A Online Appendix

A.1 Derivations

As Adrian et al. (2013), we assume that the systematic risk is represented by a stochastic vector, $(X_t)_{t\geq 0}$, that follows a stationary vector autoregression

$$X_t = \mu + \Phi X_{t-1} + v_t \tag{A.1}$$

with initial condition X_0 and whose residual terms, $(v_t)_{t\geq 0}$ follow a Gaussian distribution with variance-covariance matrix, Σ, i.e,.

$$v_t \big| \left(X_s\right)_{0\leq s\leq t} \sim \mathcal{N}\left(0, \Sigma\right). \tag{A.2}$$

Let's denote the zero coupon treasury bond price with maturity n at time t by $P_t^{(n)}$. We take the following assumptions:
Assumption 1. No-arbitrage condition holds (Dybvig and Ross, 1989), i.e.,

$$P_t^{(n)} = \mathbb{E}_t\left[M_{t+1}P_{t+1}^{n-1}\right]. \tag{A.3}$$

Assumption 2. The pricing kernel, $m_{t+1} := \log M_{t+1}$, is exponentially affine

$$m_{t+1} = -r_t^{(1)} - \frac{1}{2}||\lambda_t||^2 - \lambda_t^{\mathrm{T}}\Sigma^{-\frac{1}{2}}v_{t+1}, \tag{A.4}$$

where $r_t^{(1)} := -p_t^{(1)}$ is the continuously compounded risk-free rate, and $\lambda_t \in \mathbb{R}^K$.
Assumption 3. Market prices of risk are affine

$$\lambda_t = \Sigma^{-\frac{1}{2}}\left(\lambda_0 + \lambda_1 X_t\right), \tag{A.5}$$

where $\lambda_0 \in \mathbb{R}^K$ and $\lambda_1 \in \mathbb{R}^{K\times K}$.
Assumption 4. $\left(xr_t^{(n-1)}, v_t\right)_{t\geq 0}$ are jointly normally distributed for $n \geq 2$.
Thanks to all these assumptions, we can continue our modeling by recalling the definition of the excess holding return of a bond maturing in n periods, i.e.,

$$xr_{t+1}^{(n-1)} := p_{t+1}^{(n-1)} - p_t^{(n)} - r_t^{(1)}, \tag{A.6}$$

where $n-1$ indicates the $n-1$ periods remaining since time $t+1$ with respect to which the return is computed. Now, (A.3) can be rewritten as

$$\begin{aligned} 1 &= \mathbb{E}_t\left[\exp\left\{m_{t+1} + p_{t+1}^{(n-1)} - p_t^{(1)}\right\}\right] \\ &= \mathbb{E}_t\left[\exp\left\{-r_t^{(1)} - \frac{1}{2}||\lambda_t||^2 - \lambda_t^{\mathrm{T}}\Sigma^{-\frac{1}{2}}v_{t+1} + xr_{t+1}^{(n)} + r_t^{(1)}\right\}\right] \\ &= \mathbb{E}_t\left[\exp\left\{xr_{t+1}^{(n)} - \frac{1}{2}||\lambda_t||^2 - \lambda_t^{\mathrm{T}}\Sigma^{-\frac{1}{2}}v_{t+1}\right\}\right] \\ &= \exp\left\{\mathbb{E}_t\left[\xi_{t+1}\right] + \frac{1}{2}\mathbb{V}\left[\xi_{t+1}\right]\right\}, \end{aligned} \tag{A.7}$$

where $\xi_{t+1} \coloneqq xr_{t+1}^{(n)} - \frac{1}{2}||\lambda_t||^2 - \lambda_t^{\mathrm{T}}\Sigma^{-\frac{1}{2}}v_{t+1}$, and

$$\mathbb{E}_t\left[\xi_{t+1}^{(n-1)}\right] = \mathbb{E}_t\left[xr_{t+1}^{(n-1)}\right] - \frac{1}{2}||\lambda_t||^2 \tag{A.8}$$

$$\begin{aligned} \mathbb{V}_t\left[\xi_{t+1}^{(n-1)}\right] &= \mathbb{V}_t\left[xr_{t+1}^{(n-1)} - \lambda_t^{\mathrm{T}}\Sigma^{-\frac{1}{2}}v_{t+1}\right] \\ &= \mathbb{V}_t\left[xr_{t+1}^{(n-1)}\right] + \mathbb{V}_t\left[\lambda_t^{\mathrm{T}}\Sigma^{-\frac{1}{2}}v_{t+1}\right] - 2\mathrm{cov}\left(xr_{t+1}^{(n-1)}, \lambda_t^{\mathrm{T}}\Sigma^{-\frac{1}{2}}v_{t+1}\right) \\ &= \mathbb{V}_t\left[xr_{t+1}^{(n-1)}\right] + \lambda_t^{\mathrm{T}}\Sigma^{-\frac{1}{2}}\mathbb{V}_t\left[v_{t+1}\right]\Sigma^{-\frac{1}{2}}\lambda_t - 2\lambda_t^{\mathrm{T}}\Sigma^{-\frac{1}{2}}\mathrm{cov}_t\left(xr_{t+1}^{(n-1)}, v_{t+1}\right) \\ &= \mathbb{V}_t\left[xr_{t+1}^{(n-1)}\right] + ||\lambda_t||^2 - 2\lambda_t^{\mathrm{T}}\Sigma^{\frac{1}{2}}\beta_t^{(n-1)}. \end{aligned} \tag{A.9}$$

where

$$\beta_t^{(n-1)} \coloneqq \Sigma^{-1}\mathrm{cov}_t\left(xr_{t+1}^{(n-1)}, v_{t+1}\right) \in \mathbb{R}^K. \tag{A.10}$$

Therefore, no-arbitrage condition (A.3) is equivalent to

$$0 = \mathbb{E}_t\left[xr_{t+1}^{(n-1}\right] + \frac{1}{2}\mathbb{V}_t\left[xr_{t+1}^{(n)}\right] - \lambda_t^{\mathrm{T}}\Sigma^{\frac{1}{2}}\beta_t^{(n-1)}, \tag{A.11}$$

which gives us the following expression for the expected returns:

$$E_t\left[xr_{t+1}^{(n-1)}\right] = \lambda_t^{\mathrm{T}}\Sigma^{\frac{1}{2}}\beta_t^{(n-1)} - \frac{1}{2}\mathbb{V}_t\left[xr_{t+1}^{(n)}\right]. \tag{A.12}$$

Assumption 5. $\beta_t^{(n)} = \beta^{(n)}$ for every $t \geq 0$.

If we were to decompose the unexpected excess return, $xr_{t+1}^{(n-1)} - \mathbb{E}_t\left[xr_{t+1}^{(n-1)}\right]$ into a component that is correlated with v_{t+1} and another component which is conditionally orthogonal, $\varepsilon_{t+1}^{(n-1)}$ (return pricing error), we could simply write the following OLS-wise form

$$xr_{t+1}^{(n-1)} - \mathbb{E}_t\left[xr_{t+1}^{(n-1)}\right] = v_{t+1}^{\mathrm{T}}\gamma^{(n-1)} + \varepsilon_{t+1}^{(n-1)}. \tag{A.13}$$

and try to figure out who the $\gamma^{(n-1)}$ is. To do so, notice that

$$\beta_t^{(n-1)} = \Sigma^{-1}\left(\mathbb{E}\left[xr_{t+1}^{(n-1)}v_{t+1}\right] - \mathbb{E}\left[xr_{t+1}^{(n-1)}\right]\mathbb{E}_t\left[v_{t+1}\right]\right) = \Sigma^{-1}\mathbb{E}\left[xr_{t+1}^{(n-1)}v_{t+1}\right]$$

and

$$\gamma^{(n-1)} = \left(\mathbb{E}\left[v_{t+1}^{\mathrm{T}}v_{t+1}\right]\right)^{-1}\mathbb{E}\left[v_{t+1}xr_{t+1}^{(n-1)}\right] = \Sigma^{-1}\mathbb{E}\left[xr_{t+1}^{(n-1)}v_{t+1}\right],$$

because $\mathbb{E}\left[v_{t+1}^{\mathrm{T}}v_{t+1}\right] = \Sigma$. Therefore, $\gamma^{(n)} = \beta^{(n)}$ for every $n \geq 0$. With this identity in our hands,

$$\begin{aligned}
\mathbb{V}\left[xr_{t+1}^{(n-1)}\right] &= \mathbb{E}_t\left[\left(xr_{t+1}^{(n-1)} - \mathbb{E}_t\left[xr_{t+1}^{(n-1)}\right]\right)^2\right] \\
&= \mathbb{E}_t\left[\left(v_{t+1}^{\mathrm{T}}\beta^{(n-1)} + \varepsilon_{t+1}^{n-1}\right)^2\right] \\
&= \mathbb{E}_t\left[\left(v_{t+1}^{\mathrm{T}}\beta^{(n-1)}\right)^2 + 2v_{t+1}^{\mathrm{T}}\beta^{(n-1)}\varepsilon_{t+1}^{(n-1)} + \left(\varepsilon_{t+1}^{(n-1)}\right)^2\right] \\
&= \left(\beta^{(n-1)}\right)^{\mathrm{T}}\mathbb{E}_t\left[v_{t+1}v_{t+1}^{\mathrm{T}}\right]\beta^{(n-1)} + \sigma^2 \\
&= \left(\beta^{(n-1)}\right)^{\mathrm{T}}\Sigma\beta^{(n-1)} + \sigma^2,
\end{aligned}$$

Finally,

$$\begin{aligned}
xr_{t+1}^{(n-1)} = (\lambda_0 + \lambda_1 X_t)^{\mathrm{T}}\beta^{(n-1)} - \frac{1}{2}\left(\left(\beta^{(n-1)}\right)^{\mathrm{T}}\Sigma\beta^{(n-1)} + \sigma^2\right) \\
+ v_{t+1}^{\mathrm{T}}\beta^{(n-1)} + \varepsilon_{t+1}^{(n-1)}.
\end{aligned} \tag{A.14}$$

A.2 Estimation

We can then rewrite (A.14) as

$$xr_{t+1}^{(n-1)} = (\lambda_0 + \lambda_1 X_t)^{\mathrm{T}} B_{n-1} - \frac{1}{2}\left(B_{n-1}^{\mathrm{T}}\Sigma B_{n-1} + \sigma^2\right) + v_{t+1}^{\mathrm{T}}B_n + e_{t+1}^{(n-1)} \tag{A.15}$$

and therefore have a vectorial form:

$$\mathbf{xr} = \left(\lambda_0 \mathbb{1}_{T\times 1}^{\mathrm{T}} + \lambda_1 \mathbf{X}_{-}^{\mathrm{T}}\right)^{\mathrm{T}}\mathbf{B} - \frac{1}{2}\left(\mathbf{B}^{*}\mathrm{vec}\left(\Sigma\right) + \sigma^2\mathbb{1}_{K\times 1}\right)\mathbb{1}_{T}^{\mathrm{T}} + \mathbf{V}^{\mathrm{T}}\mathbf{B} + \mathbf{E} \tag{A.16}$$

where

1. $\mathbf{xr} \in \mathbb{R}^{T\times N}$.
2. $\lambda_0 \in \mathbb{R}^{K}$, $\lambda_1 \in \mathbb{R}^{K\times K}$,
3. $\mathbf{X}_{-} = [X_1 \mid X_2 \mid \cdots \mid X_{T-1}]^{\mathrm{T}} \in \mathbb{R}^{T\times K}$,
4. $\mathbf{B} \in \mathbb{R}^{K\times N}$,
5. $\mathbf{B}^{*} = \left[\mathrm{vec}\left(B_1 B_1^{\mathrm{T}}\right) \mid \cdots \mid \mathrm{vec}\left(B_n B_n^{\mathrm{T}}\right)\right]^{\mathrm{T}} \in \mathbb{R}^{K^2\times N}$,
6. $\mathbf{V} \in \mathbb{R}^{T\times K}$ and $\mathbf{E} \in \mathbb{R}^{T\times N}$.

So we take (A.16) as our reference point in the estimation process that we do in four stepsby extending Adrian et al. (2013) procedure:

1. Construct the pricing factors $(X_t)_{t=1}^{T}$. First, model the trend in the one-period (three-month) rate is captured by projecting it on the proxy for the age structure of the population, potential output growth and the survey-based measure of long-run inflation expectations.

Second, derive the cyclical components of yields at any maturity by considering the difference between yields and the trend in the three-month rate. Third consider as price factors the first k principal components of de-trended yields.

2. Model the pricing factors, $(X_t)_{t=1}^T$ via a VAR and estimate the VAR coefficients $\mu \in \mathbb{R}^K$ and $\Phi \in \mathbb{R}^K$ in (A.1) using OLS. Then take $(\hat{v}_t)_{t=1}^T$ from $\hat{v}_t := X_t - \hat{X}_t \in \mathbb{R}^K$, where $\hat{X}_t = \mu + \Phi X_{t-1}$ for every $t = 1, \dots, T$. Stack the time series $(v_t)_{t=1}^T$ into the matrix $\hat{\mathbf{V}} \in \mathbb{R}^{T\times K}$. The variance-covariance matrix is thus

$$\hat{\Sigma} = \frac{\hat{\mathbf{V}}^{\mathrm{T}}\hat{\mathbf{V}}}{T} \tag{A.17}$$

3. Perform the regression according to (A.16), i.e.,

$$\mathbf{xr} = a\mathbb{1}_{T\times K}\mathbb{1}_{K\times N} + \hat{\mathbf{V}}b + \mathbf{X}_{-}c + \mathbf{E} \tag{A.18}$$

where $a \in \mathbb{R}$, $b, c \in \mathbb{R}^{K\times N}$. Collect everything into single matrices

$$\mathbf{Z} = \left[\mathbb{1}_{T\times 1} \mid \hat{\mathbf{V}} \mid \mathbf{X}_{-}\right] \in \mathbb{R}^{T\times(2K+1)} \tag{A.19}$$

$$d = [a\mathbb{1}_{K\times 1} \mid b \mid c]^{\mathrm{T}} \in \mathbb{R}^{(2K+1)\times N} \tag{A.20}$$

so we can write $\mathbf{xr} = \mathbf{Z}d + \mathbf{E}$ and therefore

$$\hat{d} = \left(\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\right)^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{xr}. \tag{A.21}$$

Then, collect the residuals from this regression into the matrix

$$\hat{\mathbf{E}} = \mathbf{xr} - \mathbf{Z}\hat{d} \in \mathbb{R}^{T\times N}. \tag{A.22}$$

and estimate

$$\hat{\sigma}^2 = \frac{\operatorname{tr}\left(\hat{\mathbf{E}}^{\mathrm{T}}\hat{\mathbf{E}}\right)}{NT}. \tag{A.23}$$

Finally, we construct $\hat{\mathbf{B}}^*$ from $\hat{b}$.

4. Estimate the price of risk parameters, λ_0 and λ_1 via cross-sectional regression. Recall from (A.16) that

$$a = \left(\lambda_0\mathbb{1}_{T\times 1}^{\mathrm{T}}\right)^{\mathrm{T}}\mathbf{B} - \frac{1}{2}\left(\mathbf{B}^*\mathrm{vec}\left(\Sigma\right) + \sigma^2\mathbb{1}_{K\times 1}\right)\mathbb{1}_T^{\mathrm{T}} \tag{A.24}$$

$$c = \lambda_1^{\mathrm{T}}\mathbf{B} \tag{A.25}$$

If we transpose them, we can estimate λ_0 and λ_1 via OLS, i.e.,

$$\hat{\lambda}_0 = \left(\hat{\mathbf{B}}\hat{\mathbf{B}}^{\mathrm{T}}\right)^{-1}\hat{\mathbf{B}}\left[\hat{a}^{\mathrm{T}} + \frac{1}{2}\mathbb{1}_{T\times 1}\left(\mathbf{B}^*\mathrm{vec}\left(\Sigma\right) + \sigma^2\mathbb{1}_{N\times 1}\right)^{\mathrm{T}}\right] \tag{A.26}$$

$$\hat{\lambda}_1 = \left(\hat{\mathbf{B}}\hat{\mathbf{B}}^{\mathrm{T}}\right)^{-1}\hat{\mathbf{B}}\hat{c}^{\mathrm{T}} \tag{A.27}$$

A.3 Recursion for the Term Structure

Consider the generating process for log excess returns in our model:

$$xr_{t+1}^{(n-1)} = (\lambda_0 + \lambda_1 X_t)^{\mathrm{T}} \beta^{(n-1)} - \frac{1}{2}\left(\left(\beta^{(n-1)}\right)^{\mathrm{T}} \Sigma \beta^{(n-1)} + \sigma^2\right) + v_{t+1}^{\mathrm{T}} \beta^{(n-1)} + \varepsilon_{t+1}^{(n-1)}. \tag{A.28}$$

We need now to find two sequences of coefficients, $(A_n)_{n=1}^N$ and $(B_n)_{n=1}^N$, that allow us to express bond prices as exponentially affine in the vector of state variables, X_t, plus a trend term, $p_t^{*,(n)}$, i.e.,

$$p_t^{(n)} = p_t^{*,(n)} + A_n + X_t^{\mathrm{T}} B_n + e_t^{(n)}, \tag{A.29}$$

where $p_t^{(n)} \coloneqq \log P_t^{(n)}$. Notice that

$$p_t^{(1)} = -r_t^{(1)} = -r_t^{*,(1)} - \delta_0 - X_t^{\mathrm{T}} \delta_1, \tag{A.30}$$

motivating that $A_1 = -\delta_0$, $B_1 = -\delta_1$, and $p_t^{1,*} = -r_t^{*,(1)}$. For any $n > 1$,

$$\begin{aligned}
xr_{t+1}^{(n-1)} &= p_{t+1}^{*,(n-1)} + A_{n-1} + X_{t+1}^{\mathrm{T}} B_{n-1} + e_{t+1}^{(n-1)} \\
&- p_t^{*,(n)} - A_n - X_t^{\mathrm{T}} B_n - e_t^{(n)} \\
&+ p_t^{*,(1)} + A_1 + X_t^{\mathrm{T}} B_1 + e_t^{(1)} \\
&= p_{t+1}^{*,(n-1)} + A_{n-1} + (\mu + \Phi X_t + v_{t+1})^{\mathrm{T}} B_{n-1} + e_{t+1}^{(n-1)} \\
&- p_t^{*,(n)} - A_n - X_t^{\mathrm{T}} B_n - e_t^{(n)} \\
&+ p_t^{*,(1)} + A_1 + X_t^{\mathrm{T}} B_1 + e_t^{(1)} \\
&= xr_{t+1}^{*,(n-1)} + (A_{n-1} - A_n + A_1 + \mu^{\mathrm{T}} B_{n-1}) \\
&+ X_t^{\mathrm{T}} \left(\Phi^{\mathrm{T}} B_{n-1} - B_n + B_1\right) + \left(e_{t+1}^{n-1} - e_t^{(n)} + e_t^{(1)}\right) + v_{t+1}^{\mathrm{T}} B_{n-1}.
\end{aligned} \tag{A.31}$$

Hence, the following must hold

$$\begin{aligned}
&xr_{t+1}^{*,(n-1)} + (A_{n-1} - A_n + A_1 + \mu^{\mathrm{T}} B_{n-1}) \\
&+X_t^{\mathrm{T}} \left(\Phi^{\mathrm{T}} B_{n-1} - B_n + B_1\right) + \left(e_{t+1}^{n-1} - e_t^{(n)} + e_t^{(1)}\right) \\
&=(\lambda_0 + \lambda_1) X_t^{\mathrm{T}} \beta^{(n-1)} - \frac{1}{2}\left(\left(\beta^{(n-1)}\right)^{\mathrm{T}} \Sigma \beta^{(n-1)} + \sigma^2\right) + v_{t+1} \beta^{(n-1)} + \varepsilon_{t+1}^{(n-1)}
\end{aligned}$$

i.e.,

$$\begin{aligned}
A_{n-1} - A_n + A_1 + \mu^{\mathrm{T}} B_{n-1} &= \lambda_0^{\mathrm{T}} \beta^{(n-1)} - \frac{1}{2}\left(\left(\beta^{(n-1)}\right)^{\mathrm{T}} \Sigma \beta^{(n-1)} + \sigma^2\right) \\
\Phi^{\mathrm{T}} B_{n-1} - B_n + B_1 &= \lambda_1^{\mathrm{T}} \beta^{(n-1)} \\
u_{t+1}^{n-1} - u_t^{(n)} + u_t^{(1)} + v_{t+1}^{\mathrm{T}} B_{n-1} &= \varepsilon_{t+1}^{(n-1)} \\
xr_{t+1}^{*,(n-1)} &= 0 \\
v_{t+1}^{\mathrm{T}} \beta^{(n-1)} &= v_{t+1}^{\mathrm{T}} B_{n-1}
\end{aligned}$$

and therefore

$$\begin{aligned}
A_n &= A_{n-1} + \mu^{\mathrm{T}} B_{n-1} - \lambda_0^{\mathrm{T}} \beta^{(n-1)} + \frac{1}{2}\left(\left(\beta^{(n-1)}\right)^{\mathrm{T}} \Sigma \beta^{(n-1)} + \sigma^2\right) + A_1 \\
B_n &= \Phi^{\mathrm{T}} B_{n-1} + B_1 - \lambda_1^{\mathrm{T}} \beta^{(n-1)} \\
p_t^{*,(n)} &= p_{t+1}^{*,(n-1)} - r_t^{*,(1)} \\
\beta^{(n)} &= B_n
\end{aligned}$$

The last equation simplifies everything even more:

$$A_n = A_{n-1} + (\mu - \lambda_0)^{\mathrm{T}} B_{n-1} + \frac{1}{2}\left(B_{n-1}^{\mathrm{T}} \Sigma B_{n-1} + \sigma^2\right) - \delta_0 \tag{A.32}$$

$$B_n = (\Phi - \lambda_1)^{\mathrm{T}} B_{n-1} - \delta_1 \tag{A.33}$$

$$p_t^{(n),*} = p_{t+1}^{(n-1),*} - r_t^{*,(1)} \tag{A.34}$$

Equation (A.34) for the price stochastic trend implies that

$$r_t^{*,(n)} = \frac{1}{n} \sum_{i=0}^{n-1} r_{t+i}^{*,(1)}. \tag{A.35}$$

On the other hand, these equations are fully deterministic, meaning that one can iterate all the equations back to get expressions that depend only on the initial values, A_1 and B_1. First,

$$\begin{aligned}
B_n &= (\Phi - \lambda_1)^{\mathrm{T}} \left((\Phi - \lambda_1)^{\mathrm{T}} B_{n-2} - \delta_1\right) - \delta_1 \\
&= \cdots \\
&= \left[(\Phi - \lambda_1)^{\mathrm{T}}\right]^{n-1} B_1 - \sum_{j=1}^{n-2} \left[(\Phi - \lambda_1)^{\mathrm{T}}\right]^j \delta_1. \qquad \text{(A.36)} \\
&= -\sum_{j=1}^{n-1} \left[(\Phi - \lambda_1)^{\mathrm{T}}\right]^j \delta_1
\end{aligned}$$

Second,

$$\begin{aligned}
A_n &= A_{n-2} + (\mu - \lambda_0)^{\mathrm{T}} (B_{n-1} + B_{n-2}) + \frac{1}{2}\left(B_{n-1}^{\mathrm{T}} \Sigma B_{n-1} + B_{n-2}^{\mathrm{T}} \Sigma B_{n-2}\right) + 2\left(\frac{1}{2}\sigma^2 - \delta_0\right) \\
&= A_{n-2} + (\mu - \lambda_0)^{\mathrm{T}} (B_{n-1} + B_{n-2}) \\
&\quad + \frac{1}{2}\left([B_{n-1} + B_{n-2}]^{\mathrm{T}} \Sigma [B_{n-1} + B_{n-2}]\right) + 2\left(\frac{1}{2}\sigma^2 - \delta_0\right) \\
&= A_1 + (\Phi - \lambda_1)^{\mathrm{T}} \sum_{j=1}^{n-1} B_{n-j} + \frac{1}{2}\left(\sum_{j=1}^{n-1} B_{n-j}\right)^{\mathrm{T}} \Sigma \left(\sum_{j=1}^{n-1} B_{n-j}\right) + (n-1)\left(\frac{1}{2}\sigma^2 - \delta_0\right)
\end{aligned}$$

It's not difficult to see that

$$\sum_{j=1}^{n-1} B_{n-j} = \sum_{j=1}^{n-1}\sum_{k=1}^{n-j} \left[(\Phi - \lambda_1)^{\mathrm{T}}\right]^j \delta_1 = \sum_{j=1}^{n-1}(n-j)\left[(\Phi - \lambda_1)^{\mathrm{T}}\right]^j \delta_1. \tag{A.37}$$

That allows us to write

$$\begin{aligned} A_n = {} & (\Phi - \lambda_1)^{\mathrm{T}} \sum_{j=1}^{n-1}(n-j)\left[(\Phi - \lambda_1)^{\mathrm{T}}\right]^j \\ & + \frac{1}{2}\left(\sum_{j=1}^{n-1}(n-j)(\Phi - \lambda_1)^j\right)\Sigma\left(\sum_{j=1}^{n-1}(n-j)\left[(\Phi - \lambda_1)^{\mathrm{T}}\right]^j\right) \\ & + n\left(\frac{1}{2}\sigma^2 - \delta_0\right). \end{aligned} \tag{A.38}$$

A.4 Recursion for Term Premia

Remember that

$$TP_t^{(n)} = u_t^{(n)} - \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_t\left[u_{t+i}^{(1)}\right], \tag{A.39}$$

where $u_t^{(n)} = r_t^{(n)} - r_t^{*,(n)}$. The affine model implies that

$$u_t^{(n)} = -n\left(A_n + X_t^{\mathrm{T}} B_n + e_t^{(n)}\right). \tag{A.40}$$

In particular, for $n = 1$,

$$u_t^{(1)} = -A_1 - X_t^{\mathrm{T}} B_1 - e_t^{(1)}. \tag{A.41}$$

Hence,

$$\mathbb{E}_t\left[u_{t+i}^{(1)}\right] = -A_1 - \mathbb{E}_t\left[X_{t+i}^{\mathrm{T}}\right] B_1. \tag{A.42}$$

Now, since $X_{t+i} = \mu + \Phi X_{t+i-1} + v_{t+i}$, then, we can iterate backwards to get

$$\begin{aligned} X_{t+i} &= \mu + \Phi X_{t+i-1} + v_{t+i} \\ &= \mu + \Phi\left(\mu + \Phi X_{t+i-2} + v_{t+i-1}\right) + v_{t+i} \\ &= (1+\Phi)\mu + \Phi^2 X_{t+i-2} + \Phi v_{t+i-1} + v_{t+i} \\ &= \cdots \\ &= \left(\sum_{j=0}^{i-1}\Phi^j\right)\mu + \Phi^i X_t + \sum_{j=0}^{i-1}\Phi^j v_{t+i-j}. \end{aligned} \tag{A.43}$$

Since $\mathbb{E}_t[v_s] = 0$ for every $s > t$, then

$$\mathbb{E}_t[X_{t+i}] = \widetilde{\Phi}_i\mu + \Phi^i X_t, \tag{A.44}$$

where

$$\widetilde{\Phi}_i = \left(\sum_{j=0}^{i-1} \Phi^j \right). \tag{A.45}$$

Hence,

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_t \left[u_t^{(1)} \right] &= -A_1 - \frac{1}{n} \sum_{i=1}^{n} \left(\widetilde{\Phi}_i \mu + \Phi^i X_t \right)^{\mathrm{T}} B_1 \\
&= -A_1 - \frac{1}{n} B_1^{\mathrm{T}} \left(\sum_{i=1}^{n} \widetilde{\Phi}_i \right) \mu - \frac{1}{n} B_1^{\mathrm{T}} \left(\sum_{i=1}^{n} \Phi^i \right) X_t \\
&= -A_1 - \frac{1}{n} B_1^{\mathrm{T}} \left(\sum_{i=1}^{n} \widetilde{\Phi}_i \right) \mu - \frac{1}{n} B_1^{\mathrm{T}} \widetilde{\Phi}_n X_t \\
&= \Xi_n + \Psi_n X_t
\end{aligned} \tag{A.46}$$

where

$$\Xi_n = -\frac{1}{n} A_1 - \frac{1}{n} B_1^{\mathrm{T}} \left(\sum_{i=1}^{n} \widetilde{\Phi}_i \right) \mu \tag{A.47}$$

$$\Psi_n = -\frac{1}{n} B_1^{\mathrm{T}} \widetilde{\Phi}_i \tag{A.48}$$

Hence,

$$TP_t^{(n)} = u_t^{(n)} + \Xi_n + \Psi_n X_t \tag{A.49}$$