left(\mu_g - \Sigma_{gh}\Sigma_h\nu_h - \eta\vartheta\lambda_0\right) \\
&\quad - \sum_{i=1}^{H} \nu_{h,i} \log\left(1 - \mathrm{e}'_i\Sigma'_h Z_4\right) + \frac{(\vartheta - 1)}{\vartheta} \sum_{i=1}^{H} \nu_{h,i} \log\left(1 - \mathrm{e}'_i\Sigma'_h\vartheta Z_3\right) \\
&\quad - \frac{(\vartheta - 1)\vartheta}{2} Z'_1\Sigma_{0,g}\Sigma'_{0,g}Z_1 + \frac{1}{2} Z'_2\Sigma_{0,g}\Sigma'_{0,g}Z_2 \\
\bar{b}_{1,g} &= -\left(\Phi_g - \eta\vartheta\lambda_g\right)' \eta Z_c \\
\bar{b}_{1,h} &= -\left(\Phi_{gh} - \Sigma_{gh}\Phi_h\right)' \eta Z_c \\
&\quad + \left(\sum_{i=1}^{H} \frac{\mathrm{e}'_i\Sigma'_h Z_4}{1 - \mathrm{e}'_i\Sigma'_h Z_4} \mathrm{e}'_i\Sigma_h^{-1}\Phi_h\right)' - (\vartheta - 1)\left(\sum_{i=1}^{H} \frac{\mathrm{e}'_i\Sigma'_h Z_3}{1 - \vartheta\mathrm{e}'_i\Sigma'_h Z_3} \mathrm{e}'_i\Sigma_h^{-1}\Phi_h\right)' \\
&\quad + \frac{1}{2}\left(I_H \otimes Z_2\right)' \tilde{\Sigma}_g\tilde{\Sigma}'_g\left(\iota_H \otimes Z_2\right) - \frac{(\vartheta - 1)\vartheta}{2}\left(I_H \otimes Z_1\right)' \tilde{\Sigma}_g\tilde{\Sigma}'_g\left(\iota_H \otimes Z_1\right)
\end{aligned}
$$

where bond prices only exist if $\mathrm{e}'_i\Sigma'_h Z_4 < 1$ for $i = 1, \ldots, H$. At maturity n, we use the fact that $P_t^{(n)} = \mathrm{E}_t\left[\exp\left(m_{t+1}\right) P_{t+1}^{(n-1)}\right]$. The bond loadings are

$$
\begin{aligned}
\bar{a}_n &= \bar{a}_{n-1} + \bar{a}_1 + \sum_{i=1}^{H} \nu_{h,i} \log\left(\frac{1 - \mathrm{e}'_i\Sigma'_h Z_4}{1 - \mathrm{e}'_i\Sigma'_h\left(\Sigma'_{gh}\bar{b}_{n-1,g} + \bar{b}_{n-1,h} + Z_4\right)}\right) \\
&\quad + \left(\mu_g - \Sigma_{gh}\Sigma_h\nu_h - \eta\vartheta\lambda_0\right)' \bar{b}_{n-1,g} + \frac{1}{2}\bar{b}'_{n-1,g}\Sigma_{0,g}\Sigma'_{0,g}\bar{b}_{n-1,g} + \bar{b}'_{n-1,g}\Sigma_{0,g}\Sigma'_{0,g}Z_2 \\
\bar{b}_{n,g} &= \left(\Phi_g - \eta\vartheta\lambda_g\right)' \bar{b}_{n-1,g} + \bar{b}_{1,g} \\
\bar{b}_{n,h} &= \left(\Phi_{gh} - \Sigma_{gh}\Phi_h\right)' \bar{b}_{n-1,g} + \bar{b}_{1,h} \\
&\quad + \left(\sum_{i=1}^{H}\left(\frac{\mathrm{e}'_i\Sigma'_h\left(\Sigma'_{gh}\bar{b}_{n-1,g} + \bar{b}_{n-1,h} + Z_4\right)}{1 - \mathrm{e}'_i\Sigma'_h\left(\Sigma'_{gh}\bar{b}_{n-1,g} + \bar{b}_{n-1,h} + Z_4\right)} - \frac{\mathrm{e}'_i\Sigma'_h Z_4}{1 - \mathrm{e}'_i\Sigma'_h Z_4}\right) \mathrm{e}'_i\Sigma_h^{-1}\Phi_h\right)' \\
&\quad + \frac{1}{2}\left(I_H \otimes \bar{b}_{n-1,g}\right)' \tilde{\Sigma}_g\tilde{\Sigma}'_g\left(\iota_H \otimes \bar{b}_{n-1,g}\right) + \left(I_H \otimes Z_2\right)' \tilde{\Sigma}_g\tilde{\Sigma}'_g\left(\iota_H \otimes \bar{b}_{n-1,g}\right)
\end{aligned}
$$

Real yields are $y_t^{(n)} = a_n + b'_{n,g} g_t + b'_{n,h} h_t$ with $a_n = -\frac{1}{n}\bar{a}_n$, $b_{n,g} = -\frac{1}{n}\bar{b}_{n,g}$ and $b_{n,h} = -\frac{1}{n}\bar{b}_{n,h}$.

Appendix D.2   Nominal bonds

Similar to the solution for the real bond, we guess and then verify. The solution for zero coupon nominal bonds is $P_t^{\$,(n)} = \exp\left(\bar{a}_n^{\$} + \bar{b}_{n,g}^{\$,\prime}g_t + \bar{b}_{n,h}^{\$,\prime}h_t\right)$ for some unknown coefficients $\bar{a}_n^{\$}$ and $\bar{b}_{n,g}^{\$}$ and $\bar{b}_{n,h}^{\$}$. For maturity $n = 1$, the payoff is guaranteed to be $P_{t+1}^{\$,(0)} = 1$ in the next period, in which case $P_t^{\$,(1)} = \mathrm{E}_t\left[M_{t+1}^{\$}\right]$. The solutions are

$$
\begin{aligned}
\bar{a}_1^{\$} &= \ln(\beta) - (\eta Z_c + Z_\pi)'\left(\mu_g - \Sigma_{gh}\Sigma_h\nu_h - \eta\vartheta\lambda_0\right) \\
&\quad + \frac{(\vartheta-1)}{\vartheta}\sum_{i=1}^{H}\nu_{h,i}\log\left(1 - \mathrm{e}_i'\Sigma_h'\vartheta Z_3\right) - \sum_{i=1}^{H}\nu_{h,i}\log\left(1 - \mathrm{e}_i'\Sigma_h'Z_5\right) \\
&\quad - \frac{(\vartheta-1)\vartheta}{2}Z_1'\Sigma_{0,g}\Sigma_{0,g}'Z_1 + \frac{1}{2}Z_2'\Sigma_{0,g}\Sigma_{0,g}'Z_2 + \frac{1}{2}Z_\pi'\Sigma_{0,g}\Sigma_{0,g}'Z_\pi - Z_2'\Sigma_{0,g}\Sigma_{0,g}'Z_\pi \\
\bar{b}_{1,g}^{\$} &= -(\Phi_g - \eta\vartheta\lambda_g)'(\eta Z_c + Z_\pi) \\
\bar{b}_{1,h}^{\$} &= -(\Phi_{gh} - \Sigma_{gh}\Phi_h)'(\eta Z_c + Z_\pi) \\
&\quad - (\vartheta-1)\left(\sum_{i=1}^{H}\frac{\mathrm{e}_i'\Sigma_h'Z_3}{1 - \vartheta\mathrm{e}_i'\Sigma_h'Z_3}\mathrm{e}_i'\Sigma_h^{-1}\Phi_h\right)' + \left(\sum_{i=1}^{H}\frac{\mathrm{e}_i'\Sigma_h'Z_5}{1 - \mathrm{e}_i'\Sigma_h'Z_5}\mathrm{e}_i'\Sigma_h^{-1}\Phi_h\right)' \\
&\quad + \frac{1}{2}(I_H \otimes Z_\pi)'\tilde{\Sigma}_g\tilde{\Sigma}_g'(\iota_H \otimes Z_\pi) - (I_H \otimes Z_2)'\tilde{\Sigma}_g\tilde{\Sigma}_g'(\iota_H \otimes Z_\pi) \\
&\quad + \frac{1}{2}(I_H \otimes Z_2)'\tilde{\Sigma}_g\tilde{\Sigma}_g'(\iota_H \otimes Z_2) - \frac{(\vartheta-1)\vartheta}{2}(I_H \otimes Z_1)'\tilde{\Sigma}_g\tilde{\Sigma}_g'(\iota_H \otimes Z_1)
\end{aligned}
$$

where bond prices only exist if $\mathrm{e}_i'\Sigma_h'Z_5 < 1$ for $i = 1, \ldots, H$. At longer maturities n, we use the fact that $P_t^{\$,(n)} = \mathrm{E}_t\left[\exp\left(m_{t+1}^{\$}\right)P_{t+1}^{\$,(n-1)}\right]$. The bond loadings are

$$
\begin{aligned}
\bar{a}_n^{\$} &= \bar{a}_{n-1}^{\$} + \bar{a}_1^{\$} + (\mu_g - \Sigma_{gh}\Sigma_h\nu_h - \eta\vartheta\lambda_0)'\bar{b}_{n-1,g}^{\$} \\
&\quad + \sum_{i=1}^{H}\nu_{h,i}\log\left(\frac{1 - \mathrm{e}_i'\Sigma_h'Z_5}{1 - \mathrm{e}_i'\Sigma_h'\left(\Sigma_{gh}'\bar{b}_{n-1,g}^{\$} + \bar{b}_{n-1,h}^{\$} + Z_5\right)}\right) \\
&\quad + \frac{1}{2}\bar{b}_{n-1,g}^{\$,\prime}\Sigma_{0,g}\Sigma_{0,g}'\bar{b}_{n-1,g}^{\$} + \bar{b}_{n-1,g}^{\$}\Sigma_{0,g}\Sigma_{0,g}'(Z_2 - Z_\pi) \\
\bar{b}_{n,g}^{\$} &= (\Phi_g - \eta\vartheta\lambda_g)'\bar{b}_{n-1,g}^{\$} + \bar{b}_{1,g}^{\$} \\
\bar{b}_{n,h}^{\$} &= (\Phi_{gh} - \Sigma_{gh}\Phi_h)'\bar{b}_{n-1,g}^{\$} + \bar{b}_{1,h}^{\$} \\
&\quad + \left(\sum_{i=1}^{H}\left(\frac{\mathrm{e}_i'\Sigma_h'\left(\Sigma_{gh}'\bar{b}_{n-1,g}^{\$} + \bar{b}_{n-1,h}^{\$} + Z_5\right)}{1 - \mathrm{e}_i'\Sigma_h'\left(\Sigma_{gh}'\bar{b}_{n-1,g}^{\$} + \bar{b}_{n-1,h}^{\$} + Z_5\right)} - \frac{\mathrm{e}_i'\Sigma_h'Z_5}{1 - \mathrm{e}_i'\Sigma_h'Z_5}\right)\mathrm{e}_i'\Sigma_h^{-1}\Phi_h\right)' \\
&\quad + \frac{1}{2}\left(I_H \otimes \bar{b}_{n-1,g}^{\$}\right)'\tilde{\Sigma}_g\tilde{\Sigma}_g'\left(\iota_H \otimes \bar{b}_{n-1,g}^{\$}\right) + \left(I_H \otimes \bar{b}_{n-1,g}^{\$}\right)'\tilde{\Sigma}_g\tilde{\Sigma}_g'(\iota_H \otimes (Z_2 - Z_\pi))
\end{aligned}
$$

Nominal yields are $y_t^{\$,(n)} = a_n^{\$} + b_{n,g}^{\$,\prime}g_t + b_{n,h}^{\$,\prime}h_t$ with $a_n^{\$} = -\frac{1}{n}\bar{a}_n^{\$}$, $b_{n,g}^{\$} = -\frac{1}{n}\bar{b}_{n,g}^{\$}$ and $b_{n,h}^{\$} = -\frac{1}{n}\bar{b}_{n,h}^{\$}$. The nominal short term interest rate is

$$
r_t^{\$} = y_t^{\$,(1)} = a_1^{\$} + b_{1,g}^{\$,\prime}g_t + b_{1,h}^{\$,\prime}h_t \tag{D.18}
$$

Appendix E Proof of Propositions

Appendix E.1 General case

Define the fixed point problem

$$\kappa_1 = \frac{\exp(\bar{p}c)}{1 + \exp(\bar{p}c)}$$

$$\kappa_0 = \ln(1 + \exp(\bar{p}c)) - \kappa_1 \bar{p}c$$

$$D'_g = (1 - \eta) Z'_c (\Phi_g - \vartheta \eta \lambda_g) (I - \kappa_1 (\Phi_g - \vartheta \eta \lambda_g))^{-1}$$

$$Z_1 = (1 - \eta) Z_c + \kappa_1 D_g$$

$$Z_3 = \Sigma'_{gh} ((1 - \eta) Z_c + \kappa_1 D_g) + \kappa_1 D_h$$

$$D_h = (\Phi_{gh} - \Sigma_{gh} \Phi_h)' Z_1 + \frac{\vartheta}{2} (\iota_H \otimes Z_1)' \tilde{\Sigma}_g \tilde{\Sigma}'_g (I_H \otimes Z_1) + \sum_{i=1}^{H} \frac{e'_i \Sigma'_h Z_3}{1 - \vartheta e'_i \Sigma'_h Z_3} \Phi'_h \Sigma_h^{-1,'} e_i$$

$$(1 - \kappa_1) D_0 = \ln(\beta) + \kappa_0 + Z'_1 (\mu_g - \Sigma_{gh} \Sigma_h \nu_h - \eta \vartheta \lambda_0)$$

$$- \frac{1}{\vartheta} \sum_{i=1}^{H} \nu_{h,i} \ln(1 - e'_i \Sigma'_h \vartheta Z_3) + \frac{\vartheta}{2} Z'_1 \Sigma_{0,g} \Sigma'_{0,g} Z_1$$

$$f(\bar{p}c) = D_0 + D'_g \bar{\mu}_g + D'_h \bar{\mu}_h$$

which is solved if $\bar{p}c = f(\bar{p}c)$.

Assumptions 1

The vector of coefficients D_h is a solution to the system of non-linear equations in (C.14). The system of equations does not necessarily have a real solution for a given parameter vector θ.

In the special case when Σ_h and Φ_h are diagonal, each loading (C.15) reduces to a quadratic equation given by (C.16) that can be solved separately for each element i. The solutions are in (C.17). The fixed point problem only has a solution when $D_{h,i}$ is real. The coefficient $D_{h,i}$ is real if and only if the parameters satisfy

$$\left(\kappa_1 \Phi_{h,i} - \kappa_1 \vartheta \Sigma_{h,i} \bar{D}_i - 1 + \vartheta \Sigma_{h,i} \bar{Z}_{3,i} \right)^2 - 4 \kappa_1 \vartheta \Sigma_{h,i} \left[\bar{D}_i \left(1 - \vartheta \Sigma_{h,i} \bar{Z}_{3,i} \right) + \bar{Z}_{3,i} \Phi_{h,i} \right] \geq 0$$

for $i = 1, \ldots, H$.

In order to solve for the price to consumption ratio pc_t, the conditional expectation in (C.13) must exist. This condition is

$$\vartheta e'_i \Sigma'_h \left[\Sigma'_{gh} ((1 - \eta) Z_c + \kappa_1 D_g) + \kappa_1 D_h \right] < 1 \qquad i = 1, \ldots, H \tag{E.19}$$

This defines another set of restrictions across the parameters θ of the model.

Proof of Proposition 1

First, derive the limiting property for $\bar{p}c \to -\infty$: $\lim_{\bar{p}c \to -\infty} \kappa_1 = 0$ and $\lim_{\bar{p}c \to -\infty} \kappa_0 = 0$. In this case, both D_0 and D_h are finite due to $\vartheta e'_i \Sigma'_h Z_3 < 1$ in Assumption 1. Therefore, $\tilde{p}c$ is finite, so $\lim_{\bar{p}c \to -\infty} (\bar{p}c - \tilde{p}c) \to -\infty$.

Next, derive the limiting property for $\bar{p}c \to \infty$: $\lim_{\bar{p}c \to \infty} \kappa_1 = 1$ and $\lim_{\bar{p}c \to \infty} \kappa_0 = 0$. This implies D_g is finite as long as the eigenvalue of $(\Phi_g - \vartheta \eta \lambda_g)$ for consumption is smaller than 1. D_h is finite due to $\vartheta e'_i \Sigma'_h Z_3 < 1$. And $\lim_{\bar{p}c \to \infty} (1 - \kappa_1) D_0 = \lim_{\bar{p}c \to \infty} \ln(\beta) + Z'_1 (\mu_g - \Sigma_{gh} \Sigma_h \nu_h - \eta \vartheta \lambda_0) - \frac{1}{\vartheta} \sum_{i=1}^{H} \nu_{h,i} \ln(1 - e'_i \Sigma'_h \vartheta Z_3) + \frac{\vartheta}{2} Z'_1 \Sigma_{0,g} \Sigma'_{0,g} Z_1$. The right hand side is finite due to $\vartheta e'_i \Sigma'_h Z_3 < 1$. Therefore, $\lim_{\bar{p}c \to \infty} \kappa_1 = 1$ leads to an infinite D_0. The condition $\lim_{\bar{p}c \to \infty} D_0 \to -\infty$ implies $\lim_{\bar{p}c \to \infty} (\bar{p}c - \tilde{p}c) \to \infty$, which together $\lim_{\bar{p}c \to -\infty} (\bar{p}c - \tilde{p}c) \to -\infty$ guarantees there exists a solution for the fixed point problem.

With $\kappa_1 < 1$, the condition $\lim_{\bar{p}c \to \infty} D_0 \to -\infty$ is equivalent to

$$\beta < \lim_{\bar{p}c \to \infty} \exp\left[- \left(Z_1' \left(\mu_g - \Sigma_{gh}\Sigma_h\nu_h - \eta\vartheta\lambda_0 \right) - \frac{1}{\vartheta}\sum_{i=1}^{H}\nu_{h,i}\ln\left(1 - e_i'\Sigma_h'\vartheta Z_3\right) + \frac{\vartheta}{2}Z_1'\Sigma_{0,g}\Sigma_{0,g}'Z_1 \right) \right].$$

Therefore, the boundary condition is

$$\bar{\beta} = \exp\left[- \left(Z_1^{\infty\prime} \left(\mu_g - \Sigma_{gh}\Sigma_h\nu_h - \eta\vartheta\lambda_0 \right) - \frac{1}{\vartheta}\sum_{i=1}^{H}\nu_{h,i}\ln\left(1 - e_i'\Sigma_h'\vartheta Z_3^{\infty}\right) + \frac{\vartheta}{2}Z_1^{\infty\prime}\Sigma_{0,g}\Sigma_{0,g}'Z_1^{\infty} \right) \right],$$

where
$$Z_1^{\infty} = (1-\eta)\,Z_c + D_g^{\infty}$$
$$D_g^{\infty\prime} = (1-\eta)\,Z_c'\left(\Phi_g - \vartheta\eta\lambda_g\right)\left(I - (\Phi_g - \vartheta\eta\lambda_g)\right)^{-1},$$
$$Z_3^{\infty} = \bar{Z}_3^{\infty} + D_h^{\infty}$$
$$\bar{Z}_3^{\infty} = \Sigma_{gh}'Z_1^{\infty}$$
$$D_{h,i}^{\infty} = -\frac{1}{2}\left(\frac{\Phi_{h,i}-1}{\vartheta\Sigma_{h,i}} - \bar{D}_i^{\infty} + \bar{Z}_{3,i}^{\infty}\right) - \sqrt{\frac{1}{4}\left(\frac{\Phi_{h,i}-1}{\vartheta\Sigma_{h,i}} - \bar{D}_i^{\infty} + \bar{Z}_{3,i}^{\infty}\right)^2 - \frac{1}{\vartheta\Sigma_{h,i}}\left[\bar{D}_i^{\infty}\left(1 - \vartheta\Sigma_{h,i}\bar{Z}_{3,i}^{\infty}\right) + \bar{Z}_{3,i}^{\infty}\Phi_{h,i}\right]}$$
$$\bar{D}^{\infty} = \left(\Phi_{gh} - \Sigma_{gh}\Phi_h\right)'Z_1^{\infty} + \frac{\vartheta}{2}\left(\iota_H \otimes Z_1^{\infty}\right)'\tilde{\Sigma}_g\tilde{\Sigma}_g'\left(I_H \otimes Z_1^{\infty}\right)$$

Appendix E.2   Special case with Gaussian dynamics

The fixed point problem simplifies to

$$\kappa_1 = \frac{\exp(\bar{p}c)}{1 + \exp(\bar{p}c)}$$
$$\kappa_0 = \ln\left(1 + \exp(\bar{p}c)\right) - \kappa_1\bar{p}c$$
$$D_g' = (1-\eta)\,Z_c'\Phi_g^{\mathbb{Q}^{\$}}\left(I - \kappa_1\Phi_g^{\mathbb{Q}^{\$}}\right)^{-1}$$
$$Z_1 = (1-\eta)\,Z_c + \kappa_1 D_g$$
$$D_0\left(1 - \kappa_1\right) = \ln(\beta) + \kappa_0 + Z_1'\mu_g^* + \frac{1}{2}\vartheta Z_1'\Sigma_{0,g}\Sigma_{0,g}'Z_1$$
$$\tilde{p}c = D_0 + D_g'\bar{\mu}_g$$

which is solved if $\bar{p}c = \tilde{p}c$.

First, the condition in Proposition 1 becomes

$$\beta < \lim_{\bar{p}c \to \infty} \exp\left(-Z_1'\mu_g^* - \frac{1}{2}\vartheta Z_1'\Sigma_{0,g}\Sigma_{0,g}'Z_1\right), \tag{E.20}$$

and $\bar{\beta}$ simplifies to

$$\bar{\beta} = \exp\left[- \left(Z_1^{\infty\prime}\mu_g^* + \frac{\vartheta}{2}Z_1^{\infty\prime}\Sigma_{0,g}\Sigma_{0,g}'Z_1^{\infty} \right) \right],$$

where $Z_1^{\infty} \equiv \lim_{\bar{p}c \to \infty} Z_1(\bar{p}c) = (1-\eta)\,Z_c + D_g^{\infty}$ and $D_g^{\infty\prime} \equiv \lim_{\bar{p}c \to \infty} D_g(\bar{p}c)' = (1-\eta)\,Z_c'\Phi_g^{\mathbb{Q}^{\$}}\left(I - \Phi_g^{\mathbb{Q}^{\$}}\right)^{-1}$.

Proof of Corollary 1

1. The condition (E.20) is guaranteed by $Z_1^{\infty\prime}\mu_g^* \le 0$ and $\vartheta < 0$ for any $\beta \le 1$. And $\frac{1-\gamma}{1-\psi} > 0$ is equivalent to $\vartheta < 0$,

2. A stronger condition is

$$\beta \le 1 < \lim_{\bar{p}c \to \infty} \exp\left(-Z_1'\mu_g^* - \frac{1}{2}\vartheta Z_1'\Sigma_{0,g}\Sigma_{0,g}'Z_1\right),$$

which can be simplified to

$$\gamma \; > \; 1 + \frac{2Z_c'\left(I - \Phi_g^{\mathbb{Q}^\$}\right)^{-1}\mu_g^*}{Z_c'\left(I - \Phi_g^{\mathbb{Q}^\$}\right)^{-1}\Sigma_{0,g}\Sigma_{0,g}'\left(I - \Phi_g^{\mathbb{Q}^\$,\prime}\right)^{-1}Z_c}, \qquad \text{if} \quad \psi > 1$$

$$\gamma \; < \; 1 + \frac{2Z_c'\left(I - \Phi_g^{\mathbb{Q}^\$}\right)^{-1}\mu_g^*}{Z_c'\left(I - \Phi_g^{\mathbb{Q}^\$}\right)^{-1}\Sigma_{0,g}\Sigma_{0,g}'\left(I - \Phi_g^{\mathbb{Q}^\$,\prime}\right)^{-1}Z_c}. \qquad \text{if} \quad \psi < 1$$

hence $\bar{\gamma}(\theta^{\mathbb{P}}, \theta^{\lambda}) = 1 + \frac{2Z_c'\left(I - \Phi_g^{\mathbb{Q}^\$}\right)^{-1}\mu_g^*}{Z_c'\left(I - \Phi_g^{\mathbb{Q}^\$}\right)^{-1}\Sigma_{0,g}\Sigma_{0,g}'\left(I - \Phi_g^{\mathbb{Q}^\$,\prime}\right)^{-1}Z_c}$, does not depend on ψ.

3. We have $\frac{d\vartheta}{d\gamma} = -\frac{1}{1-\eta}$, $\frac{dD_g^{\infty\prime}}{d\gamma} = 0$ and $\frac{dZ_1^{\infty\prime}}{d\gamma} = \frac{dD_g^{\infty\prime}}{d\gamma} = 0$. Hence, the derivative of $\ln\bar{\beta}$ w.r.t. γ is

$$\frac{d\ln\bar{\beta}}{d\gamma} \;=\; \frac{1}{2(1-\eta)}Z_1^{\infty\prime}\Sigma_{0,g}\Sigma_{0,g}'Z_1^{\infty}$$

$\frac{d\ln\bar{\beta}}{d\gamma} = \frac{1}{\bar{\beta}}\frac{d\bar{\beta}}{d\gamma}$ implies that the two derivatives have the same sign. Therefore, for $\psi > 1$, then $\frac{d\bar{\beta}}{d\gamma} > 0$; for $\psi < 1$, then $\frac{d\bar{\beta}}{d\gamma} < 0$.

Appendix F   MCMC and particle filters

Appendix F.1   MCMC

Our MCMC algorithm is the particle Gibbs (PG) sampler. It iterates between two broad steps: (i) drawing the latent state variables $(g_{1:T}, h_{0:T})$ conditional on the model's parameters; and (ii) drawing the model's parameters $\theta^{\mathbb{P}}$ given the latent state variables. We make heavy use of the fact that the model is a conditionally linear Gaussian state space model.

Appendix F.1.1   Conditionally linear, Gaussian state space form

Conditional on $h_{0:T}$, the model is a linear, Gaussian state space model. We write the model using the state space form of Durbin and Koopman(2012) given by

$$
\begin{aligned}
Y_t &= Zg_t + d + \eta_t^* & \eta_t^* &\sim \mathrm{N}\left(0, H\right), & \text{(F.21)} \\
g_{t+1} &= Tg_t + c_t + R\varepsilon_{t+1}^* & \varepsilon_{t+1}^* &\sim \mathrm{N}\left(0, Q_t\right), & \text{(F.22)}
\end{aligned}
$$

where $Y_t = (\Delta c_t \;\; \pi_t)'$. The models in this paper can placed in this state space form as

$$Z \;=\; \begin{pmatrix} Z_c \\ Z_\pi \end{pmatrix} \qquad T \;=\; \Phi_g \qquad d \;=\; 0_{2\times 1} \qquad H \;=\; 0_{2\times 2}$$

$$c_t = \mu_g + \Phi_{gh}h_t + \Sigma_{gh}\varepsilon_{h,t+1} \qquad Q_t \;=\; \Sigma_{g,t}\Sigma_{g,t}'$$

For some models, there are free, estimable parameters in the matrices $(\mu_g, \Phi_{gh}, \Sigma_{gh})$. We can place these in the state vector. This allows any free parameters in $(\mu_g, \Phi_{gh}, \Sigma_{gh})$ to be drawn jointly with the state variables $g_{1:T}$. It also allows us to marginalize over them when drawing other parameters, see ? for discussion.

Appendix F.1.2   Drawing the state variables

We draw $(g_{1:T}, h_{0:T})$ from their full conditional distribution in two steps.

$$g_{1:T} \sim p\left(g_{1:T}|Y_{1:T}, h_{0:T}, \theta^{\mathbb{P}}\right)$$
$$h_{0:T} \sim p\left(h_{1:T}|Y_{1:T}, g_{1:T}, \theta^{\mathbb{P}}\right)$$

We draw $g_{1:T}$ conditional on $h_{0:T}$ from the conditionally linear, Gaussian state space model (F.21) and (F.22) using a forward filtering backward sampling algorithm or simulation smoother; see, e.g. Durbin and Koopman(2002). Conditional on the draw for $g_{1:T}$, we draw $h_{0:T}$ using a particle Gibbs sampler.

There are two PG samplers developed in the literature. The original PG sampler of Andrieu, Doucet, and Holenstein(2010) with the backward-sampling pass developed by Whiteley(2010), see Creal and Tsay(2015). And, the PG sampler with ancestor sampling (PGAS) of Lindsten, Jordan, and Schön(2145–2184). The former algorithm is simple to implement for Model #1. We describe its implementation here.

Let J be the total number of particles. In our work, we select $J = 100$. The PG sampler starts with a set of existing particles $h_{0:T}^{(1)}$ that were drawn from the previous iteration.
For $t = 1, \ldots, T$, run:

- For $j = 2, \ldots, J$, draw from a proposal: $(h_t, h_{t-1})^{(j)} \sim q\left(h_t, h_{t-1}|g_{t-1:t}, \theta^{\mathbb{P}}\right)$.

- For $j = 1, \ldots, J$, calculate the importance weight:

$$w_t^{(j)} \quad \propto \quad \frac{p\left(g_t|g_{t-1}, h_t^{(j)}, h_{t-1}^{(j)}, \theta^{\mathbb{P}}\right) p\left(h_t^{(j)}|h_{t-1}^{(j)}, \theta^{\mathbb{P}}\right)}{q\left(h_t^{(j)}, h_{t-1}^{(j)}|g_{t-1:t}, \theta^{\mathbb{P}}\right)}$$

- For $j = 1, \ldots, J$, normalize the weights: $\hat{w}_t^{(j)} = \frac{w_t^{(j)}}{\sum_{j=1}^J w_t^{(j)}}$.

- Conditionally resample the particles $\left\{h_t^{(j)}\right\}_{j=1}^J$ with probabilities $\left\{\hat{w}_t^{(j)}\right\}_{j=1}^J$. In this step, the first particle $h_t^{(1)}$ always gets resampled and may be randomly duplicated.

Implementation of the PG sampler is different than a standard particle filter due to the "conditional" resampling algorithm used in the last step. We use the conditional multinomial resampling algorithm from Andrieu, Doucet, and Holenstein(2010).

In the original PG sampler, the particles $\left\{h_t^{(j)}\right\}_{j=1}^J$ are stored for $t = 1, \ldots, T$ and a single trajectory is sampled using the probabilities from the last iteration $\left\{\hat{w}_T^{(j)}\right\}_{j=1}^J$. An important improvement upon the original PG sampler was introduced by Whiteley(2010), who suggested drawing the path of the state variables from the discrete particle approximation using the backwards sampling algorithm of Godsill, Doucet, and West(2004). On the forwards pass, we store the normalized weights and particles $\left\{\hat{w}_t^{(m)}, h_{i,t}^{(m)}\right\}_{m=1}^M$ for $t = 1, \ldots, T$. We draw a path of the state variables $(h_1^*, \ldots, h_T^*)$ from this discrete distribution.

At $t = T$, draw a particle $h_T^* = h_T^{(j)}$ with probability $\hat{w}_T^{(j)}$.
For $t = T - 1, \ldots, 0$, run:

- For $j = 1, \ldots, J$, calculate the backwards weights: $w_{t|T}^{(j)} \propto \hat{w}_t^{(j)} p\left(h_{t+1}^*|h_t^{(j)}, \theta\right)$

- For $j = 1, \ldots, J$, normalize the weights: $\hat{w}_{t|T}^{(j)} = \frac{w_{t|T}^{(j)}}{\sum_{j=1}^J w_{t|T}^{(j)}}$.

- Draw a particle $h_t^* = h_t^{(j)}$ with probability $\hat{w}_{t|T}^{(j)}$.

The draw $h_{0:T} = (h_0^*, \ldots, h_T^*)$ is a draw from the full-conditional distribution. In practice, when the dimension H of h_t is high, the number of particles J required for satisfactory performance can be quite large. In this case, we can separate each element of the state vector $h_{i,t}$ for $i = 1, \ldots, H$ and draw them one at a time.

Appendix F.1.3 Drawing the parameters

We block the parameters into groups that are highly correlated. These groups can be separated into parameters governing the dynamics of g_t and the parameters that enter the dynamics of volatility h_t.

1. **Drawing parameters in $\bar{\mu}_g, \Phi_{gh}, \Sigma_{gh}$:** We place these parameters in the state vector and draw them jointly with the Gaussian state variables.

2. **Drawing parameters in $\Phi_g, \Sigma_{0,g}$:** We use the independence Metropolis Hastings algorithm. Conditional on the volatility state variables $h_{0:T}$, the model is a linear, Gaussian state space model (F.21) and (F.22). We maximize the likelihood using the Kalman filter and calculate the Hessian at the posterior mode. We then draw from a Student's t proposal distribution with mean equal to the posterior mode and covariance matrix equal to the inverse Hessian at the mode; see, e.g. **?** for details.

3. **Drawing parameters of the volatility process $\bar{\mu}_h, \Phi_h, \Sigma_h$:** We use an independence Metropolis-Hastings step. When drawing these parameters, we can marginalize out the Gaussian state variables using the Kalman filter. Conditional on the remaining parameters of the model (which we omit), the target distribution of ν_h, Φ_h, Σ_h can be written as

$$p\left(\bar{\mu}_h, \Phi_h, \Sigma_h | Y_{1:T}, h_{0:T}\right) \quad \propto \quad p\left(Y_{1:T} | h_{0:T}, \bar{\mu}_h, \Phi_h, \Sigma_h\right) p\left(h_{0:T} | \bar{\mu}_h, \Phi_h, \Sigma_h\right) p\left(\bar{\mu}_h, \Phi_h, \Sigma_h\right)$$

where $p\left(Y_{1:T} | h_{0:T}, \bar{\mu}_h, \Phi_h, \Sigma_h\right)$ is the likelihood from the Kalman filter, $p\left(h_{0:T} | \bar{\mu}_h, \Phi_h, \Sigma_h\right)$ is the transition density of the volatility process (B.7). We maximize this target density and calculate the Hessian at the posterior mode. We then draw from a Student's t proposal distribution with mean equal to the posterior mode and covariance matrix equal to the inverse Hessian at the mode.

For Gaussian models, we draw the free parameters in $\Sigma_{0,g}$ instead of $\bar{\mu}_h, \Phi_h, \Sigma_h$.

Appendix F.2 Particle filter

To estimate the structural parameters (β, γ, ψ) and the preference parameters θ^λ, we run cross-sectional regressions on filtered estimates of the factors. In order to calculate the filtered estimates of the state variables, we use a particle filter. The particle filter we implement is the mixture Kalman filter of Chen and Liu(2000). Let $g_{t|t-1}$ denote the conditional mean and $P_{t|t-1}$ the conditional covariance matrix of the one-step ahead predictive distribution $p(g_t | Y_{1:t-1}, h_{0:t-1}; \theta)$ of a conditionally linear, Gaussian state space model. Similarly, let $g_{t|t}$ denote the conditional mean and $P_{t|t}$ the conditional covariance matrix of the filtering distribution $p(g_t | Y_{1:t}, h_{0:t}; \theta)$. Conditional on the volatilities $h_{0:T}$, these quantities can be calculated by the Kalman filter.

Let J denote the number of particles and let $Y_t = (\pi_t, \Delta c_t)$ be $N \times 1$. The particle filter then proceeds as follows:

At $t = 0$, for $i = 1, \ldots, J$, set $w_0^{(i)} = \frac{1}{J}$ and

- Draw $h_0^{(i)} \sim p\left(h_0; \theta\right)$ and calculate $\Sigma_{g,0}^{(i)} \Sigma_{g,0}^{(i),\prime}$.

- Set $g_{1|0}^{(i)} = \bar{\mu}_g + \Phi_{gh} \bar{h}_0^{(i)}$, $P_{1|0}^{(i)} = \Sigma_{g,0}^{(i)} \Sigma_{g,0}^{(i),\prime}$,

- Set $\ell_0 = 0$.

For $t = 1, \ldots, T$ do:

STEP 1: For $i = 1, \ldots, J$:

- Draw from the transition density: $h_{t+1}^{(i)} \sim p(h_{t+1} | h_t^{(i)}; \theta)$ given by:

$$
\begin{aligned}
z_{j,t+1}^{(i)} &\sim \quad \text{Poisson}\left(\mathbf{e}_j' \Sigma_h^{-1} \Phi_h h_t^{(i)}\right) \qquad j = 1, \ldots, H \\
w_{j,t+1}^{(i)} &\sim \quad \text{Gamma}\left(\nu_{h,j} + z_{j,t+1}^{(i)}, 1\right) \qquad j = 1, \ldots, H \\
h_{t+1}^{(i)} &= \quad \Sigma_h w_{t+1}^{(i)}
\end{aligned}
$$

- Calculate $c_t^{(i)}$ and $Q_t^{(i)}$ using $h_t^{(i)}$.

$$
\begin{aligned}
c_t^{(i)} &= \Phi_{gh} h_t^{(i)} + \Sigma_{gh} \varepsilon_{h,t+1}^{(i)} \\
Q_t^{(i)} &= \Sigma_{g,t}^{(i)} \Sigma_{g,t}^{(i),\prime}
\end{aligned}
$$

- Run the Kalman filter:

$$
\begin{aligned}
v_t^{(i)} &= Y_t - Z g_{t|t-1}^{(i)} - d \\
F_t^{(i)} &= Z P_{t|t-1}^{(i)} Z' + H \\
K_t^{(i)} &= P_{t|t-1}^{(i)} Z' \left(F_t^{(i)} \right)^{-1} \\
g_{t|t}^{(i)} &= g_{t|t-1}^{(i)} + K_t^{(i)} v_t^{(i)} \\
P_{t|t}^{(i)} &= P_{t|t-1}^{(i)} - K_t^{(i)} Z P_{t|t-1}^{(i)} \\
g_{t+1|t}^{(i)} &= T g_{t|t}^{(i)} + c_t^{(i)} \\
P_{t+1|t}^{(i)} &= T P_{t|t}^{(i)} T' + R Q_t^{(i)} R'
\end{aligned}
$$

- Calculate the weight: $\log\left(w_t^{(i)} \right) = \log\left(\hat{w}_{t-1}^{(i)} \right) - 0.5 N \log(2\pi) - 0.5 \log |F_t^{(i)}| - \frac{1}{2} v_t^{(i)\prime} \left(F_t^{(i)} \right)^{-1} v_t^{(i)}$.

STEP 2: Calculate an estimate of the log-likelihood: $\ell_t = \ell_{t-1} + \log\left(\sum_{i=1}^{J} w_t^{(i)} \right)$.

STEP 3: For $i = 1, \ldots, J$, calculate the normalized importance weights: $\hat{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^{J} w_t^{(j)}}$.

STEP 4: Calculate the effective sample size $E_t = \frac{1}{\sum_{j=1}^{J} \left(\hat{w}_t^{(j)} \right)^2}$.

STEP 5: If $E_t < 0.5J$, resample $\left\{ g_{t+1|t}^{(i)}, P_{t+1|t}^{(i)}, h_{t+1}^{(i)} \right\}_{i=1}^{J}$ with probabilities $\hat{w}_t^{(i)}$ and set $\hat{w}_t^{(i)} = \frac{1}{J}$.
STEP 6: Increment time and return to STEP 1.

Within the particle filter, we use the residual resampling algorithm of Liu and Chen(1998). We set $J = 100000$.