Supplementary Materials for "Asset Pricing in Production Economies with Extrapolative Expectations"

1 Numerical solution

The standard value iteration algorithm with Chebyshev regression (e.g., Judd (1998) and KL (2010)) is used to solve the model numerically. Since the quantities in the economy are cointegrated with the aggregate productivity, A_t, quantities are first scaled by the aggregate productivity. In what follows, scaled variables are denoted by hats, and log values are denoted by lower case:

$$\begin{aligned}\hat{c}_t &= \log\left(\frac{C_t}{A_t}\right),\ \hat{k}_t = \log\left(\frac{K_t}{A_t}\right),\ \ \hat{\imath}_t = \log\left(\frac{I_t}{A_t}\right),;\\ \hat{y}_t &= \log\left(\frac{Y_t}{A_t}\right) = \log\left(A_t^{-\alpha}K_t^{\alpha}\right) = \alpha\hat{k}_t,\ \ a_t = \log\left(A_t\right).\end{aligned}$$

The social planner's problem is

$$V_t\left(K_t, \hat{\mu}_t, A_t\right) = \max_{C_t, I_t} \hat{\mathbf{E}}_t\left[\left(1-\beta\right)C_t^{\frac{1-\gamma}{\theta}} + \beta\left(\hat{\mathbf{E}}_t\left[V\left(K_{t+1}, \hat{\mu}_{t+1}, A_{t+1}\right)^{1-\gamma}\right]\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}$$

Since the problem is homogeneous, the value function can be re-defined

$$V_t\left(K_t, \hat{\mu}_t, A_t\right) \equiv \hat{V}_t\left(\frac{K_t}{A_t}, \hat{\mu}_t\right)\cdot A_t.$$

The maximization problem can be rewritten as

$$\hat{V}_t\left(\frac{K_t}{A_t}, \hat{\mu}_t\right) = \max_{C_t, I_t}\left[\left(1-\beta\right)\left(\frac{C_t}{A_t}\right)^{\frac{1-\gamma}{\theta}} + \beta\left(\hat{\mathbf{E}}_t\left[\hat{V}\left(\frac{K_{t+1}}{A_{t+1}}, \hat{\mu}_{t+1}\right)^{1-\gamma}\cdot\left[\frac{A_{t+1}}{A_t}\right]^{1-\gamma}\right]\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}$$

Take logarithm on both sides of the above equation

$$\begin{aligned} & \hat{v}_t\left(\hat{k}_t, \hat{\mu}_t\right) \equiv \log \hat{V}_t\left(\frac{K_t}{A_t}, \hat{\mu}_t\right) \\ = & \frac{\theta}{1-\gamma} \max_{C_t, I_t} \log \left\{\left[(1-\beta)\left(\frac{C_t}{A_t}\right)^{\frac{1-\gamma}{\theta}}+\beta\left(\hat{\mathbf{E}}_t\left[\hat{V}\left(\frac{K_{t+1}}{A_{t+1}}, \hat{\mu}_{t+1}\right)^{1-\gamma} \cdot\left[\frac{A_{t+1}}{A_t}\right]^{1-\gamma}\right]\right)^{\frac{1}{\theta}}\right]\right\} \\ = & \frac{\theta}{1-\gamma} \max_{\hat{c}_t, \hat{\imath}_t} \log \left\{\left[(1-\beta) e^{\frac{1-\gamma}{\theta} \hat{c}_t}+\beta\left(\hat{\mathbf{E}}_t\left[e^{\left[\hat{v}_{t+1}\left(\hat{k}_{t+1}, \hat{\mu}_{t+1}\right)+g_{A, t+1}\right](1-\gamma)}\right]\right)^{\frac{1}{\theta}}\right]\right\} \end{aligned}$$

Thus, the maximization problem is

$$\hat{v}_t\left(\hat{k}_t, \hat{\mu}_t\right)=\frac{\theta}{1-\gamma} \max_{\hat{\imath}_t} \log \left\{\left[(1-\beta) e^{\frac{1-\gamma}{\theta} \hat{c}_t}+\beta\left[\hat{\mathbf{E}}_t\left(e^{(1-\gamma) \cdot\left[v_{t+1}\left(\hat{k}_{t+1}, \hat{\mu}_{t+1}\right)+\left(\hat{\mu}_t+\sigma_A \hat{\epsilon}_{A, t+1}\right)\right]}\right)\right]^{\frac{1}{\theta}}\right]\right\}, \tag{1}$$

where $\hat{\epsilon}_{A,t+1}$ is standard normal under the individual's perception. The dynamics of the state variables are

$$\hat{\mu}_{+1}=(1-\rho-\tilde{\rho}) \mu_A+\rho \hat{\mu}_t+\tilde{\rho} g_{A, t+1},$$

and

$$\hat{k}_{t+1}=\log \left\{(1-\delta_K) \exp \left(\hat{k}_t\right)+\phi\left(\exp \left(\hat{\imath}_t-\hat{k}_t\right)\right) \exp \left(\hat{k}_t\right)\right\}-g_{A, t+1}$$

However, from the individual's perspective,

$$\hat{\mu}_{t+1}=(1-\rho-\tilde{\rho}) \mu_A+(\rho+\tilde{\rho}) \hat{\mu}_t+\tilde{\rho} \hat{\epsilon}_{A, t+1}, \tag{2}$$

and

$$\hat{k}_{t+1}=\log \left((1-\delta_k) \exp \left(\hat{k}_t\right)+\phi\left(\exp \left(\hat{\imath}_t-\hat{k}_t\right)\right) \exp \left(\hat{k}_t\right)\right)-\hat{\mu}_t-\sigma_A \hat{\epsilon}_{A, t+1}. \tag{3}$$

The last budget constraint is

$$\hat{\imath}_t \quad = \quad \log \left(\exp \left(\alpha \hat{k}_t\right)-\exp \left(\hat{c}_t\right)\right). \tag{4}$$

To solve the maximization problem numerically, the value function is parameterized with a 5th order Chebyshev orthogonal polynomial over a 6×6 Chebyshev grid for the state variables $\left(\hat{k}, \hat{\mu}\right)$. The value function iteration algorithm follows KL (2010). In particular, our algorithm assumes that the value function is $\Psi_n\left(\hat{k}, \hat{\mu}\right)$ at iteration n. For each grid point for the state variables $\left(\hat{k}, \hat{\mu}\right)$, a numerical optimizer is used to find the policy $(\hat{\imath}_t^*)$ that maximizes the value function

$$\hat{v}_t^*\left(\hat{k}_t,\hat{\mu}_t\right)=\frac{\theta}{1-\gamma}\max_{\hat{\imath}_t}\log\left\{\hat{\mathbf{E}}_t\left[(1-\delta)\,e^{\frac{1-\gamma}{\theta}\hat{c}_t}+\delta\left[\hat{\mathbf{E}}_t\left(e^{(1-\gamma)\cdot\left[\Psi_n\left(\hat{k}_{t+1},\hat{\mu}_{t+1}\right)+\left(\hat{\mu}_t+\sigma_A\hat{\epsilon}_{A,t+1}\right)\right]}\right)\right]^{\frac{1}{\theta}}\right]\right\},$$

where Gauss-Hermite quadrature with 5 nodes is used to approximate the expectations operator. The new value $\hat{v}^*$ is then regressed onto state variables $\hat{k}$, and $\hat{\mu}$ in order to update the coefficients of the polynomial for the value function. This way, the new Chebyshev polynomial $\Psi_{n+1}\left(\hat{k},\hat{\mu}\right)$ is obtained at iteration $n+1$.

After obtaining the value function and the corresponding policy function, other variables of interest can be calculated. For example, the perceived consumption growth rate is

$$\begin{aligned}\hat{\mathbf{E}}_t\left[\log\left(\frac{C_{t+1}}{C_t}\right)\right] &= \hat{\mathbf{E}}_t\log\left(\frac{\hat{C}_{t+1}}{\hat{C}_t}\frac{A_{t+1}}{A_t}\right)\\ &= \hat{\mathbf{E}}_t\left(\hat{c}_{t+1}-\hat{c}_t+\hat{\mu}_t+\sigma_A\hat{\epsilon}_{A,t+1}\right).\end{aligned}$$

The risk free rate can be computed by

$$\begin{aligned} r_{f,t} &= -\log\left[\hat{\mathbf{E}}_t\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\psi}}\left(\frac{V_{t+1}\left(K_{t+1},\hat{\mu}_{t+1},A_{t+1}\right)}{\left[\hat{\mathbf{E}}_t\left(V_{t+1}^{1-\gamma}\left(K_{t+1},\hat{\mu}_{t+1},A_{t+1}\right)\right)\right]^{\frac{1}{1-\gamma}}}\right)^{\frac{1}{\psi}-\gamma}\right)\right]\\ &= -\log\left[\hat{\mathbf{E}}_t\left(\beta\left(\frac{\hat{C}_{t+1}}{\hat{C}_t}\right)^{-\frac{1}{\psi}}\left(\frac{\hat{V}_{t+1}\left(K_{t+1},\hat{\mu}_{t+1},A_{t+1}\right)}{\left[\hat{\mathbf{E}}_t\left(\hat{V}_{t+1}^{1-\gamma}\left(K_{t+1},\hat{\mu}_{t+1},A_{t+1}\right)\left(\frac{A_{t+1}}{A_t}\right)^{1-\gamma}\right)\right]^{\frac{1}{1-\gamma}}}\right)^{\frac{1}{\psi}-\gamma}\left(\frac{A_{t+1}}{A_t}\right)^{-\gamma}\right)\right].\end{aligned}$$

Long-term bond prices can be calculated recursively by

$$\begin{aligned} P_{t,t+n} &= \hat{\mathbf{E}}_t M_{t,t+n}\\ &= \hat{\mathbf{E}}_t\left[M_{t,t+1}\hat{\mathbf{E}}_{t+1}\left(M_{t+1,t+n}\right)\right]\\ &= \hat{\mathbf{E}}_t\left[M_{t,t+1}P_{t+1,t+n}\right].\end{aligned}$$

and the long-term bond yield can be easily obtained from the long-term bond price. Lastly, to find Tobin's Q, let

$$\frac{\hat{I}_t}{K_t}\equiv\phi\left(\frac{I_t}{K_t}\right)\ \text{and}\ \hat{\phi}\left(\hat{I}_t,K_t\right)\equiv I_t-\phi\left(\frac{I_t}{K_t}\right)K_t=I_t-\hat{I}_t.$$

Then, Tobin's Q is

$$
\begin{aligned}
Q &= 1 + \frac{\partial\ \hat{\phi}\left(\hat{I}_t, K_t\right)}{\hat{I}_t} = \frac{\partial I_t}{\hat{I}_t} = \frac{\partial \phi^{-1}\left(\frac{\hat{I}_t}{K_t}\right)}{\frac{\hat{I}_t}{K_t}} \\
&= \frac{1}{\phi'\left(\phi^{-1}\left(\frac{\hat{I}_t}{K_t}\right)\right)} = \frac{1}{\phi'\left(\frac{I}{K}\right)}.
\end{aligned}
$$

2 Approximate analytical solution

This section analyzes an approximate model solution using log-linearization. The detailed solution is described in Section 2.1, followed by the summarization and discussion of the findings in Section 2.2.

2.1 Detailed solution

The model is based on Kaltenbrunner and Lochstoer (2010) (KL (2010) thereafter). The major deviation is the process for the perceived productivity growth, which the agent who makes consumption and investment decisions is based on.

$$\hat{\mu}_{t+1} = (1 - \rho - \tilde{\rho})\mu_A + (\rho + \tilde{\rho})\hat{\mu}_t + \tilde{\rho}\sigma_A \hat{\epsilon}_{At+1} \tag{5}$$

where $\hat{\epsilon}_{At+1}$ is the perceived shocks in productivity.

The objective productivity process is a random walk:

$$g_{At+1} = \mu_A + \sigma_A \epsilon_{At+1} \tag{6}$$

so the objective expected productivity growth is a constant μ_A, and ϵ_{At+1} is the i.i.d. shocks under objective measure.

2.1.1 Solution under subjective measure

In this section, the model is solved under agent's subjective measure with extrapolation bias. This provides solutions for the consumption and investment decisions, as well as asset prices. The next section studies the dynamics of quantities and asset prices under the objective measure.

Our approximation starts by detrending variables by aggregate productivity (A_t). Define $\hat{Y}_t = \frac{Y_t}{A_t}$, $\hat{K}_t = \frac{K_t}{A_t}$, so the production function becomes

$$Y_t = \hat{Y}_t A_t = A_t^{1-\alpha} \hat{K}_t^{\ \alpha} A_t^{\alpha} \tag{7}$$

Or,

$$\hat{Y}_t = \hat{K_t}^{\alpha} \tag{8}$$

Define $y_t = \log \hat{Y}_t$ and $k_t = \log \hat{K}_t$, so

$$y_t = \alpha k_t \tag{9}$$

For capital accumulation,

$$K_{t+1} = (1-\delta)K_t + \phi(\frac{I_t}{K_t})K_t \tag{10}$$

After detrending,

$$\hat{K}_{t+1} = [(1-\delta)\hat{K}_t + \phi(\frac{\hat{I}_t}{\hat{K}_t})\hat{K}_t]\exp(-\hat{\mu}_t - \sigma_a \hat{\epsilon}_{At+1}) \tag{11}$$

where $\phi(x) = x - \frac{\xi}{2}x^2$.

In the steady state,

$$(1-\delta) + \frac{\hat{I}}{\hat{K}} - \frac{\xi}{2}(\frac{\hat{I}}{\hat{K}})^2 = \exp(\mu_A) \approx 1 + \mu_A \tag{12}$$

Solving the quadratic equation gives

$$\frac{\hat{I}}{\hat{K}} = \frac{1 - \sqrt{1 - 2\xi(\mu_A + \delta)}}{\xi} \tag{13}$$

The capital accumulation equation can be log-linearized as

$$\begin{aligned}
\Delta k_{t+1} &= \log(\frac{\hat{K}_{t+1}}{\hat{K}_t}) \\
&= \log((1-\delta) + \phi(\exp(i_t - k_t))) - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \\
&= \log((1-\delta) + \exp(i_t - k_t) - \frac{\xi}{2}(\exp(i_t - k_t))^2) - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \\
&\approx \frac{\frac{\hat{I}}{\hat{K}} - \xi(\frac{\hat{I}}{\hat{K}})^2}{1 - \delta + \frac{\hat{I}}{\hat{K}} - \frac{\xi}{2}(\frac{\hat{I}}{\hat{K}})^2}(i_t - k_t) - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \\
&= \frac{\frac{1-\sqrt{1-2\xi(\mu_A+\delta)}}{\xi} - \xi\left(\frac{1-\sqrt{1-2\xi(\mu_A+\delta)}}{\xi}\right)^2}{1 - \delta + \mu_A + \delta}(i_t - k_t) - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \\
&\approx \frac{\mu_A + \delta}{1 + \mu_A}(i_t - k_t) - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1}
\end{aligned} \tag{14}$$

where $i_t - k_t$ can be approximated by

$$\begin{aligned}
i_t - k_t &\approx y_t - k_t - \frac{\exp(c-y)}{1-\exp(c-y)}(c_t - y_t) \\
&= \alpha k_t - k_t - \frac{\exp(c-y)}{1-\exp(c-y)}(c_t - \alpha k_t) \\
&= k_t\left[(\alpha - 1) + \frac{\alpha\exp(c-y)}{1-\exp(c-y)}\right] - \frac{\exp(c-y)}{1-\exp(c-y)}c_t
\end{aligned} \tag{15}$$

So

$$\Delta k_{t+1} = k_t \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] \frac{\mu_A + \delta}{1+\mu_A} - \frac{\exp(c-y)}{1-\exp(c-y)} \frac{\mu_A+\delta}{1+\mu_A} c_t - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \tag{16}$$

$$\begin{aligned} k_{t+1} &= k_t \left\{ \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] \frac{\mu_A + \delta}{1+\mu_A} + 1 \right\} - \frac{\exp(c-y)}{1-\exp(c-y)} \frac{\mu_A+\delta}{1+\mu_A} c_t - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \\ &= \lambda_1 k_t + \lambda_2 c_t - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \end{aligned} \tag{17}$$

where $\lambda_1 = \left[(\alpha-1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] \frac{\mu_A+\delta}{1+\mu_A} + 1$, and $\lambda_2 = -\frac{\exp(c-y)}{1-\exp(c-y)} \frac{\mu_A+\delta}{1+\mu_A}$.

The investment return is given by:

$$\begin{aligned} R_{it+1} &= \phi' \left(\frac{I_t}{K_t} \right) \left\{ \alpha \left(\frac{A_{t+1}}{K_{t+1}} \right)^{1-\alpha} + \frac{(1-\delta) + \phi\left(\frac{I_{t+1}}{K_{t+1}} \right) - \frac{I_{t+1}}{K_{t+1}} \phi' \left(\frac{I_{t+1}}{K_{t+1}} \right)}{\phi' \left(\frac{I_{t+1}}{K_{t+1}} \right)} \right\} \\ &= \left[1 - \xi \left(\frac{\hat{I}_t}{\hat{K}_t} \right) \right] \left\{ \alpha \hat{K}_{t+1}^{\alpha-1} + \frac{(1-\delta) + \frac{\xi}{2} \left(\frac{\hat{I}_{t+1}}{\hat{K}_{t+1}} \right)^2}{1 - \xi \left(\frac{\hat{I}_{t+1}}{\hat{K}_{t+1}} \right)} \right\} \end{aligned} \tag{18}$$

so

$$\begin{aligned} r_{it+1} &= \log(R_{it+1}) \\ &= \log\left[1 - \xi \exp(i_t - k_t) \right] + \log \left\{ \alpha \exp[(\alpha-1) k_{t+1}] + \frac{(1-\delta) + \frac{\xi}{2} \left(\exp(i_{t+1} - k_{t+1}) \right)^2}{1 - \xi \exp(i_{t+1} - k_{t+1})} \right\} \\ &\approx \frac{-\xi \left(\frac{\hat{I}}{\hat{K}} \right)}{1 - \xi \left(\frac{\hat{I}}{\hat{K}} \right)} (i_t - k_t) + \frac{\alpha(\alpha-1)\hat{K}^{\alpha-1}}{\alpha \hat{K}^{\alpha-1} + \frac{(1-\delta) + \frac{\xi}{2} \left(\frac{\hat{I}}{\hat{K}} \right)^2}{1 - \xi \left(\frac{\hat{I}}{\hat{K}} \right)}} k_{t+1} \\ &+ \frac{\xi \left(\frac{\hat{I}}{\hat{K}} \right)^2 \left[1 - \xi \left(\frac{\hat{I}}{\hat{K}} \right) \right] + \xi \left(\frac{\hat{I}}{\hat{K}} \right) \left[1 - \delta + \frac{\xi}{2} \left(\frac{\hat{I}}{\hat{K}} \right)^2 \right]}{\left[\alpha \hat{K}^{\alpha-1} + \frac{(1-\delta) + \frac{\xi}{2} \left(\frac{\hat{I}}{\hat{K}} \right)^2}{1 - \xi \left(\frac{\hat{I}}{\hat{K}} \right)} \right] \left[1 - \xi \left(\frac{\hat{I}}{\hat{K}} \right) \right]^2} (i_{t+1} - k_{t+1}) \end{aligned} \tag{19}$$

In the steady state,

$$\begin{aligned} R &= \exp(r) \approx 1 + r \\ &= \left[1 - \xi \left(\frac{\hat{I}}{\hat{K}} \right) \right] \left[\alpha \hat{K}^{\alpha-1} + \frac{(1-\delta) + \frac{\xi}{2} \left(\frac{\hat{I}}{\hat{K}} \right)^2}{1 - \xi \left(\frac{\hat{I}}{\hat{K}} \right)} \right] \\ &= \sqrt{1 - 2\xi(\mu_A + \delta)} \left[\alpha \hat{K}^{\alpha-1} + \frac{(1-\delta) + \frac{\xi}{2} \frac{1 + 1 - 2\xi(\mu_A+\delta) - 2\sqrt{1 - 2\xi(\delta + \mu_A)}}{\xi^2}}{\sqrt{1 - 2\xi(\mu_A+\delta)}} \right] \\ &= \sqrt{1 - 2\xi(\mu_A + \delta)} \left[\alpha \hat{K}^{\alpha-1} + \frac{1-\delta}{\sqrt{1 - 2\xi(\mu_A+\delta)}} + \frac{1 - \xi(\mu_A+\delta) - \sqrt{1 - 2\xi(\delta+\mu_A)}}{\xi\sqrt{1 - 2\xi(\mu_A+\delta)}} \right] \end{aligned} \tag{20}$$

So

$$\alpha \hat{K}^{\alpha-1} = \frac{(r+\delta) + (\mu_A+\delta) - \frac{1-\sqrt{1-2\xi(\delta+\mu_A)}}{\xi}}{\sqrt{1-2\xi(\mu_A+\delta)}} \tag{21}$$

Substitute the above equation and the expression for steady state of investment rate into the equation for the log-linearization of investment return:

$$\begin{aligned} r_{it+1} \approx & \frac{-(1-\sqrt{1-2\xi(\mu_A+\delta)})}{\sqrt{1-2\xi(\mu_A+\delta)}}(i_t - k_t) \\ & + \frac{(\alpha-1)\frac{(r+\delta)+(\mu_A+\delta)-\frac{1-\sqrt{1-2\xi(\delta+\mu_A)}}{\xi}}{\sqrt{1-2\xi(\mu_A+\delta)}}}{\frac{(r+\delta)+(\mu_A+\delta)-\frac{1-\sqrt{1-2\xi(\delta+\mu_A)}}{\xi}}{\sqrt{1-2\xi(\mu_A+\delta)}} + \frac{(1-\delta)+\frac{\xi}{2}\frac{1+1-2\xi(\mu_A+\delta)-2\sqrt{1-2\xi(\mu_A+\delta)}}{\xi^2}}{\sqrt{1-2\xi(\mu_A+\delta)}}} k_{t+1} \\ & + \frac{\xi\left(\frac{\hat{I}}{\hat{K}}\right)\left[\left(\frac{\hat{I}}{\hat{K}}\right) - \xi\left(\frac{\hat{I}}{\hat{K}}\right)^2 + 1 - \delta + \frac{\xi}{2}\left(\frac{\hat{I}}{\hat{K}}\right)^2\right]}{\left[\frac{(r+\delta)+(\mu_A+\delta)-\frac{1-\sqrt{1-2\xi(\delta+\mu_A)}}{\xi}}{\sqrt{1-2\xi(\mu_A+\delta)}} + \frac{(1-\delta)+\frac{\xi}{2}\frac{1+1-2\xi(\mu_A+\delta)-2\sqrt{1-2\xi(\mu_A+\delta)}}{\xi^2}}{\sqrt{1-2\xi(\mu_A+\delta)}}\right][1-2\xi(\mu_A+\delta)]}(i_{t+1} - k_{t+1}) \\ = & \frac{-(1-\sqrt{1-2\xi(\mu_A+\delta)})}{\sqrt{1-2\xi(\mu_A+\delta)}}(i_t - k_t) + \frac{(\alpha-1)\left[(r+\delta)+(\mu_A+\delta)-\frac{1-\sqrt{1-2\xi(\delta+\mu_A)}}{\xi}\right]}{1+r}k_{t+1} \\ & + \frac{(1-\sqrt{1-2\xi(\mu_A+\delta)})(1+\mu_A)}{(1+r)\sqrt{1-2\xi(\mu_A+\delta)}}(i_{t+1} - k_{t+1}) \end{aligned} \tag{22}$$

Implicitly define $\tilde{\lambda}_3$, λ_4, and λ_5 and rewrite the above equation as

$$r_{it+1} \approx \tilde{\lambda}_3 k_{t+1} + \lambda_4(i_t - k_t) + \lambda_5(i_{t+1} - k_{t+1}) \tag{23}$$

For return on consumption claim, the log-linear approximate expression follows Bansal and Yaron (2004):

$$\begin{aligned} r_{at+1} &\approx \kappa_1 h_{t+1} - h_t + \log\left(\frac{C_{t+1}}{C_t}\right) \\ &= \kappa_1 h_{t+1} - h_t + \Delta c_{t+1} + \hat{\mu}_t + \sigma_A \hat{\epsilon}_{At+1} \end{aligned} \tag{24}$$

Finally, the logarithm of pricing kernel is

$$\begin{aligned} m_{t+1} &= \theta \log \delta - \frac{\theta}{\psi}\log\left(\frac{C_{t+1}}{C_t}\right) + (\theta-1) r_{at+1} \\ &\approx -\frac{\theta}{\psi}(\Delta c_{t+1} + \hat{\mu}_t + \sigma_A \hat{\epsilon}_{At+1}) + (\theta-1)(\kappa_1 h_{t+1} - h_t + \Delta c_{t+1} + \hat{\mu}_t + \sigma_A \hat{\epsilon}_{At+1}) \\ &= (\theta - 1 - \frac{\theta}{\psi})\hat{\mu}_t + (\theta-1)\kappa_1 h_{t+1} - (\theta-1)h_t + (\theta - 1 - \frac{\theta}{\psi})\Delta c_{t+1} + (\theta - 1 - \frac{\theta}{\psi})\sigma_A \hat{\epsilon}_{At+1} \\ &= -\gamma\hat{\mu}_t + (\theta-1)(\kappa_1 h_{t+1} - h_t) - \gamma \Delta c_{t+1} - \gamma \sigma_A \hat{\epsilon}_{At+1} \end{aligned} \tag{25}$$

Now conjecture that:

$$h_t = \tilde{A}_1 k_t + \tilde{A}_2 \hat{\mu}_t \tag{26}$$

and

$$c_t = B_1 k_t + B_2 \hat{\mu}_t \tag{27}$$

which implies that:

$$\begin{aligned} k_{t+1} &= \lambda_1 k_t + \lambda_2 c_t - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \\ &= \lambda_1 k_t + \lambda_2 (B_1 k_t + B_2 \hat{\mu}_t) - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \\ &= (\lambda_1 + B_1 \lambda_2) k_t + (B_2 \lambda_2 - 1) \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \\ &= D_1 k_t + D_2 \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} \end{aligned} \tag{28}$$

where $D_1 = \lambda_1 + B_1 \lambda_2$ and $D_2 = B_2 \lambda_2 - 1$. The detrended consumption growth becomes:

$$\begin{aligned} \Delta c_{t+1} &= B_1 (k_{t+1} - k_t) + B_2 (\hat{\mu}_{t+1} - \hat{\mu}_t) \\ &= B_1 (D_1 k_t + D_2 \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1} - k_t) + B_2 [(1 - \rho - \tilde{\rho}) \mu_A + (\rho + \tilde{\rho}) \hat{\mu}_t + \tilde{\rho} \sigma_A \hat{\epsilon}_{At+1} - \hat{\mu}_t] \\ &\approx (B_1 D_1 - B_1) k_t + [B_1 D_2 + B_2 (\rho + \tilde{\rho} - 1)] \hat{\mu}_t + (-B_1 + B_2 \tilde{\rho}) \sigma_A \hat{\epsilon}_{At+1} \\ &= d_{ck} k_t + d_{c\mu} \hat{\mu}_t + d_{ce} \sigma_A \hat{\epsilon}_{At+1} \end{aligned} \tag{29}$$

where $d_{ck} = B_1 D_1 - B_1$, $d_{c\mu} = B_1 D_2 + B_2 (\rho + \tilde{\rho} - 1)$, and $d_{ce} = -B_1 + B_2 \tilde{\rho}$.

The return on consumption claim can be written as:

$$\begin{aligned} r_{at+1} &= \kappa_1 h_{t+1} - h_t + \Delta c_{t+1} + \hat{\mu}_t + \sigma_A \hat{\epsilon}_{At+1} \\ &= \kappa_1 (\tilde{A}_1 k_{t+1} + \tilde{A}_2 \hat{\mu}_{t+1}) - (\tilde{A}_1 k_t + \tilde{A}_2 \hat{\mu}_t) + d_{ck} k_t + d_{c\mu} \hat{\mu}_t + (-B_1 + B_2 \tilde{\rho}) \sigma_A \hat{\epsilon}_{At+1} + \hat{\mu}_t + \sigma_A \hat{\epsilon}_{At+1} \\ &= \kappa_1 \{ \tilde{A}_1 (D_1 k_t + D_2 \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1}) + \tilde{A}_2 [(1 - \rho - \tilde{\rho}) \mu_A + (\rho + \tilde{\rho}) \hat{\mu}_t + \tilde{\rho} \sigma_A \hat{\epsilon}_{At+1}]) \} \\ &- (\tilde{A}_1 k_t + \tilde{A}_2 \hat{\mu}_t) + d_{ck} k_t + d_{c\mu} \hat{\mu}_t + (-B_1 + B_2 \tilde{\rho}) \sigma_A \hat{\epsilon}_{At+1} + \hat{\mu}_t + \sigma_A \hat{\epsilon}_{At+1} \\ &\approx (\kappa_1 \tilde{A}_1 D_1 - \tilde{A}_1 + d_{ck}) k_t + [\kappa_1 \tilde{A}_1 D_2 + \kappa_1 \tilde{A}_2 (\rho + \tilde{\rho}) - \tilde{A}_2 + d_{c\mu} + 1] \hat{\mu}_t \\ &+ (-\kappa_1 \tilde{A}_1 + \kappa_1 \tilde{A}_2 \tilde{\rho} - B_1 + B_2 \tilde{\rho} + 1) \sigma_A \hat{\epsilon}_{At+1} \\ &= d_{rk} k_t + d_{r\mu} \hat{\mu}_t + d_{re} \sigma_A \hat{\epsilon}_{At+1} \end{aligned} \tag{30}$$

where $d_{rk} = \kappa_1 \tilde{A}_1 D_1 - \tilde{A}_1 + d_{ck}$, $d_{r\mu} = \kappa_1 \tilde{A}_1 D_2 + \kappa_1 \tilde{A}_2 (\rho + \tilde{\rho}) - \tilde{A}_2 + d_{c\mu} + 1$, and $d_{re} = -\kappa_1 \tilde{A}_1 + \kappa_1 \tilde{A}_2 \tilde{\rho} - B_1 + B_2 \tilde{\rho} + 1$.

Rewrite investment return as:

$$
\begin{aligned}
r_{it+1} &\approx \tilde{\lambda}_3 k_{t+1} + \lambda_4(i_t - k_t) + \lambda_5(i_{t+1} - k_{t+1}) \\
&= \tilde{\lambda}_3(\lambda_1 k_t + \lambda_2 c_t - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1}) + \lambda_4 \left\{ k_t \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] - \frac{\exp(c-y)}{1-\exp(c-y)} c_t \right\} \\
&+ \lambda_5 \left\{ k_{t+1} \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] - \frac{\exp(c-y)}{1-\exp(c-y)} c_{t+1} \right\} \\
&= \tilde{\lambda}_3[\lambda_1 k_t + \lambda_2(B_1 k_t + B_2 \hat{\mu}_t) - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1}] \\
&+ \lambda_4 \left\{ k_t \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] - \frac{\exp(c-y)}{1-\exp(c-y)} (B_1 k_t + B_2 \hat{\mu}_t) \right\} \\
&+ \lambda_5(D_1 k_t + D_2 \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1}) \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] \\
&- \lambda_5 \frac{\exp(c-y)}{1-\exp(c-y)} \left\{ B_1[\lambda_1 k_t + \lambda_2(B_1 k_t + B_2 \hat{\mu}_t) - \hat{\mu}_t - \sigma_A \hat{\epsilon}_{At+1}] + B_2[(1-\rho-\tilde{\rho})\mu_A + (\rho+\tilde{\rho})\hat{\mu}_t + \tilde{\rho}\sigma_A \hat{\epsilon}_{At+1}] \right\} \\
&\approx \{ \lambda_1 \tilde{\lambda_3} + B_1 \lambda_2 \tilde{\lambda}_3 + \lambda_4 \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} - \frac{B_1 \exp(c-y)}{1-\exp(c-y)} \right] + \lambda_5 D_1 \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] \\
&- \lambda_5 \frac{\exp(c-y)}{1-\exp(c-y)} [B_1(\lambda_1 + \lambda_2 B_1)] \} k_t + \{ \tilde{\lambda}_3(\lambda_2 B_2 - 1) - \lambda_4 B_2 \frac{\exp(c-y)}{1-\exp(c-y)} + \lambda_5 D_2 \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] \\
-\lambda_5 & \frac{\exp(c-y)}{1-\exp(c-y)} [B_1(\lambda_2 B_2 - 1) + B_2(\rho + \tilde{\rho})] \} \hat{\mu}_t \\
&+ \left\{ -\tilde{\lambda}_3 - \lambda_5 \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] + \lambda_5 \frac{\exp(c-y)}{1-\exp(c-y)} (B_1 - B_2 \tilde{\rho}) \right\} \sigma_A \hat{\epsilon}_{At+1} \\
&= \left\{ \tilde{\lambda}_3 D_1 + (\lambda_4 + \lambda_5 D_1) \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} - \frac{B_1 \exp(c-y)}{1-\exp(c-y)} \right] \right\} k_t \\
&+ \left\{ \frac{\exp(c-y)}{1-\exp(c-y)} [-\lambda_4 B_2 + \lambda_5 D_2(\alpha - B_1) - \lambda_5 B_2(\rho + \tilde{\rho})] + D_2[\tilde{\lambda_3} + \lambda_5(\alpha - 1)] \right\} \hat{\mu}_t \\
&+ \left[\frac{\lambda_5 \exp(c-y)}{1-\exp(c-y)} (-\alpha + B_1 - B_2 \tilde{\rho}) - \tilde{\lambda}_3 - \lambda_5(\alpha - 1) \right] \sigma_A \hat{\epsilon}_{At+1} \\
&= d_{ik} k_t + d_{i\mu} \hat{\mu}_t + d_{ie} \sigma_A \hat{\epsilon}_{At+1}
\end{aligned} \tag{31}
$$

where d_{ik}, $d_{i\mu}$ and d_{ie}.

The pricing kernel becomes:

$$
\begin{aligned}
m_{t+1} &= -\gamma\hat{\mu}_t + (\theta-1)(\kappa_1 h_{t+1} - h_t) - \gamma\Delta c_{t+1} - \gamma\sigma_A\hat{\epsilon}_{At+1} \\
&= -\gamma\hat{\mu}_t + (\theta-1)[\kappa_1(\tilde{A}_1 k_{t+1} + \tilde{A}_2\hat{\mu}_{t+1}) - (\tilde{A}_1 k_t + \tilde{A}_2\hat{\mu}_t)] - \gamma[d_{ck}k_t + d_{c\mu}\hat{\mu}_t + (-B_1 + B_2\tilde{\rho})\sigma_A\hat{\epsilon}_{At+1}]) - \gamma\sigma_A\hat{\epsilon}_{At+1} \\
&= -\gamma\hat{\mu}_t + (\theta-1)\{\kappa_1(\tilde{A}_1(D_1 k_t + D_2\hat{\mu}_t - \sigma_A\hat{\epsilon}_{At+1}) + \tilde{A}_2[(1-\rho-\tilde{\rho})\mu_A + (\rho+\tilde{\rho})\hat{\mu}_t + \tilde{\rho}\sigma_A\hat{\epsilon}_{At+1}]) - (\tilde{A}_1 k_t + \tilde{A}_2\hat{\mu}_t)\} \\
&- \gamma[d_{ck}k_t + d_{c\mu}\hat{\mu}_t + (-B_1 + B_2\tilde{\rho})\sigma_A\hat{\epsilon}_{At+1}]) - \gamma\sigma_A\hat{\epsilon}_{At+1} \\
&\approx [(\theta-1)(\kappa_1\tilde{A}_1 D_1 - \tilde{A}_1) - \gamma d_{ck}]k_t + \{-\gamma + (\theta-1)[\kappa_1\tilde{A}_1 D_2 + \kappa_1\tilde{A}_2(\rho+\tilde{\rho}) - \tilde{A}_2] - \gamma d_{c\mu}\}\hat{\mu}_t \\
&- [\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho}) + \gamma(-B_1 + B_2\tilde{\rho}) + \gamma]\sigma_A\hat{\epsilon}_{At+1} \\
&= [(\theta-1)(d_{rk} - d_{ck}) - \gamma d_{ck}]k_t + [-\gamma + (\theta-1)(d_{r\mu} - d_{c\mu} - 1) - \gamma d_{c\mu}]\hat{\mu}_t \\
&- [\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho}) + \gamma(-B_1 + B_2\tilde{\rho}) + \gamma]\sigma_A\hat{\epsilon}_{At+1} \\
&= [(\theta-1)d_{rk} - \frac{\theta}{\psi}d_{ck}]k_t + [(\theta-1)d_{r\mu} - \frac{\theta}{\psi} - \frac{\theta}{\psi}d_{c\mu}]\hat{\mu}_t - [\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho}) + \gamma(-B_1 + B_2\tilde{\rho}) + \gamma]\sigma_A\hat{\epsilon}_{At+1}
\end{aligned} \tag{32}
$$

So the changes in pricing kernel can be decomposed into a short-run and a long-run component:

$$
m_{t+1} - \hat{E}_t[m_{t+1}] = -[\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho}) + \gamma(-B_1 + B_2\tilde{\rho}) + \gamma]\sigma_A\hat{\epsilon}_{At+1} \tag{33}
$$

The short-run risk component is $\gamma(1 - B_1 + B_2\tilde{\rho})\sigma_A\hat{\epsilon}_{At+1}$, and the long-run risk component is $\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho})\sigma_A\hat{\epsilon}_{At+1}$. In the case when $\rho + \tilde{\rho} = 1$ and $\tilde{\rho} = 0$, the short-run risk component becomes $\gamma(1 - B_1)\sigma_A\hat{\epsilon}_{At+1} = \gamma B_2\sigma_A\hat{\epsilon}_{At+1}$, where the price of short-run risk is γ and the short-run consumption response is $\epsilon_t^c \approx B_2\sigma_A\hat{\epsilon}_{At+1}$. The long-run risk component in this special case is $\kappa_1(\theta-1)\tilde{A}_1\sigma_A\hat{\epsilon}_{At+1} = (A_1 - B_1)(\frac{1}{\psi} - \gamma)\sigma_A\hat{\epsilon}_{At+1} = (A_2 - B_2)(\gamma - \frac{1}{\psi})\sigma_A\hat{\epsilon}_{At+1}$, where the price of long-run risk is $\gamma - \frac{1}{\psi}$ and the long-run consumption response is $\epsilon_t^{vc} \approx (A_2 - B_2)\sigma_A\hat{\epsilon}_{At+1}$. These results are consistent with KL (2010).

In more general cases, the short-run risk component of the pricing kernel is $\gamma(1 - B_1 + \tilde{\rho}B_2)\sigma_A\hat{\epsilon}_{At+1}$, where the price of risk is still γ, but the short-run consumption response is augmented by this extra term $\tilde{\rho}B_2\sigma_A\hat{\epsilon}_{At+1}$ induced by the extrapolative bias. Similarly, the long-run component becomes $\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho})\sigma_A\hat{\epsilon}_{At+1} = (\gamma - \frac{1}{\psi})[-(A_1 - B_1) + \tilde{\rho}(A_2 - B_2)]\sigma_A\hat{\epsilon}_{At+1}$, where the price of long-run consumption risk is $\gamma - \frac{1}{\psi}$ and the long-run consumption response is augmented by the extra term $\tilde{\rho}(A_2 - B_2)\sigma_A\hat{\epsilon}_{At+1}$.

Now we solve the coefficients $\tilde{A}_1$, $\tilde{A}_1$, B_1, B_2 and confirm our earlier conjecture on the solution for h_t and c_t.

Apply the asset pricing equation on the consumption claim:

$$
\begin{aligned}
1 &= \hat{E}_t[\exp(m_{t+1} + r_{at+1})] \\
&= \hat{E}_t[\exp([(\theta-1)d_{rk} - \frac{\theta}{\psi}d_{ck}]k_t + [(\theta-1)d_{r\mu} - \frac{\theta}{\psi} - \frac{\theta}{\psi}d_{c\mu}]\hat{\mu}_t - [\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho}) + \gamma(-B_1 + B_2\tilde{\rho}) + \gamma]\sigma_A\hat{\epsilon}_{At+1} \\
&+ d_{rk}k_t + d_{r\mu}\hat{\mu}_t + (-\kappa_1\tilde{A}_1 + \kappa_1\tilde{A}_2\tilde{\rho} - B_1 + B_2\tilde{\rho} + 1)\sigma_A\hat{\epsilon}_{At+1})] \\
&= \hat{E}_t[\exp(\theta(d_{rk} - \frac{1}{\psi}d_{ck})k_t + \theta(d_{r\mu} - \frac{1}{\psi}d_{c\mu} - \frac{1}{\psi})\hat{\mu}_t \\
&+ \{(-\kappa_1\tilde{A}_1 + \kappa_1\tilde{A}_2\tilde{\rho} - B_1 + B_2\tilde{\rho} + 1) - [\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho}) + \gamma(-B_1 + B_2\tilde{\rho}) + \gamma]\}\sigma_A\hat{\epsilon}_{At+1}]
\end{aligned}
\tag{34}
$$

Since the above equality holds for all k_t and $\hat{\mu}_t$, the following two equations must hold:

$$d_{rk} = \frac{d_{ck}}{\psi} \tag{35}$$

$$d_{r\mu} = \frac{d_{c\mu} + 1}{\psi} \tag{36}$$

Rewrite the above two equations as:

$$
\begin{aligned}
d_{rk} &= \kappa_1\tilde{A}_1 D_1 - \tilde{A}_1 + d_{ck} = \frac{d_{ck}}{\psi} \\
\Rightarrow \tilde{A}_1 &= \frac{B_1(D_1 - 1)(1 - \frac{1}{\psi})}{1 - \kappa_1 D_1} = \frac{B_1(\lambda_1 + B_1\lambda_2 - 1)(1 - \frac{1}{\psi})}{1 - \kappa_1(\lambda_1 + B_1\lambda_2)}
\end{aligned}
\tag{37}
$$

and

$$
\begin{aligned}
d_{r\mu} &= \kappa_1\tilde{A}_1 D_2 + \kappa_1\tilde{A}_2(\rho + \tilde{\rho}) - \tilde{A}_2 + d_{c\mu} + 1 = \frac{d_{c\mu} + 1}{\psi} \\
\Rightarrow \tilde{A}_2 &= \frac{(\frac{1}{\psi} - 1)[B_1 D_2 + B_2(\rho + \tilde{\rho} - 1)] + (\frac{1}{\psi} - 1) - \kappa_1\tilde{A}_1 D_2}{\kappa_1(\rho + \tilde{\rho}) - 1} \\
&= \frac{(\frac{1}{\psi} - 1)[B_1(B_2\lambda_2 - 1) + B_2(\rho + \tilde{\rho} - 1) + 1] - \kappa_1\tilde{A}_1(B_2\lambda_2 - 1)}{\kappa_1(\rho + \tilde{\rho}) - 1}
\end{aligned}
\tag{38}
$$

Now the investment return can be priced:

$$
\begin{aligned}
1 &= \hat{E}_t[\exp(m_{t+1} + r_{it+1})] \\
&= \hat{E}_t[\exp([(\theta-1)d_{rk} - \frac{\theta}{\psi}d_{ck}]k_t + [(\theta-1)d_{r\mu} - \frac{\theta}{\psi} - \frac{\theta}{\psi}d_{c\mu}]\hat{\mu}_t - [\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho}) + \gamma(-B_1 + B_2\tilde{\rho}) + \gamma]\sigma_A\hat{\epsilon}_{At+1} \\
&+ d_{ik}k_t + d_{i\mu}\hat{\mu}_t + d_{ie}\sigma_A\hat{\epsilon}_{At+1})] \\
&= \hat{E}_t[\exp((-d_{rk} + d_{ik})k_t + (-d_{r\mu} + d_{i\mu})\hat{\mu}_t + \{d_{ie} - [\kappa_1(\theta-1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho}) + \gamma(-B_1 + B_2\tilde{\rho}) + \gamma]\}\sigma_A\hat{\epsilon}_{At+1})]
\end{aligned}
\tag{39}
$$

Again, the above equality holds for all k_t. This implies:

$$
\begin{aligned}
& d_{rk} = d_{ik} = \frac{d_{ck}}{\psi} \\
& \Rightarrow \tilde{\lambda}_3 D_1 + (\lambda_4 + \lambda_5 D_1)\left[(\alpha-1) + \frac{\alpha\exp(c-y)}{1-\exp(c-y)} - \frac{B_1\exp(c-y)}{1-\exp(c-y)}\right] = \frac{B_1 D_1 - B_1}{\psi} \\
& \Rightarrow \tilde{\lambda}_3(\lambda_1 + B_1\lambda_2) + [\lambda_4 + \lambda_5(\lambda_1 + B_1\lambda_2)]\left[(\alpha-1) + \frac{\alpha\exp(c-y)}{1-\exp(c-y)} - \frac{B_1\exp(c-y)}{1-\exp(c-y)}\right] = \frac{B_1(\lambda_1 + B_1\lambda_2) - B_1}{\psi} \\
& \Rightarrow Q_2 B_1^2 + Q_1 B_1 + Q_0 = 0
\end{aligned}
\tag{40}
$$

where

$$
Q_2 = \frac{\lambda_2\lambda_5\exp(c-y)}{1-\exp(c-y)} + \frac{\lambda_2}{\psi} \tag{41}
$$

$$
\begin{aligned}
Q_1 &= \frac{\lambda_1 - 1}{\psi} - \lambda_2\tilde{\lambda}_3 + \frac{(\lambda_4 + \lambda_1\lambda_5)\exp(c-y)}{1-\exp(c-y))} - \lambda_2\lambda_5\left[\alpha - 1 + \frac{\alpha\exp(c-y)}{1-\exp(c-y)}\right] \\
&= \frac{\exp(c-y)}{1-\exp(c-y)}(\lambda_4 + \lambda_1\lambda_5 - \alpha\lambda_2\lambda_5) + \frac{\lambda_1 - 1}{\psi} - \lambda_2\tilde{\lambda}_3 + \lambda_2\lambda_5(1-\alpha)
\end{aligned}
\tag{42}
$$

$$
Q_0 = -\lambda_1\tilde{\lambda_3} - (\lambda_4 + \lambda_1\lambda_5)\left[(\alpha-1) + \frac{\alpha\exp(c-y)}{1-\exp(c-y)}\right] \tag{43}
$$

Solving the quadratic equation gives

$$
B_1 = \frac{-Q_1 - \sqrt{Q_1^2 - 4Q_2Q_0}}{2Q_2} \tag{44}
$$

Similarly for $\hat{\mu}_t$:

$$
\begin{aligned}
&d_{r\mu} = d_{i\mu} = \frac{d_{c\mu}+1}{\psi} \\
&\Rightarrow \frac{\exp(c-y)}{1-\exp(c-y)}[-\lambda_4 B_2 + \lambda_5 D_2(\alpha - B_1) - \lambda_5 B_2(\rho+\tilde{\rho})] + D_2[\tilde{\lambda_3} + \lambda_5(\alpha-1)] = \frac{B_1 D_2 + B_2(\rho+\tilde{\rho}-1)+1}{\psi} \\
&\Rightarrow \frac{\exp(c-y)}{1-\exp(c-y)}[-\lambda_4 B_2 + \lambda_5(B_2\lambda_2 - 1)(\alpha - B_1) - \lambda_5 B_2(\rho+\tilde{\rho})] + (B_2\lambda_2-1)[\tilde{\lambda_3} + \lambda_5(\alpha-1)] \\
&\quad = \frac{B_1(B_2\lambda_2-1) + B_2(\rho+\tilde{\rho}-1)+1}{\psi} \\
&\Rightarrow B_2 = \frac{\frac{1-B_1}{\psi} + \tilde{\lambda_3} + \lambda_5(\alpha-1) + \frac{\lambda_5(\alpha-B_1)\exp(c-y)}{1-\exp(c-y)}}{\frac{\exp(c-y)}{1-\exp(c-y)}[-\lambda_4 + \lambda_2\lambda_5(\alpha-B_1) - \lambda_5(\rho+\tilde{\rho})] + \lambda_2[\tilde{\lambda_3}+\lambda_5(\alpha-1)] - \frac{B_1\lambda_2+\rho+\tilde{\rho}-1}{\psi}}
\end{aligned} \tag{45}
$$

Now we use lag operator to express the system. The perceived productivity growth follows AR(1) process:

$$
\hat{\mu}_t = \frac{\tilde{\rho}}{1-(\rho+\tilde{\rho})L}\sigma_A \hat{\epsilon}_{At} \tag{46}
$$

The capital accumulation equation implies

$$
k_t = \frac{(D_2\tilde{\rho}+\rho+\tilde{\rho})L - 1}{[1-(\rho+\tilde{\rho})L](1-D_1 L)}\sigma_A \hat{\epsilon}_{At} \tag{47}
$$

Therefore, log capital stock follows ARMA(2,1) process. For log output:

$$
y_t = \alpha k_t = \frac{\alpha[(D_2\tilde{\rho}+\rho+\tilde{\rho})L - 1]}{[1-(\rho+\tilde{\rho})L](1-D_1 L)}\sigma_A \hat{\epsilon}_{At} \tag{48}
$$

The log consumption can be written as:

$$
c_t = B_1 k_t + B_2 \hat{\mu}_t = \frac{[B_1(D_2\tilde{\rho}+\rho+\tilde{\rho}) - B_2\tilde{\rho}D_1]L + B_2\tilde{\rho} - B_1}{[1-(\rho+\tilde{\rho})L](1-D_1 L)}\sigma_A \hat{\epsilon}_{At} \tag{49}
$$

So log consumption also follows ARMA(2,1) process.

Finally, the perceived consumption growth becomes

$$
\begin{aligned}
&\log\left(\frac{C_t}{C_{t-1}}\right) = \Delta c_t + \hat{\mu}_{t-1} + \sigma_A\hat{\epsilon}_{At} = (1-L)c_t + \hat{\mu}_{t-1} + \sigma_A\hat{\epsilon}_{At} \\
&= \frac{\{[B_1(D_2\tilde{\rho}+\rho+\tilde{\rho}) - B_2\tilde{\rho}D_1]L + B_2\tilde{\rho} - B_1\}(1-L)}{[1-(\rho+\tilde{\rho})L](1-D_1 L)}\sigma_A\hat{\epsilon}_{At} + \frac{\tilde{\rho}L}{1-(\rho+\tilde{\rho})L}\sigma_A\hat{\epsilon}_{At} + \sigma_A\hat{\epsilon}_{At} \\
&= \frac{-[B_1(D_2\tilde{\rho}+\rho+\tilde{\rho}) - B_2\tilde{\rho}D_1 + D_1\rho]L^2 + [B_1(D_2\tilde{\rho}+\rho+\tilde{\rho}) - (B_2\tilde{\rho}+1)D_1 - (B_2\tilde{\rho}-B_1) - \rho]L + B_2\tilde{\rho} - B_1 + 1}{[1-(\rho+\tilde{\rho})L](1-D_1 L)}\sigma_A\hat{\epsilon}_{At}
\end{aligned} \tag{50}
$$

Note in the case when $\rho+\tilde{\rho}=1$ and $\tilde{\rho}=0$, as in KL (2010), consumption growth follows ARMA(1,1). But in more general cases when there is extrapolative bias, the perceived consumption growth follows ARMA(2,2) process.

The perceived expected consumption growth is:

$$\begin{aligned}
&\hat{E}_t\left[\log\left(\frac{C_{t+1}}{C_t}\right)\right] = E_t[\Delta c_{t+1} + \hat{\mu}_t + \sigma_A \hat{\epsilon}_{At+1}] \\
&= \hat{E}_t[(B_1D_1 - B_1)k_t + [B_1D_2 + B_2(\rho+\tilde{\rho}-1)]\hat{\mu}_t + (-B_1 + B_2\tilde{\rho})\sigma_A\hat{\epsilon}_{At+1} + \hat{\mu}_t + \sigma_A\hat{\epsilon}_{At+1}] \\
&= (B_1D_1 - B_1)k_t + [B_1D_2 + B_2(\rho+\tilde{\rho}-1)+1]\hat{\mu}_t \\
&= (B_1D_1 - B_1)\frac{(D_2\tilde{\rho}+\rho+\tilde{\rho})L-1}{[1-(\rho+\tilde{\rho})L](1-D_1L)}\sigma_A\hat{\epsilon}_{At} + [B_1D_2 + B_2(\rho+\tilde{\rho}-1)+1]\frac{\tilde{\rho}}{1-(\rho+\tilde{\rho})L}\sigma_A\hat{\epsilon}_{At} \\
&= \frac{(B_1D_1-B_1)[(D_2\tilde{\rho}+\rho+\tilde{\rho})L-1] + \tilde{\rho}[B_1D_2 + B_2(\rho+\tilde{\rho}-1)+1](1-D_1L)}{[1-(\rho+\tilde{\rho})L](1-D_1L)}\sigma_A\hat{\epsilon}_{At} \\
&= \frac{\{(B_1D_1-B_1)(D_2\tilde{\rho}+\rho+\tilde{\rho}) - \tilde{\rho}D_1[B_1D_2+B_2(\rho+\tilde{\rho}-1)+1]\}L + \tilde{\rho}[B_1D_2+B_2(\rho+\tilde{\rho}-1)+1] - (B_1D_1-B_1)}{[1-(\rho+\tilde{\rho})L](1-D_1L)}\sigma_A\hat{\epsilon}_{At}
\end{aligned} \tag{51}$$

The process for perceived expected consumption growth is AR(1) when $\rho+\tilde{\rho}=1$ and $\tilde{\rho}=0$, but in more general cases, it is ARMA(2,1).

To calculate the first-order autocorrelation for the perceived expected consumption growth, we derive the general formula for an ARMA(2,1) process $Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \theta_0\epsilon_t + \theta_1\epsilon_{t-1}$, where ϵ_t is a standard normal with mean 0 and variance σ^2. Define γ_k as $E_t[Z_tZ_{t-k}]$, and the following three equations hold:

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + (\theta_0^2 + \theta_1^2 + \phi_1\theta_0\theta_1)\sigma^2 \tag{52}$$

$$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 + \theta_0\theta_1\sigma^2 \tag{53}$$

$$\gamma_2 = \phi_1\gamma_1 + \phi_2\gamma_0 \tag{54}$$

After some algebra, the first-order autocorrelation can be written as:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{1+\phi_1^2-\phi_2^2+\frac{\theta_0^2+\theta_1^2}{\theta_0\theta_1}\phi_1}{2\phi_1+\frac{\theta_0^2+\theta_1^2}{\theta_0\theta_1}(1-\phi_2)} \tag{55}$$

For our perceived expected consumption growth,

$$\phi_1 = \rho + \tilde{\rho} + D_1 \tag{56}$$

$$\phi_2 = -(\rho + \tilde{\rho})D_1 \tag{57}$$

$$\theta_0 = \tilde{\rho}[B_1 D_2 + B_2(\rho + \tilde{\rho} - 1) + 1] - (B_1 D_1 - B_1) \tag{58}$$

$$\theta_1 = (B_1 D_1 - B_1)(D_2\tilde{\rho} + \rho + \tilde{\rho}) - \tilde{\rho}D_1[B_1 D_2 + B_2(\rho + \tilde{\rho} - 1) + 1]. \tag{59}$$

When $\rho + \tilde{\rho} \to 1$ and $\tilde{\rho} \to 0$, define $u = 1 - (\rho + \tilde{\rho})$, so that $\phi_1 = \lim_{u\to 0}(1 + D_1 - u)$, $\phi_2 = \lim_{u\to 0} -(1-u)D_1$, $\theta_0 = \lim_{\tilde{\rho}\to 0}[\tilde{\rho}(B_1 D_2 + 1) - B_1(D_1 - 1)]$, $\theta_1 = \lim_{u\to 0,\tilde{\rho}\to 0}\{B_1(D_1 - 1)(D_2\tilde{\rho} + 1 - u) - \tilde{\rho}D_1[B_1 D_2 - 1]\}$. So

$$\begin{aligned}
\rho_1 &= \lim_{u,\tilde{\rho}\to 0} \frac{\gamma_1}{\gamma_0} = \lim_{u,\tilde{\rho}\to 0} \frac{1 + \phi_1^2 - \phi_2^2 + \frac{\theta_0^2+\theta_1^2}{\theta_0\theta_1}\phi_1}{2\phi_1 + \frac{\theta_0^2+\theta_1^2}{\theta_0\theta_1}(1-\phi_2)} \\
&= \lim_{u,\tilde{\rho}\to 0} \frac{1 + (1 + D_1 - u)^2 - (1-u)^2 D_1^2 + \frac{[\tilde{\rho}(B_1D_2+1) - B_1(D_1-1)]^2 + \{B_1(D_1-1)(D_2\tilde{\rho}+1-u) - \tilde{\rho}D_1[B_1D_2-1]\}^2}{[\tilde{\rho}(B_1D_2+1) - B_1(D_1-1)]\{B_1(D_1-1)(D_2\tilde{\rho}+1-u) - \tilde{\rho}D_1[B_1D_2-1]\}}(1 + D_1 - u)}{2(1 + D_1 - u) + \frac{[\tilde{\rho}(B_1D_2+1) - B_1(D_1-1)]^2 + \{B_1(D_1-1)(D_2\tilde{\rho}+1-u) - \tilde{\rho}D_1[B_1D_2-1]\}^2}{[\tilde{\rho}(B_1D_2+1) - B_1(D_1-1)]\{B_1(D_1-1)(D_2\tilde{\rho}+1-u) - \tilde{\rho}D_1[B_1D_2-1]\}}(1 + (1-u)D_1)} \\
&= \lim_{u\to 0} \frac{1 + (1 + D_1)^2 - 2u(1 + D_1) - (1 - 2u)D_1^2 - 2(1 + D_1 - u)}{2(1 + D_1 - u) - 2(1 + (1-u)D_1)} \\
&= \lim_{u\to 0} \frac{2uD_1(D_1 - 1)}{2u(D_1 - 1)} = D_1
\end{aligned} \tag{60}$$

which is consistent with KL (2010).

2.1.2 Dynamics under objective measure

So far, the model is solved under the subjective measure, where the representative agent is subject to extrapolative bias. Under the objective measure or from the perspective of the econometrician, the dynamics of detrended variables are all different from those under the subjective measure.

The objective law of motion for the $\hat{\mu}_t$ is:

$$\begin{aligned}
\hat{\mu}_{t+1} &= (1 - \rho - \tilde{\rho})\mu_A + \rho\hat{\mu}_t + \tilde{\rho}g_{At+1} \\
&\approx \rho\hat{\mu}_t + \tilde{\rho}\sigma_A\epsilon_{At+1} \\
&= \rho L\hat{\mu}_{t+1} + \tilde{\rho}\sigma_A\epsilon_{At+1}
\end{aligned} \tag{61}$$

So $\hat{\mu}_t$ follows an AR(1) process, but its persistence under objective measure is smaller than the perceived persistence of the agent.

$$\hat{\mu}_t = \frac{\tilde{\rho}}{1 - \rho L}\sigma_A\epsilon_{At} \tag{62}$$

The capital accumulation equation under the objective measure becomes:

$$
\begin{aligned}
k_{t+1} &= \lambda_1 k_t + \lambda_2(B_1 k_t + B_2 \hat{\mu}_t) - \mu_A - \sigma_A \epsilon_{At+1} \\
&\approx (\lambda_1 + \lambda_2 B_1) k_t + \lambda_2 B_2 \hat{\mu}_t - \sigma_A \epsilon_{At+1} \\
&= (\lambda_1 + \lambda_2 B_1) L k_{t+1} + \frac{\lambda_2 B_2 \tilde{\rho} L}{1-\rho L} \sigma_A \epsilon_{At+1} - \sigma_A \epsilon_{At+1} \\
&= (\lambda_1 + \lambda_2 B_1) L k_{t+1} + \frac{(\lambda_2 B_2 \tilde{\rho} + \rho) L - 1}{1-\rho L} \sigma_A \epsilon_{At+1}
\end{aligned} \tag{63}
$$

So the detrended capital stock k_t follows ARMA(2,1) process under the objective measure:

$$
k_t = \frac{(\lambda_2 B_2 \tilde{\rho} + \rho) L - 1}{(1-\rho L)[1 - (\lambda_1 + \lambda_2 B_1) L]} \sigma_A \epsilon_{At} = \frac{(\lambda_2 B_2 \tilde{\rho} + \rho) L - 1}{(1-\rho L)[1 - D_1 L]} \sigma_A \epsilon_{At} \tag{64}
$$

Therefore, the realized log consumption growth can be written as:

$$
\begin{aligned}
&\log\left(\frac{C_{t+1}}{C_t}\right) = \Delta c_{t+1} + \mu_A + \sigma_A \epsilon_{At+1} \\
&= B_1(k_{t+1} - k_t) + B_2(\hat{\mu}_{t+1} - \hat{\mu}_t) + \mu_A + \sigma_A \epsilon_{At+1} \\
&= B_1(\lambda_1 k_t + \lambda_2(B_1 k_t + B_2 \hat{\mu}_t) - \mu_A - \sigma_A \epsilon_{At+1} - k_t) + B_2(\rho \hat{\mu}_t + \tilde{\rho} \sigma_A \epsilon_{At+1} - \hat{\mu}_t) + \mu_A + \sigma_A \epsilon_{At+1} \\
&\approx B_1(\lambda_1 + B_1 \lambda_2 - 1) k_t + (B_1 \lambda_2 B_2 + B_2 \rho - B_2) \hat{\mu}_t + (1 - B_1 + B_2 \tilde{\rho}) \sigma_A \epsilon_{At+1} \\
&= \frac{B_1(\lambda_1 + B_1 \lambda_2 - 1)[(\lambda_2 B_2 \tilde{\rho} + \rho) L - 1] L}{(1-\rho L)[1 - D_1 L]} \sigma_A \epsilon_{At+1} + \frac{\tilde{\rho}(B_1 \lambda_2 B_2 + B_2 \rho - B_2) L}{1 - \rho L} \sigma_A \epsilon_{At+1} + (1 - B_1 + B_2 \tilde{\rho}) \sigma_A \epsilon_{At+1} \\
&= \frac{[\rho D_1(B_2 + 1) - B_1(B_2 \lambda_2 \tilde{\rho} + \rho)] L^2 + [B_1(\tilde{\rho} B_2 \lambda_2 + \rho) - D_1(1 + B_2 \tilde{\rho}) + B_2(1 - \tilde{\rho}) - \rho] L + (1 - B_1 + B_2 \tilde{\rho})}{(1-\rho L)[1 - D_1 L]} \sigma_A \epsilon_{At+1}
\end{aligned} \tag{65}
$$

so it follows an ARMA(2,2) process.

The expected consumption growth under the objective measure follows an ARMA(2,1) process:

$$
\begin{aligned}
&E_t\left[\log\left(\frac{C_{t+1}}{C_t}\right)\right] \approx B_1(\lambda_1 + B_1 \lambda_2 - 1) k_t + (B_1 \lambda_2 B_2 + B_2 \rho - B_2) \hat{\mu}_t \\
&\approx B_1(\lambda_1 + B_1 \lambda_2 - 1) \frac{(\lambda_2 B_2 \tilde{\rho} + \rho) L - 1}{(1-\rho L)[1 - D_1 L]} \sigma_A \epsilon_{At} + (B_1 \lambda_2 B_2 + B_2 \rho - B_2) \frac{\tilde{\rho}}{1 - \rho L} \sigma_A \epsilon_{At} \\
&\approx \frac{[(\lambda_2 B_2 \tilde{\rho} + \rho) L - 1] B_1(\lambda_1 + B_1 \lambda_2 - 1) + (B_1 \lambda_2 B_2 + B_2 \rho - B_2) \tilde{\rho}(1 - D_1 L)}{(1-\rho L)[1 - D_1 L]} \sigma_A \epsilon_{At} \\
&= \frac{[B_1(\lambda_2 B_2 \tilde{\rho} + \rho)(\lambda_1 + B_1 \lambda_2 - 1) - D_1 \tilde{\rho}(B_1 \lambda_2 B_2 + B_2 \rho - B_2)] L + (B_1 \lambda_2 B_2 + B_2 \rho - B_2) \tilde{\rho} - B_1(\lambda_1 + B_1 \lambda_2 - 1)}{(1-\rho L)[1 - D_1 L]} \sigma_A \epsilon_{At}
\end{aligned} \tag{66}
$$

The pricing kernel under objective measure is almost the same as under subjective measure (except that now the shock is the true shock in the data generating process for consumption growth):

$$
\begin{aligned}
m_{t+1} - E_t[m_{t+1}] &= [-\gamma(1 - B_1 + B_2\tilde{\rho}) + (\theta - 1)\kappa_1(\tilde{A}_2\tilde{\rho} - \tilde{A}_1)]\sigma_A\epsilon_{At+1} \\
&= -\gamma(1 - B_1 + B_2\tilde{\rho})\sigma_A\epsilon_{At+1} - (\gamma - \frac{1}{\psi})\frac{\kappa_1(-\tilde{A}_1 + \tilde{A}_2)\tilde{\rho}}{1 - \frac{1}{\psi}}\sigma_A\epsilon_{At+1} \\
&= -\gamma(1 - B_1 + B_2\tilde{\rho})\sigma_A\epsilon_{At+1} - (\gamma - \frac{1}{\psi})[-(A_1 - B_1) + \tilde{\rho}(A_2 - B_2)]\sigma_A\epsilon_{At+1}
\end{aligned}
\tag{67}
$$

where the first term is short-run risk and the second term is long-run risk. This result is not surprising because the marginal utility of the agent is still based on the perceived dynamics of consumption growth.

The dynamics of investment return under the objective measure can also provide some information regarding the stock return predictability by valuation ratios. The investment return process under objective measure can be written as

$$
\begin{aligned}
r_{it+1} \approx\ & \tilde{\lambda}_3 k_{t+1} + \lambda_4(i_t - k_t) + \lambda_5(i_{t+1} - k_{t+1}) \\
=\ & \tilde{\lambda}_3(\lambda_1 k_t + \lambda_2 c_t - \mu_A - \sigma_A \epsilon_{At+1}) + \lambda_4 \left\{ k_t \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] - \frac{\exp(c-y)}{1-\exp(c-y)} c_t \right\} \\
& + \lambda_5 \left\{ k_{t+1} \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] - \frac{\exp(c-y)}{1-\exp(c-y)} c_{t+1} \right\} \\
=\ & \tilde{\lambda}_3[\lambda_1 k_t + \lambda_2(B_1 k_t + B_2 \hat{\mu}_t) - \sigma_A \epsilon_{At+1}] \\
& + \lambda_4 \left\{ k_t \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] - \frac{\exp(c-y)}{1-\exp(c-y)} (B_1 k_t + B_2 \hat{\mu}_t) \right\} \\
& + \lambda_5(D_1 k_t + \lambda_2 B_2 \hat{\mu}_t - \sigma_A \epsilon_{At+1}) \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] \\
& - \lambda_5 \frac{\exp(c-y)}{1-\exp(c-y)} \left\{ B_1[\lambda_1 k_t + \lambda_2(B_1 k_t + B_2 \hat{\mu}_t) - \sigma_A \epsilon_{At+1}] + B_2[\rho \hat{\mu}_t + \tilde{\rho} \sigma_A \epsilon_{At+1}] \right\} \\
\approx\ & \{ \lambda_1 \tilde{\lambda_3} + B_1 \lambda_2 \tilde{\lambda}_3 + \lambda_4 \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} - \frac{B_1 \exp(c-y)}{1-\exp(c-y)} \right] + \lambda_5 D_1 \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] \\
& - \lambda_5 \frac{\exp(c-y)}{1-\exp(c-y)} [B_1(\lambda_1 + \lambda_2 B_1)] \} k_t + \{ \tilde{\lambda}_3 \lambda_2 B_2 - \lambda_4 B_2 \frac{\exp(c-y)}{1-\exp(c-y)} + \lambda_5 \lambda_2 B_2 \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] \\
& -\lambda_5 \frac{\exp(c-y)}{1-\exp(c-y)} [B_1 \lambda_2 B_2 + B_2 \rho] \} \hat{\mu}_t \\
& + \left\{ -\tilde{\lambda}_3 - \lambda_5 \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} \right] + \lambda_5 \frac{\exp(c-y)}{1-\exp(c-y)} (B_1 - B_2 \tilde{\rho}) \right\} \sigma_A \epsilon_{At+1} \\
=\ & \left\{ \tilde{\lambda}_3 D_1 + (\lambda_4 + \lambda_5 D_1) \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} - \frac{B_1 \exp(c-y)}{1-\exp(c-y)} \right] \right\} k_t \\
& + \left\{ \frac{\exp(c-y)}{1-\exp(c-y)} [-\lambda_4 B_2 + \lambda_5 \lambda_2 B_2(\alpha - B_1) - \lambda_5 B_2 \rho] + \lambda_2 B_2 [\tilde{\lambda_3} + \lambda_5(\alpha - 1)] \right\} \hat{\mu}_t \\
& + \left[\frac{\lambda_5 \exp(c-y)}{1-\exp(c-y)} (-\alpha + B_1 - B_2 \tilde{\rho}) - \tilde{\lambda}_3 - \lambda_5(\alpha - 1) \right] \sigma_A \epsilon_{At+1} \\
=\ & \tilde{d}_{ik} k_t + \tilde{d}_{i\mu} \hat{\mu}_t + \tilde{d}_{ie} \sigma_A \epsilon_{At+1}
\end{aligned}
\tag{68}
$$

where $\tilde{d}_{ik}$, $\tilde{d}_{i\mu}$, and $\tilde{d}_{ie}$.

Now compare $\tilde{d}_{i\mu}$ with $\tilde{A}_2$. Under our calibration and a wide range of parameter values, $\tilde{d}_{i\mu}$ is negative and $\tilde{A}_2$ is always positive. Therefore, a higher valuation ratio (such as wealth-consumption ratio or price-dividend ratio) predicts a lower future return.

2.2 Summary and Discussion

The analytical approximation is still quite complex due to the extrapolation bias and the capital accumulation equation. Thus, it is still very hard to gauge the intuition without resorting to numerical analysis. Many of the quantities are endogenous. For example, B_1, B_2, D_1 and D_2 all depends on the wealth consumption ratio, which is endogenous. Thus, it is quite hard to derive analytical properties on the dynamics of the variables of interests. In many cases, the analytically approximation still have to be solved numerically. Moreover, many properties may only hold under specific parameterizations (such as high adjustment cost, high IES, etc). Thus, our discussion tries not to make general claim on the properties based on analytical approximation. Instead, the numerical solution are mostly used for our model calibration. As a result, our discussion only mentions a few key intuitions derived from the analytical approximation in the main text. Below we summarize a few results based on our analytical approximation.

a. Expected consumption growth under subjective measure

The perceived expected consumption growth follows an ARMA(2,1) process:

$$\hat{E}_t\left[\log\left(\frac{C_{t+1}}{C_t}\right)\right] = \frac{\{(B_1D_1 - B_1)(D_2\tilde{\rho} + \rho + \tilde{\rho}) - \tilde{\rho}D_1[B_1D_2 + B_2(\rho + \tilde{\rho} - 1) + 1]\}L + \tilde{\rho}[B_1D_2 + B_2(\rho + \tilde{\rho} - 1) + 1] - (B_1D_1 - B_1)}{[1 - (\rho + \tilde{\rho})L](1 - D_1L)}\sigma_A\hat{\epsilon}_{At}, \tag{69}$$

where $B_1, B_2, D_1,$ and D_2 are defined above. When $\rho + \tilde{\rho} = 1$ and $\tilde{\rho} = 0$, the above process reduces to an $AR(1)$ as in KL(2010).

The AR(1) correlation coefficient ρ_1 can be calculated for the perceived expected consumption growth:

$$\rho_1 = \frac{1 + \phi_1^2 - \phi_2^2 + \frac{\theta_0^2 + \theta_1^2}{\theta_0\theta_1}\phi_1}{2\phi_1 + \frac{\theta_0^2 + \theta_1^2}{\theta_0\theta_1}(1 - \phi_2)} \tag{70}$$

where $\phi_1 = \rho + \tilde{\rho} + D_1$, $\phi_2 = -(\rho + \tilde{\rho})D_1$, $\theta_0 = \tilde{\rho}[B_1D_2 + B_2(\rho + \tilde{\rho} - 1) + 1] - (B_1D_1 - B_1)$, and $\theta_1 = (B_1D_1 - B_1)(D_2\tilde{\rho} + \rho + \tilde{\rho}) - \tilde{\rho}D_1[B_1D_2 + B_2(\rho + \tilde{\rho} - 1) + 1]$. When $\rho + \tilde{\rho} = 1$ and $\tilde{\rho} = 0$, it can be shown that $\rho_1 = D_1$, again consistent with Equation (A41) in KL (2010).

The above formula is quite complex, and thus it is very hard (if not impossible) to derive analytical properties on the relation between the persistence and extrapolation parameter $\tilde{\rho}$. However, it is clear from the above equation that the persistence depends on both D_1 and $\rho + \tilde{\rho}$. One can view D_1 as the persistence due to capital accumulation (see Equation (24)), whereas $\rho + \tilde{\rho}$ is due to the misperception in the persistence of the expected TFP growth. Thus, both capital accumulation and misperception contribute to the persistence of the expected consumption growth. Intuitively, the endogenous capital accumulation can lead to persistence in consumption growth; the intuition is exactly the same as in KL (2010). Also since investors misperceive the persistence of the TFP growth, which also contributes to the perceived persistence in the expected consumption growth.

More important, owing to extrapolation-induced overreaction, the volatility of the perceived expected consumption growth is increasing in the extrapolation bias $\tilde{\rho}$. Specifically, after favorable TFP shocks, due to extrapolation bias, perceived expected TFP growth is raised, and thus investments are increased and current consumption is reduced (compared to the cast of no extrapolation bias). However, expected future consumption is increased due to higher perceived TFP in the future. Following adverse shocks, the opposite occurs to investment and consumption. Thus, extrapolation bias tends to raise the volatility of perceived expected consumption growth. This amplified perceived expected consumption growth is translated into an amplified long-run consumption risk and hence more volatile pricing kernel when the agent has an Epstein-Zin preference with an early resolution of uncertainty.

b. Consumption dynamics under the objective measure

The consumption growth under objective measure follows an ARMA(2,2) process:

$$\log\left(\frac{C_{t+1}}{C_t}\right) = \frac{[\rho D_1(B_2+1) - B_1(B_2\lambda_2\tilde{\rho}+\rho)]L^2 + [B_1(\tilde{\rho}B_2\lambda_2+\rho) - D_1(1+B_2\tilde{\rho}) + B_2(1-\tilde{\rho}) - \rho]L + (1-B_1+B_2\tilde{\rho})}{(1-\rho L)[1-D_1 L].}\sigma_A\epsilon_{At+1}, \tag{71}$$

This is slightly different from the perceived consumption growth under the subjective measure

$$\log\left(\frac{C_{t+1}}{C_t}\right) = \frac{-[B_1(D_2\tilde{\rho}+\rho+\tilde{\rho}) - B_2\tilde{\rho}D_1 + D_1\rho]L^2 + [B_1(D_2\tilde{\rho}+\rho+\tilde{\rho}) - (B_2\tilde{\rho}+1)D_1 - (B_2\tilde{\rho}-B_1) - \rho]L + B_2\tilde{\rho} - B_1 + 1}{[1-(\rho+\tilde{\rho})L](1-D_1 L)}\sigma_A\hat{\epsilon}_{At+1}, \tag{72}$$

which also follows an ARMA(2,2) process. In the case of $\rho + \tilde{\rho} = 1$ and $\tilde{\rho} = 0$, consumption growth follows ARMA(1,1), a result consistent with KL (2010).

The conditional volatility of consumption growth in both subjective and objective measures is $(1-B_1+B_2\tilde{\rho})$. In our calibration, $B_2 < 0$, so higher extrapolative bias tends to reduce consumption response to productivity shocks. Therefore, the conditional volatility of consumption growth decreases with $\tilde{\rho}$. The intuition is the following: When extrapolative bias is stronger, a positive productivity shock gets amplified even more in the perceived measure. The agent with a high IES wants to take advantage of this large perceived shock by increasing investment and reducing current consumption. That is, investments tend to absorb more of the payoff variation resulting from productivity shocks, leading to smoother actual consumption growth. Therefore, even though the long-run consumption risk increases with $\tilde{\rho}$ due to the amplified expected consumption growth, the short-run consumption risk becomes smaller when the extrapolative bias is stronger.

The expected consumption growth under the objective measure follows an ARMA(2,1) process:

$$
\begin{aligned}
&E_t\left[\log\left(\frac{C_{t+1}}{C_t}\right)\right] \approx B_1(\lambda_1 + B_1\lambda_2 - 1)k_t + (B_1\lambda_2 B_2 + B_2\rho - B_2)\hat{\mu}_t \\
&\approx B_1(\lambda_1 + B_1\lambda_2 - 1)\frac{(\lambda_2 B_2\tilde{\rho} + \rho)L - 1}{(1-\rho L)[1 - D_1 L]}\sigma_A\epsilon_{At} + (B_1\lambda_2 B_2 + B_2\rho - B_2)\frac{\tilde{\rho}}{1-\rho L}\sigma_A\epsilon_{At} \\
&\approx \frac{[(\lambda_2 B_2\tilde{\rho} + \rho)L - 1]B_1(\lambda_1 + B_1\lambda_2 - 1) + (B_1\lambda_2 B_2 + B_2\rho - B_2)\tilde{\rho}(1 - D_1 L)}{(1-\rho L)[1 - D_1 L]}\sigma_A\epsilon_{At} \\
&= \frac{[B_1(\lambda_2 B_2\tilde{\rho} + \rho)(\lambda_1 + B_1\lambda_2 - 1) - D_1\tilde{\rho}(B_1\lambda_2 B_2 + B_2\rho - B_2)]L + (B_1\lambda_2 B_2 + B_2\rho - B_2)\tilde{\rho} - B_1(\lambda_1 + B_1\lambda_2 - 1)}{(1-\rho L)[1 - D_1 L]}\sigma_A\epsilon_{At}
\end{aligned}
\tag{73}
$$

Our numerical analysis shows that the conditional volatility of the expected consumption growth is about two times bigger under the subjective measure than the objective measure when $\tilde{\rho}$ takes our benchmark value of 0.02. The amplified conditional volatility of the perceived expected consumption growth is an important source of the volatile pricing kernel implied from our model compared with the one without extrapolative bias.

d. Impulse response functions

The above discussion can be more explicit by looking at the impulse response functions of consumption and expected consumption growth under the subjective and objective measures, as well as the situation where there is no extrapolative bias (KL (2010)). To calculate the impulse response, the process of each variable is expressed using lag operators, and the coefficient of the order T in the Taylor expansion around zero represents the impulse response in period T.

d1. Subjective measure

Under the subjective measure, the perceived TFP growth is:

$$\hat{\mu}_t = \frac{\tilde{\rho}}{1-(\rho+\tilde{\rho})L}\sigma_A\hat{\epsilon}_{At}, \tag{74}$$

the perceived TFP level is:

$$\hat{a}_t = \frac{1-\rho L}{[1-(\rho+\tilde{\rho})L](1-L)}\sigma_A\hat{\epsilon}_{At}, \tag{75}$$

the (Log) Consumption level is equal to:

$$\log(C_t) = c_t + \hat{a}_t = \left[\frac{[B_1(D_2\tilde{\rho} + \rho + \tilde{\rho}) - B_2\tilde{\rho}D_1]L + B_2\tilde{\rho} - B_1}{[1-(\rho+\tilde{\rho})L](1-D_1 L)} + \frac{1-\rho L}{[1-(\rho+\tilde{\rho})L](1-L)}\right]\sigma_A\hat{\epsilon}_{At} \tag{76}$$

and the perceived expected consumption growth is:

$$\hat{E}_t\left[\log\left(\frac{C_{t+1}}{C_t}\right)\right] = \frac{\{(B_1D_1-B_1)(D_2\tilde{\rho}+\rho+\tilde{\rho})-\tilde{\rho}D_1[B_1D_2+B_2(\rho+\tilde{\rho}-1)+1]\}L+\tilde{\rho}[B_1D_2+B_2(\rho+\tilde{\rho}-1)+1]-(B_1D_1-B_1).}{[1-(\rho+\tilde{\rho})L](1-D_1L)}\sigma_A\hat{\epsilon}_{At} \tag{77}$$

d2. Objective measure

Under the objective measure, the perceived TFP growth by the agent becomes:

$$\hat{\mu}_t = \frac{\tilde{\rho}}{1-\rho L}\sigma_A\epsilon_{At}, \tag{78}$$

the perceived TFP level by the agent is:

$$\hat{a}_t = \frac{1-\rho L+\tilde{\rho}L}{(1-\rho L)(1-L)}\sigma_A\epsilon_{At}, \tag{79}$$

and the (log) consumption level is:

$$\log(C_t) = c_t + a_t = \left[\frac{(B_1\lambda_2B_2\tilde{\rho}+B_1\rho-B_2\tilde{\rho}D_1)L+B_2\tilde{\rho}-B_1}{(1-\rho L)(1-D_1L)}+\frac{1}{1-L}\right]\sigma_A\epsilon_{At}. \tag{80}$$

It should be noted that log consumption is equal to $c_t + a_t$, not $c_t + \hat{a}_t$. From the perspective of the econometrician, a_t follows a random walk, in contrast with an AR(1) process from the perspective of the agent.

The perceived expected consumption growth under the objective measure is

$$E_t\left[\log\left(\frac{C_{t+1}}{C_t}\right)\right] = \frac{[B_1(\lambda_2B_2\tilde{\rho}+\rho)(\lambda_1+B_1\lambda_2-1)-D_1\tilde{\rho}(B_1\lambda_2B_2+B_2\rho-B_2)]L+(B_1\lambda_2B_2+B_2\rho-B_2)\tilde{\rho}-B_1(\lambda_1+B_1\lambda_2-1)}{(1-\rho L)[1-D_1L]}\sigma_A\epsilon_{At}. \tag{81}$$

d3. No bias

In the case of no extrapolative bias, ρ and $\tilde{\rho}$ are set to zero, which corresponds to the setup in KL(2010). The impulse response can be calculated from either subjective or objective measure, since they should give the same solution.

d4. Results

Figure 1 reports the impulse response functions to a positive one-standard-deviation TFP shock. We first compare the the impulse response functions for the perceived (log) TFP level and perceived TFP growth by the agent, (log) consumption level, and the agent's perceived consumption growth under subjective and objective measures, when the representative agent is subject to extrapolative bias. Because $\rho + \tilde{\rho}$ is close to one in our benchmark calibration, the expected TFP growth is very persistent (top two panels). This is in contrast with a much less persistent TFP growth rate under the objective measure, whose dynamics are the weighted average of the perceived and realized TFP growth (i.i.d.). The endogenous consumption responses are plotted in the bottom two panels. Due to the difference in persistence and magnitude in the perceived TFP growth under these two measures, the expected consumption growth is also much larger and more persistent under subjective measure than the objective measure.

We next compare the cases with and without extrapolative bias. Because the underlying process for the TFP is random walk, there is no predictable component in TFP when the individual is rational. A positive realized TFP shock increases the TFP level permanently, but does not change the expected TFP growth. More interestingly, when there is extrapolation bias, the consumption response to TFP shocks is weaker than in the case of no extrapolation bias as in KL (2010). This is because higher perceived TFP growth induces more investment now and less current consumption. However, the perceived future consumption growth by the agent would be higher with extrapolation bias than without extrapolation bias. Therefore, the introduction of extrapolative bias substantially increases the amount of long-run consumption risks under the subjective measure.

Again, since asset prices are determined by the dynamics under the subjective measure, the pricing kernel in our extrapolation model is more volatile than that in a standard long-run risk model (KL (2010)). At the same time, the smaller consumption response with extrapolative bias allows us to increase the capital adjustment cost, which helps raise the volatility of stock returns or the quantity of risk in the economy. In addition, by comparing the impulse responses of consumption under the objective measure with and without the extrapolative bias, it is seen that the consumption dynamics under these two situations are similar and the expected consumption growth under those two cases is less volatile and less persistent than that under the subjective measure with extrapolation bias. Since the realized consumption is generated by the dynamics under the objective measure, it is smooth relative to the output. In sum, these features combined explain why our model generates a large and volatile equity premium while keeping a smooth consumption process relative to the output.

e. Pricing kernel under subjective measure

With the extrapolative bias, the innovation in pricing kernel becomes

$$m_{t+1} - \hat{E}_t[m_{t+1}] = -[\kappa_1(\theta - 1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho}) + \gamma(1 - B_1 + B_2\tilde{\rho})]\sigma_A\hat{\epsilon}_{At+1} \tag{82}$$

The short-run risk component is $\gamma(1 - B_1 + B_2\tilde{\rho})\sigma_A\hat{\epsilon}_{At+1}$, and the long-run risk component is $\kappa_1(\theta - 1)(\tilde{A}_1 - \tilde{A}_2\tilde{\rho})\sigma_A\hat{\epsilon}_{At+1}$. In the case when $\rho + \tilde{\rho} = 1$ and $\tilde{\rho} = 0$, the short-run risk component becomes $\gamma(1 - B_1)\sigma_A\hat{\epsilon}_{At+1} =$

$\gamma B_2 \sigma_A \hat{\epsilon}_{At+1}$, where the price of short-run risk is γ and the short-run consumption response is $\epsilon_t^c \approx B_2 \sigma_A \hat{\epsilon}_{At+1}$. The long-run risk component in this special case is $\kappa_1(\theta - 1)\tilde{A}_1 \sigma_A \hat{\epsilon}_{At+1} = (A_1 - B_1)(\frac{1}{\psi} - \gamma)\sigma_A \hat{\epsilon}_{At+1} = (A_2 - B_2)(\gamma - \frac{1}{\psi})\sigma_A \hat{\epsilon}_{At+1}$,where the price of long-run risk is $\gamma - \frac{1}{\psi}$ and the long-run consumption response is $\epsilon^{vc} \approx (A_2 - B_2)\, \sigma_A \hat{\epsilon}_{At+1}$. These results are consistent with KL (2010).

In more general cases, the short-run risk component of the pricing kernel is $\gamma(1 - B_1 + \tilde{\rho} B_2)\sigma_A \hat{\epsilon}_{At+1}$, where the price of risk is still γ, but the short-run consumption response is augmented by this extra term $\tilde{\rho} B_2 \sigma_A \hat{\epsilon}_{At+1}$ induced by the extrapolative bias. In our benchmark calibration, $B_2 < 0$. Therefore, the introduction of extrapolation bias reduces the short-run consumption risk. On the other hand, the long-run component becomes $\kappa_1(\theta - 1)(\tilde{A}_1 - \tilde{A}_2 \tilde{\rho})\sigma_A \hat{\epsilon}_{At+1} = (\gamma - \frac{1}{\psi})[-(A_1 - B_1) + \tilde{\rho}(A_2 - B_2)]\sigma_A \hat{\epsilon}_{At+1}$, where the price of long-run consumption risk is $\gamma - \frac{1}{\psi}$ and the long-run consumption response is augmented by the extra term $\tilde{\rho}(A_2 - B_2)\sigma_A \hat{\epsilon}_{At+1}$. $(A_2 - B_2)$ under our benchmark calibration is positive, so a large extrapolative bias $\tilde{\rho}$ induces a larger long-run consumption risk and a higher risk premium.

f. Perceived return on dividend claim under subjective measure

The return on the dividend claim can be written as

$$\begin{aligned} r_{it+1} \approx & \left\{ \tilde{\lambda}_3 D_1 + (\lambda_4 + \lambda_5 D_1) \left[(\alpha - 1) + \frac{\alpha \exp(c - y)}{1 - \exp(c - y)} - \frac{B_1 \exp(c - y)}{1 - \exp(c - y)} \right] \right\} k_t \\ & + \left\{ \frac{\exp(c - y)}{1 - \exp(c - y)} [-\lambda_4 B_2 + \lambda_5 D_2(\alpha - B_1) - \lambda_5 B_2(\rho + \tilde{\rho})] + D_2[\tilde{\lambda_3} + \lambda_5(\alpha - 1)] \right\} \hat{\mu}_t \\ & + \left[\frac{\lambda_5 \exp(c - y)}{1 - \exp(c - y)} (-\alpha + B_1 - B_2 \tilde{\rho}) - \tilde{\lambda}_3 - \lambda_5(\alpha - 1) \right] \sigma_A \hat{\epsilon}_{At+1} \\ & \equiv d_{ik} k_t + d_{i\mu} \hat{\mu}_t + d_{ie} \sigma_A \hat{\epsilon}_{At+1} \end{aligned} \tag{83}$$

The perceived risk exposure to aggregate productivity shocks (or the conditional volatility of equity return) is d_{ie}, and it is related to our extrapolative bias parameter $\tilde{\rho}$. In our calibrations, λ_5 is positive and B_2 is negative. So a higher $\tilde{\rho}$ increases the perceived risk exposure, and hence the risk premium, of the dividend claim.

g. Return predictability

Our return predictability can be understood by looking at the valuation ratio and investment return under the objective measure.

$$h_t = \tilde{A}_1 k_t + \tilde{A}_2 \hat{\mu}_t \tag{84}$$

$$
\begin{aligned}
r_{it+1} \approx & \left\{ \tilde{\lambda}_3 D_1 + (\lambda_4 + \lambda_5 D_1) \left[(\alpha - 1) + \frac{\alpha \exp(c-y)}{1-\exp(c-y)} - \frac{B_1 \exp(c-y)}{1-\exp(c-y)} \right] \right\} k_t \\
& + \left\{ \frac{\exp(c-y)}{1-\exp(c-y)} [-\lambda_4 B_2 + \lambda_5 \lambda_2 B_2 (\alpha - B_1) - \lambda_5 B_2 \rho] + \lambda_2 B_2 [\tilde{\lambda_3} + \lambda_5(\alpha - 1)] \right\} \hat{\mu}_t \\
& + \left[\frac{\lambda_5 \exp(c-y)}{1-\exp(c-y)} (-\alpha + B_1 - B_2 \tilde{\rho}) - \tilde{\lambda}_3 - \lambda_5 (\alpha - 1) \right] \sigma_A \epsilon_{At+1} \\
& \equiv \tilde{d}_{ik} k_t + \tilde{d}_{i\mu} \hat{\mu}_t + \tilde{d}_{ie} \sigma_A \epsilon_{At+1}
\end{aligned}
\tag{85}
$$

Under our calibration, $\tilde{d}_{i\mu}$ is negative and $\tilde{A}_2$ is positive. Therefore, when the perceived productivity growth is high, the valuation ratio is high. In fact, it is too much above the fundamental value that the future return needs to be negative to correct such an mispricing due to extrapolative bias. This explains why a higher valuation ratio (such as wealth-consumption ratio or price-dividend ratio) in the model is able to predict a lower future return, and it is different from time varying risk premium channel in the standard rational framework.

Figure 1: Impulse response under subjective and objective measures with extrapolative bias and without extrapolative bias

This figure plots the impulse response functions for the perceived (log) TFP level and perceived TFP growth by the agent, (log) consumption level, and perceived consumption growth by the agent to a positive one-standard-deviation TFP growth shock from the model under the subjective and objective measures with and without extrapolative bias. The impulse response functions are estimated from the log-linearization as in KL (2010) under the benchmark parameterization. The process of each variable is expressed using lag operators, and the coefficient of the order T in the Taylor expansion around zero represents the impulse response in quarter T. The top left panel plots the response function for the agent's perception of the (log) TFP level. The top right panel plots the response function for the perceived TFP growth by the agent. The bottom left panel presents the response function for the (log) consumption level, and the bottom right panel presents the response function for the perceived consumption growth by the agent. The solid and dashed lines are for result under the subjective and objective measures, respectively, for the specification where there is extrapolative bias. The dotted line is for the specification where there is no extrapolative bias.