Term Structure of Interest Rates under Long-Run Risks and Incomplete Information

Ji Zhou *

Alex Paseka †

July 15, 2013

Abstract

In this paper we derive the term structure of default-free zero rates under the Epstein-Zin utility function, non-i.i.d. consumption growth, and incomplete information about fundamentals. We extend the continuous-time long-run risks model of Eraker (2008) to an incomplete information environment in which agents learn about unobservable persistent component of the conditional mean of consumption growth. In equilibrium, agents learn about the conditional mean of consumption growth and price zero-coupon bonds simultaneously under a new measure, which is generated by information observed by agents. We derive analytic formulas for a zero coupon bond price and that for spot rates under the new measure.

*Corresponding author: Department of Accounting and Finance, University of Manitoba, Winnipeg, Manitoba, Canada. Tel. (204)474-6985, Fax (204)474-7545, *Email address*: umzhouj0@cc.umanitoba.ca

†Corresponding author: Department of Accounting and Finance, University of Manitoba, Winnipeg, Manitoba, Canada.

1 Related Literature

The Long-run Risk model is firstly proposed by Bansal & Yaron (2004). In the LRR model, current shock to expected growth have a persistent effect on the expectations about consumption and dividend growth. Furthermore, the conditional volatility of consumption is time varying. Therefore, investors who are exposed to these two types of risk require higher premium for holding equities. BY form the LRR model based on an Epstein & Zin (1989) recursive preferences.

2 Long-Run Risk Model with Learning

Eraker (2008) extends Bansal & Yaron (2004) economy to general affine state variable processes. Both Eraker and Bansal and Yaron assume that the conditional mean of consumption growth, x_t, is observable. In our model we assume that agents do not directly observe the conditional mean of consumption growth, x_t. Unlike Eraker (2008), we model how agents learn about the value of x_t by extracting information from observing both consumption growth and dividend growth. The optimal forecast of the conditional mean of consumption growth is solved for jointly with equilibrium prices. The Markov property of the state variables allows us to break down the agent's optimization problem into two stages. First, we use a linear version (Kalman-Bucy filter) of a general filter to derive the process of an agent's best estimate of x_t, $\widehat{x}_t$. Second, given the optimal estimate, the model is effectively reduced to a full-information model in which all expectations are computed under an information set available to the agents.

We retain main features of models by Bansal & Yaron (2004) and Eraker (2008) such as recursive preferences and non-i.i.d. consumption and dividend growth. We list the assumptions below.

Assumption 1 *We assume an infinitely lived representative agent economy with stochastic differential utility preferences of Duffie & Epstein (1992a):*

$$J(G_t, W_t, t) = E_t \int_t^{\infty} f(C_s, J_s) ds^1 \qquad (1)$$

with the normalized aggregator

$$f(C_t, J_t) = \frac{\beta}{1 - \frac{1}{\psi}}(1-\gamma)J_t \left[\left(\frac{C_t}{((1-\gamma)J_t)^{\frac{1}{1-\gamma}}} \right)^{1-\frac{1}{\psi}} - 1 \right] \qquad (2)$$

where C_t is consumption at time t, J_t is the recursive utility at time t, β is the subjective discount rate, γ is the relative risk aversion coefficient, and ψ is the elasticity of intertemporal substitution (EIS).

Assumption 2 *Following Eraker (2008), we model the processes of consumption growth and dividend growth with drift being linear function of the persistent long-run risk component x_t:*

$$d\ln(C) \equiv dg_c = (\mu_c + x - \frac{V}{2})dt + \sqrt{V} dw_c, \qquad (3)$$

$$dg_d = (\mu_d + \phi x - \varphi_d^2 \frac{V}{2})dt + \varphi_d \sqrt{V} dw_d, \qquad (4)$$

where w_c and w_d are standard Wiener processes, $Edw_c dw_d = \rho_{cd}$, and V is conditional instantaneous variance of consumption growth.

Assumption 3 *The conditional mean of consumption growth is not observed by agents. They only know that it follows a mean reverting process with zero long-run mean:*

$$dx = -\rho x dt + \varphi_e \sqrt{V} dw_e \qquad (5)$$

where w_e are standard Wiener processes and $Edw_c dw_e = \rho_{cx}$, $Edw_e dw_d = \rho_{xd}$.

[1] Duffie & Epstein (1992a) consider a more general case of non-zero variance multiplier: $J(G, W, t) = E_t \int_t^{\infty} \left(f(C_s, J_s) + \frac{1}{2} A(J_s) \sigma_J^2(s) \right) ds$ where σ_J^2 is utility variance, and $A(J_t)$ is the variance multiplier. They show how one can transform to an ordinally equivalent preference structure with zero variance multiplier.

Assumption 4 *We assume that all state variable processes are homoscedastic, i.e., V is constant over time.*

2.1 Optimal Inference about the Conditional Mean of Consumption Growth

Due to the latent nature of process x_t, we assume that agents estimate the value of x_t from observations on both the consumption growth and the dividend growth. As a standard practice, we model the optimal forecast of the latent process by minimizing the posterior variance of the latent process. As a result, the estimation process is described by Kalman-Bucy filter. The following theorem summarizes the solution to the optimal filtering problem.

Theorem 1 *Assuming that the latent mean of consumption growth is described by (5) and an agent's inference is based on the observations of consumption and dividend growth as described in (3) and (4) the agent's best estimate of x_t, $\widehat{x}_t$ is given by*

$$d\widehat{x}_t = -\rho\widehat{x}_t dt + \Sigma_x \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}, \tag{6}$$

where $S_t = E_t[(x_t - \widehat{x}_t)^2]$ is the posterior variance of process x_t, and matrix Σ_x is shown below:

$$\Sigma_x = \frac{1}{1-\rho_{cd}^2}\left[\frac{S_t}{\varphi_d\sqrt{V}}a^T + \varphi_e\sqrt{V}b^T\right] \tag{7}$$

$$a^T = \begin{bmatrix} \varphi_d - \phi\rho_{cd}, & \phi - \varphi_d\rho_{cd} \end{bmatrix} \tag{8}$$

$$b^T = \begin{bmatrix} \rho_{cx} - \rho_{dx}\rho_{cd}, & \rho_{dx} - \rho_{cx}\rho_{cd} \end{bmatrix} \tag{9}$$

Processes $w_c^*(t)$ and $w_d^*(t)$ are standard Wiener processes under a filtration generated by *observation on consumption and dividend growth on the interval* $[0,t]$, *i.e.,* $w_c^*(t)$,$w_d^*(t) \in \mathcal{F}_t^{g_c,g_d}$*:*

$$\begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix} = \begin{bmatrix} \frac{x-\widehat{x}}{\sqrt{V}}dt + dw_c \\ \frac{\phi(x-\widehat{x})}{\varphi_d\sqrt{V}}dt + \rho_{cd}dw_c + \sqrt{1-\rho_{cd}^2}dw_d^{\perp} \end{bmatrix}$$

where $Edw_e dw_d^{\perp} = 0$.

Proof: see Appendix.

Given the optimal process $\widehat{x}_t$, processes g_c and g_d under the new probability measure corresponding to the information available to agents have the following form:

$$dg_c = (\mu_c + \widehat{x} - \frac{V}{2})dt + \sqrt{V}dw_c^*, \quad (10)$$

$$dg_d = (\mu_d + \phi\widehat{x} - \varphi_d^2\frac{V}{2})dt + \varphi_d\sqrt{V}dw_d^*. \quad (11)$$

2.2 Value Function

We make a further assumption about the asset composition of the economy.

Assumption 5 *There are* n *traded assets in the economy with price processes described by following stochastic differential equations (SDEs) under information available to agents:*

$$\frac{dS_i}{S_i} = \mu_i dt + \sigma_i^T \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}, \; i = \overline{1,n}$$

where σ_i *is a* 1×2 *row.*

To simplify further discussion we introduce the following notation:

$$\underset{n\times 2}{\Sigma} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ ... \\ \sigma_n \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ ... \\ \mu_n \end{pmatrix}, \pi = \begin{pmatrix} \mu_1 - r \\ \mu_2 - r \\ ... \\ \mu_n - r \end{pmatrix}$$

where r is the risk free rate of return.

An agent's wealth process (budget constraint) is represented by a self-financing portfolio-consumption pair $(\{\omega\}_{i=1}^{n}, C)$ as follows:

$$\begin{aligned} dW &= W\sum_{i=1}^{n}\omega_i\frac{dS_i}{S_i} - Cdt = W\omega^T\left((\pi + rI)\,dt + \Sigma\begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}\right) - Cdt \\ &= \left(W\omega^T\pi + Wr - C\right)dt + W\omega^T\Sigma\begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix} \end{aligned} \tag{12}$$

We also denote the 3×1 vector of state variables g_c, g_d, $\widehat{x}$, and W as G. The SDE that G obeys follows immediately from (3), (4), and (6):

$$dG = \begin{pmatrix} dg_c \\ dg_d \\ d\widehat{x} \end{pmatrix} = \widehat{\mu}_G dt + \Sigma_G\begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix} \tag{13}$$

with obvious definitions for $\widehat{\mu}_G$ and Σ_G.

Duffie & Epstein (1992a) prove that the Bellman equation for optimal consumption-portfolio process is

$$\sup_{C,\omega} D^{(C,\omega)}J\,(G, W, t) + f(C, J) = 0 \tag{14}$$

where

$$D^{(C,\omega)}J\left(G,W,t\right)=J_t+\widehat{\mu}_G^T J_G+\left(W\omega^T\pi+Wr-C\right)J_W+\frac{1}{2}tr(\Upsilon)$$

with

$$\Upsilon=\begin{pmatrix}\Sigma_G\\ W\omega^T\Sigma\end{pmatrix}^T\begin{pmatrix}J_{GG} & J_{GW}\\ J_{WG} & J_{WW}\end{pmatrix}\begin{pmatrix}\Sigma_G\\ W\omega^T\Sigma\end{pmatrix}$$

The Bellman optimality condition (14) implies that for a given consumption process the optimal differential utility satisfies the following PDE:

$$J_t+\widehat{\mu}_G^T J_G+\frac{1}{2dt}dG^T J_{GG^T}dG+f \;=\; 0 \tag{15}$$

$$J(\infty,G) \;=\; 0^2 \tag{16}$$

In general, this PDE does not have an analytical solution. To this end, we follow Campbell & Viceira (2002) and Zhu (2006) and use a log-linear approximation of the normalized aggregator:

$$f\approx h(1-\gamma)J[\ln C-\frac{1}{1-\gamma}\ln J+H] \tag{17}$$

where h is the long-term mean of consumption to wealth ratio (see Apprendix for details of the derivation). There are two shortcomings of the proposed approximation. First, the elasticity of intertemproal substitiution ψ only appears in H, and has no impact on pricing. This shortcoming can be overcome by using higher-order terms but at the cost of losing the closed form expression for the utility function. Second, dividend growth has no direct impact on the utility function and on pricing results as we show below. The only impact of dividend growth is through its effect on the optimal estimate of the conditional mean of consumption growth due to the nature of agents' learning process.

Inserting approximation (17) into PDE (15), we look for a solution, $J(t, g, d, \widehat{x})$, of the PDE in an exponential-affine form:

$$J(t, g, d, \widehat{x}) = \exp(\xi_{0t} + \xi_{1t} g_t + \xi_{2t} d_t + \xi_{3t} \widehat{x}_t) \tag{18}$$

where $g = \ln C$ and $d = g_d$.

Substitution of this function into (15) leads to an affine function of the state variables, G, that must evaluate to zero for arbitrary values of the state variables. This situation is only possible if coefficients the state variables are all zero, which gives rise to a system of four ordinary differential equations (ODEs) for coefficients ξ_0, ξ_1, ξ_2, and ξ_3. Subject to appropriate initial condition on the lifetime utility function, the ODEs have the following solutions (see Appendix for details):

$$\xi_1 = 1 - \gamma \tag{19}$$

$$\xi_2 = 0 \tag{20}$$

$$\xi_3 = \frac{1 - \gamma}{h + \rho} \tag{21}$$

As we point out above, dividend growth does not have a *direct* effect on either the utility function or the pricing of assets as $\xi_2 = 0$.

2.3 Pricing Kernel and Term Structure of Interest Rates

Duffie & Epstein (1992a) show that the stochastic discount factor π is given by the following expression:

$$\pi_t = \exp\left(\int_0^t f_J ds\right) f_C \tag{22}$$

Appling Ito's lemma to (22) and using the expressions for the normalized aggregator approximation in (17) and the state processes in (3), (4), and (6), we obtain the process for the pricing kernel, π_t:

$$\frac{d\pi}{\pi} = -\left[h + \mu_g + \widehat{x} + (\xi_1 - 1)\,V + \xi_3\sqrt{V}\Sigma_{1x}\right] dt + \begin{pmatrix} (\xi_1 - 1)\sqrt{V} + \xi_3\Sigma_{1x} \\ \xi_3\Sigma_{2x} \end{pmatrix}^T \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}, \tag{23}$$

We derive expressions for Σ_{1x} and Σ_{2x} in Appendix.

The real short term risk-free rate, r, is given by the drift in expression (23):

$$\begin{aligned} r_t &= -E\left(\frac{d\pi}{\pi}\right)/dt \\ &= \Omega + \widehat{x} \end{aligned} \tag{24}$$

where $\Omega \equiv h + \mu_g + (\xi_1 - 1)\,V + \xi_3\sqrt{V}\Sigma_{1x}$. For the case of an infinitely lived agent that we consider in our model $\xi_1 = 1 - \gamma$, which implies that $\Omega \equiv h + \mu_g - \gamma V + \xi_3\sqrt{V}\Sigma_{1x}$ (see Appendix for details).

Equation (24) shows that, the real short rate is only a function of time t and the agents' posterior estimate of the conditional mean of consumption growth, $\widehat{x}$. Ito's lemma applied to equation (24) implies that

$$\begin{aligned} dr &= \dot{\Omega} dt + d\widehat{x} \\ &= \left(\dot{\Omega} - \rho\widehat{x}\right) dt + \Sigma_x \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix} \end{aligned} \tag{25}$$

where $\dot{\Omega}$ is the first order derivative of Ω with respect to t.

Now we derive the price of a zero-coupon bond using the pricing kernel in (23). Price, P, of any asset must satisfy the fundamental pricing equation:

$$E\left(\frac{dP}{P}\right) = rdt - E\left(\frac{dP}{P}\frac{d\pi}{\pi}\right) \tag{26}$$

Given the affine nature of both the SDE and the short rate, equations (23), (24), and (26) imply that the solution for the bond price has an exponential-affine form as a function of posterior mean of conditional consumption growth, $\widehat{x}$ (see Appendix for detailed arguments). Since $\widehat{x}$ is affine (see (24)) in the short rate, r, then bond price itself is an affine function of the short rate.

Theorem 2 *The price of a zero-coupon bond with face value of* \$1 *and time to maturity* $\tau = s - t$ *is given by an exponential-affine function of the short rate:*

$$P = \exp(A(\tau) + B(\tau)\, r).$$

where functions $A(\tau)$ *and* $B(\tau)$ *are given by*

$$\begin{aligned} A(\tau) &= \Omega(B_t + \tau) + \frac{b}{4\rho^3}(e^{-2\rho\tau} - 4e^{-\rho\tau} - 2\rho\tau + 3) + \frac{\lambda}{\rho^2}(e^{-\rho\tau} + \rho\tau - 1) \\ B(\tau) &= -\frac{1}{\rho}\left(1 - e^{-\rho(s-t)}\right) \end{aligned}$$

Proof: see Appendix.

Then the yield is simply

$$y = -\frac{1}{s-t}(A + Br) \tag{27}$$

Since Ω and λ are linear function of ξ_1 and ξ_3, they are constants as well. So we can move them out of the integral when we are calculating A in the yield. Therefore, we get a much simpler expression of A :

$$
\begin{aligned}
A &= (-\Omega B)\,|_t^s - b\int_t^s \frac{B^2}{2}du + \Omega\int_t^s du - \lambda\int_t^s Bdu \\
&= \Omega(B_t+\tau) + \frac{b}{4\rho^3}(e^{-2\rho\tau} - 4e^{-\rho\tau} - 2\rho\tau + 3) + \frac{\lambda}{\rho^2}(e^{-\rho\tau} + \rho\tau - 1).
\end{aligned}
$$

The yield turns out to be a linear combination of the correlation between $\widehat{x}$ and g, a, and the variance of $\widehat{x}$, b :

$$
\begin{aligned}
y &= -\frac{1}{\tau}\left(A + B_t r\right) \\
&= -\frac{1}{\tau}\left(\Omega(2B_t+\tau) + \frac{b}{4\rho^3}(e^{-2\rho\tau} - 4e^{-\rho\tau} - 2\rho\tau + 3) + \frac{\lambda}{\rho^2}(e^{-\rho\tau} + \rho\tau - 1) + B_t x\right) \\
&= m_0 + m_1 b + m_2 a,
\end{aligned}
$$

and the conditional variance and unconditional variance of the yield are the same as those of the full model:

$$
\begin{aligned}
var(y_{t\to s}|\widehat{x}_{t-1}) &= \frac{B_t^2 b}{s^2}, \\
var(y_{t\to s}|\widehat{x}_0) &= \frac{1}{2\rho s^2}B_t^2 b\left(1 - e^{-2\rho t}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
a &= \sqrt{V}\,(C1 + \rho_{gd} C2) \\
b &= C1^2 + 2\rho_{gd} C1 C2 + C2^2 \\
\tau &= s - t \\
m_0 &= -\frac{1}{\tau}\left((h + \mu_g + (\xi_1 - 1)V)\,(2B_t + \tau) + B_t x\right) \\
m_1 &= -\frac{1}{\tau}\left(\frac{1}{4\rho^3}(e^{-2\rho\tau} - 4e^{-\rho\tau} - 2\rho\tau + 3) + \frac{1}{\rho^2}\xi_3(e^{-\rho\tau} + \rho\tau - 1)\right), \\
m_2 &= -\frac{1}{\tau}\left(\xi_3(2B_t + \tau) + \frac{1}{\rho^2}(\xi_1 - 1)(e^{-\rho\tau} + \rho\tau - 1)\right).
\end{aligned}
$$

2.4 Model Implication

To calibrate the model, we pick parameters values from Constantinides & Ghosh (2011). Since our model is continuous-time, the mean-reverting speed of x is different from that in the discreted-time model. Therefore, we set $\rho = -\log\left(\rho^{CG}\right)$ so that the first moment of x in our model matches CG's estimation.

As one can see from equation, the yield is a linear combination of a, the value of the covariance between unobserved long-run risk factor $\widehat{x}$ and consumption growth g, and b, the variance of $\widehat{x}$. Here, we start from examining the properties of a and b.

Value of a is determined by all 3 correlations between state variables ρ_{cx}, ρ_{dx}, and ρ_{cd}. In figure, we plot value of a for different combinations of values of ρ_{cx}, ρ_{dx}, and ρ_{cd}. For our parameter values, ρ_{cx}'s impact on a changes with ρ_{dx}, and ρ_{cd}. In general, a is increasing on ρ_{cx}. However, when ρ_{cd} is close to zero, a tends to be flatter. And, for large positive value of ρ_{cd}, a can become a monotonically decreasing function on ρ_{cx}. This pattern is weaker when ρ_{dx} goes from negative to zero, and for positive ρ_{dx}, a is always increasing over ρ_{cx}. a over ρ_{dx} in general is hump shaped. When consumption growth and dividend growth are largely negatively correlated, because of positive definite restriction, we only get right half of the curve, which is decreasing. When ρ_{cd} is large and positive, we get the left half of the curve, which is increasing. The impact of ρ_{cd} is similar to that of ρ_{dx}. When ρ_{dx} is large and positive, increasing ρ_{cd} increases a, and the opposite holds for negative ρ_{dx}.

Figure demonstrates the impact of change in those key correlations on b. The derivative of b over ρ_{cx} is always positive, which means the higher the correlation between consumption growth and x, the higher the variance of $\widehat{x}$. The derivative over ρ_{dx} is negative when ρ_{dx} is close to -1. It's increasing and later becomes positive when ρ_{dx} is large and positive. ρ_{cd}'s impact on b is similar to that of ρ_{dx}. Basically, b is a parabola over ρ_{cd}, which is first decreasing and later becomes increasing. Also, because of positive definite restriction, for some extreme values of ρ_{dx} and ρ_{cx}, we only have part of the parabola. For instance, when ρ_{dx} is large and positive, we can only get left half of the curve, which is in general decreasing on ρ_{cd}.

In this simplified model, both ξ_1 and ξ_3 are all negative constant, so the total risk premium $-\lambda$ is simply a linear combination of a and b. In figure, we examine the behavior of the total risk premium. We find that the total risk premium is always increasing over ρ_{cx} for both a and b are increasing over ρ_{cx}. Since $h + \rho$ is much smaller than 1, $-\xi_3$ is significantly larger than $-\xi_1$. Because of this, $-\lambda$ is in general dominated by b. Therefore, change in ρ_{dx} affects the risk premium in the similar way as it affects b. And the same thing holds for ρ_{cd}.

To analyze the yield, we still need to examine the properties of loading on a, m_2, and that on b, m_2. In table, we show values of m_1 and m_2 for different parameters values. For $\gamma = 2$, m_1 is always positive. Intuitively, for larger the variance of $\widehat{x}$, investors require higher risk premium, and therefore the yield on a zero coupon bond is higher. m_2 is negative for small τ, which implies that for a zero coupon bond with small time-to-maturity, the yield is decreasing over the covariance between the consumption growth and $\widehat{x}$. This is due to the fact that both Ω and λ contains a. While the second term of m_2 is always positive, the first term, which is succeed from Ω, is negative. Intuitively, for a larger value of a, although the risk premium is higher, the expectation of the risk free rate is lower. And, as τ increases, the first term becomes less negative, which makes m_2 positive for a large τ. For $\gamma = 2$, when

$\tau > 15$, m_2 becomes positive. That is for a large time-to-maturity, negative impact of the expectation on risk free rate is overcome by the positive effect of the risk premium.

Together with our analysis of a and b, we can now say something on the impact of changes in key correlations on the yield. Figure show yields on 1-year, 5-year, 10-year, and 20-year zero coupon bonds for different value of ρ_{cx}, ρ_{dx}, and ρ_{cd}. In general, this model produces a downward-sloping real yield curve. The real yield of a 1-year zero coupon bond is between 2.8% and 2.9%, and that of a 20-year zero coupon bond is between 1.5% and 1.6%. The impact of ρ_{cx} is the most trivial one. Large ρ_{cx} leads to larger a and b, which makes the yield higher for large τ. When τ is small, since m_2 is negative, the negative impact of a compensates that of b, and leads to a flatter yield curve. ρ_{dx}'s effect on the yield is similar to that of ρ_{dx} on b. For large τ this is because the yield's loading on b is much larger than that on a. For small τ, since m_2 is negative, impact of changing a works in the same direction as that of changing b. When ρ_{dx} is negative and large, increasing ρ_{dx} moves the whole yield curve downwards. The opposite holds for positive ρ_{dx}. ρ_{dx} also affects the impact of ρ_{cd}. If ρ_{dx} is around zero, the yield curve first moves downwards then upwards when ρ_{cd} goes from negative to positive. For large positive ρ_{dx}, the yield is decreasing over ρ_{cd}. And, the opposite holds for large negative ρ_{dx}.

Since both conditional variance and unconditional variance are only function of b, the impact of changes in key correlations on conditional variance and that on unconditional variance are again the same as that on b. Figure shows the unconditional variance of the yield when $\rho_{cx} = \rho_{dx} = \rho_{cd} = 0$. To produce figure, we set $s = 21$ and move t from 1 to $(s - 1)$. The unconditional variance first of all increases, and in our case, for time-to-maturity larger than 12, it's decreasing over τ. This is because, although B^2 is increasing over τ, $(1 - e^{-2\rho t})$ shrinks as τ goes from 1 to $(s - 1)$. However, we don't see such phenomenon in Figure if we assume period 0 is in the long past, that is t starts from a really large number and $s = t + \tau$. In this case, B^2 is still increasing over τ, and $(1 - e^{-2\rho t})$ is really flat over τ. Therefore, the unconditional variance is monotonically increasing over τ. The conditional variance of the yield is always increasing over τ. And, as being mentioned above, this is because B^2 increases over τ.

2.5 Comparison of the Incomplete and Complete Information Models

In this section, we examine the differences between our learning-based model with the complete-information model. The purpose of this section is to investigate the impact of introducing learning into the LRR model on both the risk premium and the yield curve. To simplify discussion, we still use the simplified model where the agent is infinitely lived.

The key difference between two types of model is that they have different expressions of a and b. For the complete-information model, a and b have much simpler expression:

$$
\begin{aligned}
a_{nl} &= \varphi_e V \rho_{cx}, \\
b_{nl} &= \varphi_e^2 V.
\end{aligned}
$$

Both Ω and λ are derived via universal approach, so both model with learning and that without learning share the same expression of the risk premium and the mean of risk-free rate.

The difference between a and a_{nl} is simply S. Figure demonstrates the impact of changes in key correlations on $\Delta a = a - a_{nl}$. Δa is monotonically decreasing over ρ_{cx}. When $\rho_{cx} = \rho_{dx}(\rho_{cd}) = 0$, Δa is a hump-shaped function over $\rho_{cd}(\rho_{dx})$. When ρ_{dx} (ρ_{cd}) is negative (positive), we tend to get the right (left) half of the curve, which is decreasing (increasing) over ρ_{cd} (ρ_{dx}). This pattern is weak when ρ_{cx} is large and positive. The interesting finding is that, b is smaller than b_{nl}. In Figure, we plot the difference between b and b_{nl} against different combinations of values of key correlations. As we can see, $\Delta b = b - b_{nl}$ shrinks as ρ_{cx} goes from -1 to 1. The impact of change in ρ_{cd} (ρ_{dx}) is similar to that on $-\Delta a$.Since the total risk premium has more loads on b (b_{nl}), $\Delta\lambda = -\lambda - (-\lambda_{nl})$ is dominated by Δb.

Therefore, the impact of change in $\rho_{cx}(\rho_{dx}$ or $\rho_{cd})$ on $\Delta\lambda$ is similar to that on Δb.

Now we can move to the difference in the yield Δy between two models. Since m_0, m_1, and m_2 are the same for both models, we can write Δy as

$$\begin{aligned}\Delta y &= y - y_{nl} \\ &= m_1 \Delta b + m_2 \Delta a.\end{aligned}$$

From the analysis above we know that Δa is similar in shape to $-\Delta b$. Also, we have mentioned that for small τ, m_2 is negative, and when τ is large, m_2 is much smaller than m_1. Having these two facts, together with the expression of Δy above, we conclude that Δy has the similar shape to Δb.

3 Appendix

3.1 The agent's posterior estimate of the conditional mean of consumption growth, $\widehat{x}_t$.

Because the repersentative agent cannot observe the true conditional mean of consumption growth, $\widehat{x}_t$, she seeks to extract new information about the value of x_t by observing both consumption and dividend growth. To this end, we write the (two-dimensional) measurement (observation) process and the latent process together as

$$dG^0 = \mu_G dt + \Omega dw \tag{28}$$

where the complete state vector (including both the observed and the unobserved state variables), G^0, its mean, μ_G, the volatility matrix, Ω, are given by

$$
\begin{aligned}
G^0 &= \begin{pmatrix} g_c \\ g_d \\ x \end{pmatrix} \\
\mu_G &= \begin{pmatrix} \mu_c + x - V/2 \\ \mu_d + \phi x - \varphi_d^2 V/2 \\ -\rho x \end{pmatrix} \equiv \begin{pmatrix} \underset{2\times 1}{\varphi_1} \\ \varphi \end{pmatrix} && (29) \\
\Omega &= \sqrt{V} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \varphi_d & 0 \\ 0 & 0 & \varphi_e \end{pmatrix} \equiv \begin{pmatrix} \underset{2\times 3}{\psi_1} \\ \psi \end{pmatrix} && (30)
\end{aligned}
$$

and $dw = \begin{pmatrix} dw_c \\ dw_d \\ dw_x \end{pmatrix}$ is a vector of Brownian motion increments with correlation matrix

$\Lambda = \begin{pmatrix} 1 & \rho_{cd} & \rho_{cx} \\ \rho_{cd} & 1 & \rho_{dx} \\ \rho_{cx} & \rho_{dx} & 1 \end{pmatrix}$.

If we Cholesky decompose the correlation matrix, we can simplify expression (28) for the dynamics of the state variables as follows:

$$
dG^0 = \mu_G dt + \Omega \Sigma dw \tag{31}
$$

where Σ is Cholesky decomposition of Λ, i.e.,

$$
\Sigma = \begin{pmatrix} 1 & 0 & 0 \\ \rho_{cd} & \sqrt{1-\rho_{cd}^2} & 0 \\ \rho_{cx} & \frac{1}{\sqrt{1-\rho_{cd}^2}} \left(\rho_{dx} - \rho_{cd}\rho_{cx}\right) & \sqrt{\frac{1}{\rho_{cd}^2 - 1} \left(\rho_{dx} - \rho_{cd}\rho_{cx}\right)^2 - \rho_{cx}^2 + 1} \end{pmatrix}
$$

Now, however, dw in (31) is the vector of independent Brownian motion increments. For the Cholesky decomposition of Λ to make sense, we need to impose the positive-definite restriction on Λ, which in this case amounts to the following restriction on correlations:

$$\rho_{cd}^2 + \rho_{cx}^2 + \rho_{dx}^2 - 2\rho_{cd}\rho_{cx}\rho_{dx} < 1$$

To summurize our notation we now write the measurement process as

$$d \begin{bmatrix} dg_c \\ dg_d \end{bmatrix} = \varphi_1(x)dt + \psi_1 \Sigma dw, \tag{32}$$

and the latent process as

$$dx = \varphi(y, x, t)dt + \psi(y, t)\Sigma dw \tag{33}$$

where $dw = (dw_c, dw_d, dw_e)^T$ is the vector of intependent Brownian motion increments.

As the agent collects new observations on consumption and dividend growth, she updates her estimate $\widehat{x}_t$ of conditional mean of consumption growth according to the linear Kalman-Bucy filter:

$$d\widehat{x} = -\rho\widehat{x}dt + E_t\left(\left(\varphi_1 - \widehat{\varphi}_1\right)^T x + \psi\Lambda\psi_1^T\right)\left(\psi_1\Lambda\psi_1^T\right)^{-1}\left(d\begin{pmatrix} g_c \\ g_d \end{pmatrix} - \widehat{\varphi}_1 dt\right) \tag{34}$$

where $\widehat{\varphi}_1 = E_t\varphi_1$.

Given definitions in (29) and (30) as well as the definition of the correlation matrix Λ, we arrive at the following intermediate results:

$$\begin{aligned}
\varphi_1 - \widehat{\varphi}_1 &= \begin{pmatrix} x - \widehat{x} \\ \phi\,(x - \widehat{x}) \end{pmatrix} \\
E_t\left((\varphi_1 - \widehat{\varphi}_1)^T x\right) &= S \begin{pmatrix} 1 \\ \phi \end{pmatrix}^T, \text{ where } S = E_t\,(x_t - \widehat{x}_t)^2 \\
\left(\psi_1 \Lambda \psi_1^T\right)^{-1} &= \frac{1}{V\varphi_d^2\,(1-\rho_{cd}^2)} \begin{pmatrix} \varphi_d^2 & -\varphi_d\rho_{cd} \\ -\varphi_d\rho_{cd} & 1 \end{pmatrix} \\
\psi\Lambda\psi_1^T &= V\varphi_e \begin{pmatrix} \rho_{cx} & \varphi_d\rho_{dx} \end{pmatrix}
\end{aligned}$$

Substituting (32) and (33) into equation (34) we obtain the desired posterior estimate of the conditional mean of consumption growth, $\widehat{x}_t$:

$$\begin{aligned}
d\widehat{x} &= -\rho\widehat{x}dt + \\
&\quad \left[S_t \begin{bmatrix} 1 & \phi \end{bmatrix} + V\varphi_e \begin{bmatrix} \rho_{cx} & \varphi_d\rho_{dx} \end{bmatrix}\right] \\
&\quad \times \frac{1}{\varphi_d^2 V(1-\rho_{cd}^2)} \begin{bmatrix} \varphi_d^2 & -\varphi_d\rho_{cd} \\ -\varphi_d\rho_{cd} & 1 \end{bmatrix} \left(d \begin{bmatrix} g_c \\ g_d \end{bmatrix} - \widehat{\varphi}_1 dt \right) \\
&= -\rho\widehat{x}dt + \Sigma_x \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}
\end{aligned}$$

where

$$
\begin{aligned}
S_t &= E_t[(x_t - \widehat{x}_t)^2] && (35) \\
\Sigma_x &= \begin{bmatrix} \frac{1}{1-\rho_{cd}^2}\left[S\frac{(\varphi_d - \phi\rho_{cd})}{\varphi_d\sqrt{V}} + \varphi_e\sqrt{V}(\rho_{cx} - \rho_{dx}\rho_{cd})\right] \\ \frac{1}{1-\rho_{cd}^2}\left[S\frac{(\phi - \varphi_d\rho_{cd})}{\varphi_d\sqrt{V}} + \varphi_e\sqrt{V}(\rho_{dx} - \rho_{cx}\rho_{cd})\right] \end{bmatrix}^T && (36) \\
\begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix} &= \begin{bmatrix} \frac{x-\widehat{x}}{\sqrt{V}}dt + dw_c \\ dw_d + \frac{\phi(x-\widehat{x})}{\varphi_d\sqrt{V}}dt \end{bmatrix} = \begin{bmatrix} \frac{x-\widehat{x}}{\sqrt{V}}dt + dw_c \\ \frac{\phi(x-\widehat{x})}{\varphi_d\sqrt{V}}dt + \rho_{cd}dw_c + \sqrt{1-\rho_{cd}^2}dw_d^{\perp} \end{bmatrix}. && (37)
\end{aligned}
$$

where $Edw_d^{\perp}dw_c = 0$. Brownian motion increments in (37) denote the unexpected component of the conditional mean of consumption growth conditional on agents' information set.

Posterior variance, S_t, in equation (35) ia a solution of a deterministic ODE, which follows from applying Ito's lemma to the definition of the variance in (35) and taking expectations:

$$
\begin{aligned}
dS &= Ed\left(x - \widehat{x}\right)^2 \\
&= E\left[2\left(x - \widehat{x}\right)d\left(x - \widehat{x}\right) + \left[d\left(x - \widehat{x}\right)\right]^2\right] \\
&= \left\{-\alpha S^2 - 2(\rho + \xi)S + \Gamma^2\right\}dt, && (38)
\end{aligned}
$$

where

$$
\begin{aligned}
\alpha &= \frac{\phi^2 + \varphi_d^2 - 2\phi\varphi_d\rho_{cd}}{V\varphi_d^2\left(1 - \rho_{cd}^2\right)} \\
\xi &= \frac{\varphi_e}{\varphi_d\left(1 - \rho_{cd}^2\right)}\left(\phi\left(\rho_{dx} - \rho_{cd}\rho_{cx}\right) + \varphi_d\left(\rho_{cx} - \rho_{cd}\rho_{dx}\right)\right) \\
\Gamma^2 &= V\varphi_e^2\left(1 - \frac{1}{1-\rho_{cd}^2}\left(\rho_{dx} - \rho_{cd}\rho_{cx}\right)^2 - \rho_{cx}^2\right).
\end{aligned}
$$

We do all analysis on the assumption of the steady state, i.e., $\frac{dS}{dt} = 0$. In this case, $S = \frac{(\rho+\xi)+\sqrt{(\rho+\xi)^2+\alpha\Gamma^2}}{\alpha}$.

3.2 Log-linear approximation of f

Porteus-Kreps aggregator is defined in (2). For further discussion in this section we rewrite the expression for the aggregator by defining the following function of consumption and differential utility:

$$\Psi(C,J) = \left(\frac{C}{((1-\gamma)J)^{\frac{1}{1-\gamma}}} \right)^{1-\frac{1}{\psi}} \tag{39}$$

With this definition the normailzed aggregator, f, has the form:

$$f = \frac{\beta}{1-\frac{1}{\psi}}(1-\gamma)J\left[\Psi - 1\right] \tag{40}$$

Solution of the Bellman equation (14) (**see P.129 Campbell and Viceira 2001**) has the following form:

$$J = H(t,G)\frac{W^{1-\gamma}}{1-\gamma} \tag{41}$$

The envelope condition requires that

$$f_C = J_W \tag{42}$$

The form of the Bellman solution in (41) and the first-order optimality condition (42) imply the following restriction:

$$\beta\frac{1-\gamma}{1-\frac{1}{\psi}}J\Psi_C = (1-\gamma)J/W \tag{43}$$

Using the fact that $\Psi_C = (1-\frac{1}{\psi})\Psi/C$ and inserting it into the envelope condition (43) we have

$$\beta\Psi = e^{c-w} \tag{44}$$

Next, we expand the consumption-wealth ratio into a Taylor series around its unconditional mean $c_0 - w_0$. As long as the consumption wealth ratio does not vary a lot from its long-term mean, the log-linear approximation of $\beta\Psi$ given below will remain approximately valid:

$$\beta\Psi = h_0 + h \ln \beta\Psi \tag{45}$$

where

$$\begin{aligned} h &= e^{c_0 - w_0} &(46)\\ h_0 &= e^{c_0 - w_0}[1 - (c_0 - w_0)] &(47) \end{aligned}$$

With this in mind, the approximation of the normalized aggregator has the form:

$$\begin{aligned} f &\approx \frac{1-\gamma}{1-\frac{1}{\psi}} J[h_0 + h\ln\beta + h\ln\Psi - \beta] \\ &= h(1-\gamma)J[\ln C - \frac{1}{1-\gamma}\ln J + H] \end{aligned}$$

where

$$H = \frac{(h_0 + h\ln\beta - \beta)}{h(1-\frac{1}{\psi})} - \frac{1}{1-\gamma}\ln(1-\gamma) \tag{48}$$

3.3 Derivation of the Value Function, J

Once we substitute the log-linear approximation (17) into the PDE (15), it becomes a parabolic PDE with coefficients affine in state variables. It is natural to look for a solution, $J(t, g, d, \widehat{x})$, of (15) in an exponential affine form:

$$J(t, g, d, x) = \exp(\xi_{0t} + \xi_{1t} g_t + \xi_{2t} d_t + \xi_{3t}\widehat{x}_t) \tag{49}$$

Upon the substitution of (49) into (15) the PDE becomes an identity that must hold for arbitrary values of state variables (symbol ∂_t denotes the partial derivative with respect to time variable, t):

$$\partial_t \xi_{0t} + \partial_t \xi_{1t} g + \partial_t \xi_{2t} d + \partial_t \xi_{3t} \widehat{x} + \widehat{\mu}_G^T \begin{pmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \end{pmatrix} + \frac{1}{2dt} tr \left(dG dG^T \frac{J_{GG^T}}{J} \right) + \frac{f}{J} = 0 \qquad (50)$$

where the expressions for the mean of the state vector, $\widehat{\mu}_G^T$, and the aggregator, f, immediately follow from (29), (17), and (49):

$$\begin{aligned} \widehat{\mu}_G^T &= \begin{pmatrix} \mu_c + \widehat{x} - V/2 \\ \mu_d + \phi \widehat{x} - \varphi_d^2 V/2 \\ -\rho \widehat{x} \end{pmatrix}^T = \begin{pmatrix} \mu_c - V/2 \\ \mu_d - \varphi_d^2 V/2 \\ 0 \end{pmatrix}^T + G^T \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & \phi \\ 0 & 0 & -\rho \end{pmatrix}^T \\ \frac{f}{J} &= h(1-\gamma) \left[\ln C - \frac{1}{1-\gamma} \ln J + H \right] = h(1-\gamma) \left[H - \frac{\xi_{0t}}{1-\gamma} + G^T \begin{pmatrix} 1 - \frac{\xi_{1t}}{1-\gamma} \\ -\frac{\xi_{2t}}{1-\gamma} \\ -\frac{\xi_{3t}}{1-\gamma} \end{pmatrix} \right] \end{aligned}$$

Because identity (50) must hold for all values of state variables, coefficients on the state variables must all be zero.

Collecting terms containing the state variable vector we find that

$$G^T \left[\partial_t \begin{pmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & \phi \\ 0 & 0 & -\rho \end{pmatrix}^T \begin{pmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \end{pmatrix} + h(1-\gamma) \begin{pmatrix} 1 - \frac{\xi_{1t}}{1-\gamma} \\ -\frac{\xi_{2t}}{1-\gamma} \\ -\frac{\xi_{3t}}{1-\gamma} \end{pmatrix} \right] = 0$$

The above identity holds if and only if the following conditions are satisfied:

$$\begin{aligned} \partial_t \xi_{1t} + h(1-\gamma) - h\xi_{1t} &= 0 \\ \partial_t \xi_{2t} - h\xi_{2t} &= 0 \\ \partial_t \xi_{3t} + \xi_{1t} + \phi \xi_{2t} - (\rho + h)\xi_{3t} &= 0 \end{aligned} \tag{51}$$

subject to termial condition (16), i.e.,

$$\xi_{1t}(\infty) = \xi_{2t}(\infty) = \xi_{3t}(\infty) = 0$$

Since ξ_0 does not affect any further results, we do not attempt to solve for ξ_0 and ignore it in our further discussions. This is a system of joint ODEs for coefficients ξ_{1t}, ξ_{2t}, and ξ_{3t}.

Solving the system for an infinitely lived agent, we finally have

$$\begin{aligned} \xi_1 &= 1-\gamma \\ \xi_2 &= 0 \\ \xi_3 &= \frac{1-\gamma}{h+\rho} \end{aligned}$$

3.4 Derivation of the Pricing Kernel in (23)

The normalized Porteus-Kreps aggregator has the following approximate form:

$$f \approx h(1-\gamma)J[\ln C - \frac{1}{1-\gamma}\ln J + H]$$

Duffie & Epstein (1992b) show that the SDF can be expressed in terms of the aggregator as in (22). Applying Ito's lemma to (22) we have

$$
\begin{aligned}
\frac{d\pi}{\pi} &= f_J dt + \frac{df_C}{f_C} \\
&= \left(\frac{f}{J} - h\right) dt - \frac{dC}{C} + \frac{dJ}{J} + \frac{(dC)^2}{C^2} - \frac{dJ}{J}\frac{dC}{C}
\end{aligned} \tag{52}
$$

Recall that the value function, $J(t, g, d, \widehat{x})$, and the consumption process, C, obey the following SDEs:

$$
\frac{dC}{C} = (\mu_g + \widehat{x})dt + \sqrt{V} dw_c^* \tag{53}
$$

$$
dJ = -f dt + J_C \sqrt{V} C dw_c^* + J_d \varphi_d \sqrt{V} dw_d^* + J_{\widehat{x}} \Sigma_x \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix} \tag{54}
$$

Also, from (18) we have

$$
\begin{aligned}
\frac{J_C}{J} &= \frac{\xi_1}{C} \\
\frac{J_d}{J} &= \xi_2 = 0 \\
\frac{J_{\widehat{x}}}{J} &= \xi_3
\end{aligned}
$$

Combining the above results we have the following expression for the pricing kernel:

$$
\frac{d\pi}{\pi} = -\left[h + \mu_g + \widehat{x} + (\xi_1 - 1) V + \xi_3 \sqrt{V} \Sigma_{1x}\right] dt + \begin{pmatrix} (\xi_1 - 1)\sqrt{V} + \xi_3 \Sigma_{1x} \\ \xi_3 \Sigma_{2x} \end{pmatrix}^T \begin{bmatrix} dw_c^* \\ dw_d^* \end{bmatrix}
$$

where Σ_{1x} and Σ_{2x} are the two components of 1×2 vector Σ_x.

3.5 Derivation of zero coupon bond price

By the definition of the state-price process

$$P(t,s)=\frac{E_t\pi_s}{\pi_t}=\frac{E_t\left(f_C(s)\exp(\int_t^s f_J(u)du)\right)}{f_C(t)} \tag{55}$$

Using the expression for the value function () and differentiating () we have that

$$\begin{aligned} f_C &= h(1-\gamma)\exp(\xi_{0t}+(\xi_{1t}-1)\ln C_t+\xi_{3t}\widehat{x}_t) && (56)\\ f_J(t) &= -h\left(\xi_{0t}+(\gamma-1+\xi_{1t})\ln C_t+\xi_{3t}\widehat{x}_t+1+(\gamma-1)H\right) && (57)\end{aligned}$$

Upon inserting f_C and f_J into the expression for the bond price (55) the latter takes the following form:

$$P(t,s)=E_t\exp\left[\begin{array}{c} -(\ln C_s-\ln C_t)+(\xi_{0s}-\xi_{0t})+(\xi_{1s}\ln C_s-\xi_{1t}\ln C_t)+\\ (\xi_{3s}x_s-\xi_{3t}\widehat{x}_t)+\int_t^s f_J du\end{array}\right] \tag{58}$$

Given the dynamics of the consumption process in (3) we can write that

$$\xi_{1s}\ln C_s=\xi_{1s}\left(\ln C_t+F\left(\widehat{x}\right)\right) \tag{59}$$

where function F is only a funtion of $\widehat{x}$ and τ ($\tau=s-t$). Further, with the help of (59) the integral in (55) simplifies as follows:

$$
\begin{aligned}
\int_t^s f_J du &= -h \int_t^s \left(\xi_{0u} + (\gamma - 1 + \xi_{1u}) \ln C_u + \xi_{3u}\widehat{x}_u + 1 + (\gamma - 1) H\right) du \\
&= \underbrace{-h \ln C_t \int_t^s (\gamma - 1 + \xi_{1u})\, du}_{\text{not a function of } \widehat{x}} \\
&\quad \underbrace{-h \int_t^s (\xi_{0u} + 1 + (\gamma - 1) H) - h \int_t^s \left((\gamma - 1 + \xi_{1u}) F(\widehat{x}) + \xi_{3u}\widehat{x}_u\right) du}_{\text{not a function of } C} \quad (60)
\end{aligned}
$$

It is clear now that the bond price in (58) is reduced to an expectation of the product of two processes completely separable in consumption and the latent state variable:

$$
P(t,s) = E_t\left(Z_C Z_x\right)
$$

The terms depending on consumption only in the expression for the bond price further simplify as follows:

$$
Z_C = \exp\left(\ln C_t \left[(\xi_{1s} - \xi_{1t}) - h \int_t^s (\gamma - 1 + \xi_{1u})\, du\right]\right) = 1 \quad (61)
$$

The last result is a direct consequence of (51).

To sum up, the zero-coupon bond price P is only function of x and τ:

$$
\begin{aligned}
P(t,s) &= E_t\left(Z_x\right) = P(\tau, \widehat{x}_t) \\
P(0, \widehat{x}_t) &= 1
\end{aligned}
$$

Since $r(t) = \Omega(t) + x(t)$, we can rewrite P as a function of r and τ, $P(\tau, r_t)$. The fundamental pricing equation then implies the following ODE:

$$
E\left(\frac{dP}{P}\right) = r dt - E\left(\frac{dP}{P}\frac{d\pi}{\pi}\right)
$$

Ito's lemma imlies that

We look for solution in exponential affine form

$$P = \exp(A(\tau) + B(\tau)\, r)$$

Plugging into the ODE, we get

$$\begin{aligned} A_t + rB_t + B\left(\overset{\bullet}{\Omega} - \rho x\right) + \frac{1}{2}B^2(C_1^2 + 2\rho_{gd}C_1C_2 + C_2^2) - r + B\lambda &= 0 \\ & s.t. \\ A\,(t = s) &= 0 \\ B\,(t = s) &= 0 \end{aligned}$$

where

$$\lambda = (\xi_1 - 1)\left(C_1\sqrt{V} + \rho_{gd}C_2\sqrt{V}\right) + \xi_3(C_1^2 + 2\rho_{gd}C_1C_2 + C_2^2).$$

Collecting the terms, we have

$$B_t - B\rho - 1 = 0 \tag{62}$$

$$A_t + \Omega B_t + B\overset{\bullet}{\Omega} + \frac{1}{2}B^2(C_1^2 + 2\rho_{gd}C_1C_2 + C_2^2) - \Omega + B\lambda = 0 \tag{63}$$

Solving ODE (**??**) and (62), we get the expression of B and A :

$$\begin{aligned} B &= -\frac{1}{\rho}\left(1 - e^{-\rho(s-t)}\right) \\ A &= -\left(m_0 t + m_1\frac{e^{ht}}{h} + m_2\frac{e^{(\rho+h)t}}{\rho+h} + m_3\frac{e^{\rho t}}{\rho} + m_4\frac{e^{(h+2\rho)t}}{h+2\rho} + m_5\frac{e^{2\rho t}}{2\rho}\right) + con \\ con &= m_0 s + m_1\frac{e^{hs}}{h} + m_2\frac{e^{(\rho+h)s}}{\rho+h} + m_3\frac{e^{\rho s}}{\rho} + m_4\frac{e^{(h+2\rho)s}}{h+2\rho} + m_5\frac{e^{2\rho s}}{2\rho} \end{aligned}$$

where

$$\begin{aligned}
m_0 &= -\frac{con_1}{\rho} + \frac{y}{2\rho^2} \\
m_1 &= -\frac{con_2}{\rho} e^{-hT} \\
m_2 &= \frac{con_2}{\rho} e^{-(\rho s + hT)} - \frac{con_3}{\rho} e^{-(h+\rho)T} \\
m_3 &= \frac{con_1}{\rho} e^{-\rho s} - \frac{y}{\rho^2} e^{-\rho s} \\
m_4 &= \frac{con_3}{\rho} e^{-(\rho s + (h+\rho)T)} \\
m_5 &= \frac{y}{2\rho^2} e^{-2\rho s} \\
y &= C_1^2 + 2\rho_{gd} C_1 C_2 + C_2^2 \\
z &= C_1 + \rho_{gd} C_2 \\
v &= \sqrt{V} \\
con_1 &= \rho h + \rho \mu_g - \gamma(vz + v^2\rho) + \frac{1-\gamma}{h+\rho}(\rho vz + y) \\
con_2 &= (\gamma - 1)\left(vz + v^2\rho + \frac{\rho vz + y}{\rho} + hv^2 + \frac{vzh}{\rho} \right) \\
con_3 &= \frac{(1-\gamma)h}{\rho}\left(\frac{\rho vz + y}{h+\rho} + vz \right)
\end{aligned}$$

References

Bansal, R., & Yaron, A. (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, *59*, 1481–1509.

Campbell, J., & Viceira, L. (2002). *Strategic asset allocation: portfolio choice for long-term investors*. Oxford University Press, USA.

Constantinides, G. M., & Ghosh, A. (2011). Asset pricing tests with long run risks in consumption growth. *The Review of Asset Pricing Studies*, *1*, 96–136.

Duffie, D., & Epstein, L. G. (1992a). Asset pricing with stochastic differential utility. *The Review of Financial Studies*, *5*, 411–436.

Duffie, D., & Epstein, L. G. (1992b). Stochastic differential utility. *Econometrica*, *60*, 353–394.

Epstein, L. G., & Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, *57*, 937–969.

Eraker, B. (2008). Affine general equilibrium models. *Management Science*, *54*, 2068–2080.

Zhu, Y. (2006). Dynamic volatility strategy with recursive utility. *Available at SSRN 890040*, .