

Introduction to Term Structure Models

Jean-Paul Renne and Alain Monfort

2024-01-21

Contents

1	Affine processes	7
1.1	Information in the Economy: The “factors”	7
1.2	Dynamic models and Laplace transform (L.T.)	8
1.3	Laplace Transform and moments/cumulants	9
1.4	Additional properties of the Laplace transform	10
1.5	Affine (or Car) processes	12
1.6	Markov chains	18
1.7	Wishart autoregressive (WAR) processes	19
1.8	Building affine processes	20
1.9	Multi-horizon Laplace transform	26
1.10	VAR representation and conditional moments	31
1.11	Truncated Laplace transforms of affine processes	32
1.12	Appendices	34
2	Pricing and risk-neutral dynamics	43
2.1	Consumption-based Capital Asset Pricing Model (CCAPM) and stochastic discount factor (SDF)	43
2.2	Recursive utilities	45
2.3	SDF: Absence of Arbitrage Approach	45
2.4	The risk-neutral (R.N.) dynamics	48
2.5	Typology of Econometric Asset Pricing Models	50

3	The term structure or risk-free yields	55
3.1	Introduction	55
3.2	Swap rates and forward rates	57
3.3	The Affine Case	59
3.4	Gaussian Affine Term Structure Model	61
3.5	Non-Negative Affine Term Structure Model	66
4	Estimation of affine asset-pricing models	75
4.1	State-Space Model	75
4.2	Kalman-Filter-Based Approach	77
4.3	About non-constant conditional matrix Σ	88
4.4	Inversion Technique	89
4.5	A Typical Small-Sample Issue	96
5	Forward, futures, dividends, commodity pricing, and convenience yields	99
5.1	Forward Contracts	99
5.2	Futures Contracts, Futures Prices	101
5.3	Convenience Yields	102
5.4	Pricing with Affine Models	104
5.5	Historical dynamics	104
5.6	A Gaussian VAR Model	105
5.7	Applications	107
6	References	109

Introduction to Term Structure Models

Modeling dynamic term structures serves as a practical and indispensable tool in the realm of finance. It enables investors, institutions, and policymakers to make informed decisions, manage risk effectively, and allocate resources wisely. By understanding how interest rates and yields evolve over time, these models offer a clear lens through which to assess market trends and price financial instruments accurately.

This course has been developed by Jean-Paul Renne and Alain Monfort. It is illustrated by R codes using various packages that can be obtained from CRAN. This `TSMmodels` package is available on GitHub. To install it, one needs to employ the `devtools` library:

```
install.packages("devtools") # in case this library has not been loaded yet
library(devtools)
install_github("jrenne/TSMmodels")
library(AEC)
```

Useful (R) links:

- Download R:
 - R software: <https://cran.r-project.org> (the basic R software)
 - RStudio: <https://www.rstudio.com> (a convenient R editor)
- Tutorials:

- Rstudio: <https://dss.princeton.edu/training/RStudio101.pdf> (by Oscar Torres-Reyna)
- R: https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf (by Emmanuel Paradis)
- My own tutorial: https://jrenne.shinyapps.io/Rtuto_publiShiny/

Chapter 1

Affine processes

1.1 Information in the Economy: The “factors”

On each date $t = 1, 2, \dots, T$, agents receive new information by observing *factors*, also called *states*. We denote the (K -dimensional) vector of factors by w_t . Vector w_t is usually random. On date t , vector w_t is supposed to be perfectly observed by the agents (investors), but can be only partially observed, or unobserved by the econometrician.

Naturally, w_t can be decomposed into different sub-vectors of different natures. For instance, we can have $w_t = (y'_t, z'_t)'$ with y_t : observable vector of (geometric) returns, z_t : regime, unobserved by the econometrician.

Some of the components of w_t can be prices. For instance, one component could be a short-term rate, a stock return, or an exchange rate. It can also include macroeconomic variables (inflation, GDP growth), or agent-specific variables.

1.2 Dynamic models and Laplace transform (L.T.)

The objective of a dynamic model is to describe the random changes in w_t . The dynamics can be historical or risk-neutral (see Chapter ??). The dynamics we consider are parametric, in the sense that the conditional distribution $w_{t+1}|\underline{w}_t$ (with $\underline{w}_t = \{w_t, w_{t-1}, \dots\}$) depends on a vector of parameters θ . In practice, it may be the case that θ is unknown by the econometrician (see Chapter ??). The choice (or estimation) of a conditional distribution is equivalent to the choice (or estimation) of a conditional *Laplace transforms*:

$$\varphi(u|\underline{w}_t, \theta) = \mathbb{E}_\theta[\exp(u'w_{t+1})|\underline{w}_t], \quad u \in \mathbb{R}^K, \quad (1.1)$$

or a conditional *log Laplace transforms*:

$$\psi(u|\underline{w}_t, \theta) = \log\{\mathbb{E}_\theta[\exp(u'w_{t+1})|\underline{w}_t]\}, \quad u \in \mathbb{R}^K.$$

Example 1.1 (Conditionally Bernoulli process). If $w_{t+1}|\underline{w}_t \sim \mathcal{I}[p(\underline{w}_t, \theta)]$, then:

$$\varphi(u|w_t) = \mathbb{E}[\exp(uw_{t+1}) | \underline{w}_t] = p_t \exp(u) + 1 - p_t$$

with $p_t = p(\underline{w}_t, \theta)$.

Example 1.2 (Conditionally Binomial process). If $w_{t+1}|\underline{w}_t \in \mathcal{B}(n, p_t)$, then:

$$\varphi(u|w_t) = [p_t \exp(u) + 1 - p_t]^n.$$

Example 1.3 (Conditionally Poisson process). If $w_{t+1}|\underline{w}_t \sim \mathcal{P}(\lambda_t)$, then:

$$\begin{aligned} \varphi(u|w_t) &= \sum_{j=0}^{\infty} \frac{1}{j!} \exp(-\lambda_t) \lambda_t^j \exp(uj) = \exp(-\lambda_t) \exp[\lambda_t \exp(u)] \\ &= \exp\{\lambda_t [\exp(u) - 1]\}. \end{aligned}$$

Example 1.4 (Conditionally normal (or Gaussian) process). If $w_{t+1}|\underline{w}_t \sim \mathcal{N}(m(\underline{w}_t, \theta), \Sigma(\underline{w}_t, \theta))$, then:

$$\begin{aligned} u'w_{t+1}|\underline{w}_t &\sim \mathcal{N}(u'm(\underline{w}_t, \theta), u'\Sigma(\underline{w}_t, \theta)u). \\ \Rightarrow \begin{cases} \varphi(u|\underline{w}_t, \theta) &= \exp \left[u'm(\underline{w}_t, \theta) + \frac{1}{2}u'\Sigma(\underline{w}_t, \theta)u \right] \\ \psi(u|\underline{w}_t, \theta) &= u'm(\underline{w}_t, \theta) + \frac{1}{2}u'\Sigma(\underline{w}_t, \theta)u. \end{cases} \end{aligned}$$

1.3 Laplace Transform and moments/cumulants

Here are some properties of the Laplace transform (Eq. (1.1)):

- $\varphi(0|\underline{w}_t, \theta) = 1$ and $\psi(0|\underline{w}_t, \theta) = 0$.
- It is defined in a convex set E (containing 0).
- If the interior of E is non empty, all the (conditional) moments exist.

As mentioned above, knowing the (conditional) Laplace transform is equivalent to knowing the (conditional) moments—if they exist. In the scalar case, we have that:

- the moment of order n closely relates to the n^{th} derivatives of φ :

$$\left[\frac{\partial^n \varphi(u|\underline{w}_t, \theta)}{\partial u^n} \right]_{u=0} = \mathbb{E}_\theta[w_{t+1}^n | \underline{w}_t],$$

- the cumulant of order n closely relates to the n^{th} derivatives of ψ :

$$\left[\frac{\partial^n \psi(u|\underline{w}_t, \theta)}{\partial u^n} \right]_{u=0} = K_n(\underline{w}_t, \theta).$$

In particular, what precedes implies that:

$$\begin{cases} K_1(\underline{w}_t, \theta) &= \mathbb{E}_\theta[w_{t+1} | \underline{w}_t] \\ K_2(\underline{w}_t, \theta) &= \mathbb{V}ar_\theta[w_{t+1} | \underline{w}_t]. \end{cases}$$

Accordingly, φ and ψ are respectively called conditional **moment** and **cumulant** generating function.

In the multivariate case, we have:

$$\begin{aligned} \left[\frac{\partial \psi}{\partial u}(u|\underline{w}_t, \theta) \right]_{u=0} &= \mathbb{E}_\theta[w_{t+1} | \underline{w}_t] \\ \left[\frac{\partial^2 \psi}{\partial u \partial u'}(u|\underline{w}_t, \theta) \right]_{u=0} &= \mathbb{V}ar_\theta[w_{t+1} | \underline{w}_t]. \end{aligned}$$

Example 1.5 (Conditionally normal (or Gaussian) process). Consider Example 1.5. Applying the previous formula, we have, in the scalar case:

- $\psi(u|\underline{w}_t, \theta) = um(\underline{w}_t, \theta) + \frac{1}{2}u^2\sigma^2(\underline{w}_t, \theta).$
- $\left[\frac{\partial\psi}{\partial u}(u|\underline{w}_t, \theta) \right]_{u=0} = m(\underline{w}_t, \theta).$
- $\left[\frac{\partial^2\psi}{\partial u^2}(u|\underline{w}_t, \theta) \right]_{u=0} = \sigma^2(\underline{w}_t, \theta).$

and, in the multidimensional normal case:

- $\psi(u|\underline{w}_t, \theta) = u'm(\underline{w}_t, \theta) + \frac{1}{2}u'\Sigma(\underline{w}_t, \theta)u.$
- $\left[\frac{\partial\psi}{\partial u}(u|\underline{w}_t, \theta) \right]_{u=0} = m(\underline{w}_t, \theta).$
- $\left[\frac{\partial^2\psi}{\partial u\partial u'}(u|\underline{w}_t, \theta) \right]_{u=0} = \Sigma(\underline{w}_t, \theta).$

In both cases, cumulants of order > 2 equal to 0.

1.4 Additional properties of the Laplace transform

Here are additional properties of multivariate Laplace transform:

- If $w_t = (w'_{1t}, w'_{2t})'$, $u = (u'_1, u'_2)'$:

$$\begin{aligned}\mathbb{E}_\theta[\exp(u'_1 w_{1,t+1}|\underline{w}_t)] &= \varphi(u_1, 0|\underline{w}_t, \theta) \\ \mathbb{E}_\theta[\exp(u'_2 w_{2,t+1}|\underline{w}_t)] &= \varphi(0, u_2|\underline{w}_t, \theta).\end{aligned}$$

- If $w_t = (w'_{1t}, w'_{2t})'$, and if w_{1t} and w_{2t} are conditionally independent:

$$\begin{aligned}\varphi(u|\underline{w}_t, \theta) &= \varphi(u_1, 0|\underline{w}_t, \theta) \times \varphi(0, u_2|\underline{w}_t, \theta) \\ \psi(u|\underline{w}_t, \theta) &= \psi(u_1, 0|\underline{w}_t, \theta) + \psi(0, u_2|\underline{w}_t, \theta).\end{aligned}$$

- If w_{1t} and w_{2t} have the same size and if

$$\varphi(u_1, u_2|\underline{w}_t, \theta) = \mathbb{E}_\theta[\exp(u'_1 w_{1,t+1} + u'_2 w_{2,t+1}|\underline{w}_t),$$

1.4. ADDITIONAL PROPERTIES OF THE LAPLACE TRANSFORM 11

then the conditional Laplace transform of $w_{1,t+1} + w_{2,t+1}$ given \underline{w}_t is $\varphi(u, u|\underline{w}_t, \theta)$. In particular, if w_{1t} and w_{2t} are conditionally independent and have the same size, the conditional Laplace transform and Log-Laplace transform of $w_{1,t+1} + w_{2,t+1}$ are respectively:

$$\varphi(u, 0|\underline{w}_t, \theta) \times \varphi(0, u|\underline{w}_t, \theta), \quad \text{and} \quad \psi(u, 0|\underline{w}_t, \theta) + \psi(0, u|\underline{w}_t, \theta).$$

Lemma 1.1 (Conditional zero probability for non-negative processes). *If w_t is univariate and nonnegative its (conditional) Laplace transform $\varphi_t(u) = \mathbb{E}_t[\exp(uw_{t+1})]$ is defined for $u \leq 0$ and*

$$\mathbb{P}_t(w_{t+1} = 0) = \lim_{u \rightarrow -\infty} \varphi_t(u).$$

Proof. We have $\varphi_t(u) = \mathbb{P}_t(w_{t+1} = 0) + \int_{w_{t+1} > 0} \exp(uw_{t+1}) d\mathbb{P}_t(w_{t+1})$. The Lebesgue theorem ensures that the last integral converges to zero when u goes to $-\infty$. \square

Lemma 1.2 (Conditional zero probability for non-negative multivariate processes). *Assume that:*

- $w_{1,t}$ is valued in \mathbb{R}^d ($d \geq 1$),
- $w_{2,t}$ is valued in $\mathbb{R}^+ = [0, +\infty)$,
- $\mathbb{E}_t[\exp(u'_1 w_{1,t+1} + u_2 w_{2,t+1})]$ exists for a given u_1 and $u_2 \leq 0$.

Then, we have:

$$\mathbb{E}_t[\exp(u'_1 w_{1,t+1}) \mathbf{1}_{\{w_{2,t+1}=0\}}] = \lim_{u_2 \rightarrow -\infty} \mathbb{E}_t[\exp(u'_1 w_{1,t+1} + u_2 w_{2,t+1})]. \quad (1.2)$$

Proof. We have that

$$\begin{aligned} & \lim_{u_2 \rightarrow -\infty} \mathbb{E}_t[\exp(u'_1 w_{1,t+1} + u_2 w_{2,t+1})] \\ &= \mathbb{E}_t[\exp(u'_1 w_{1,t+1}) \mathbf{1}_{\{w_{2,t+1}=0\}}] + \\ & \quad \lim_{u_2 \rightarrow -\infty} \mathbb{E}_t[\exp(u'_1 w_{1,t+1} + u_2 w_{2,t+1}) \mathbf{1}_{\{w_{2,t+1}>0\}}], \end{aligned}$$

and since in the second term on the right-hand side $\exp(u_2 w_{2,t+1}) \mathbf{1}_{\{w_{2,t+1}>0\}} \rightarrow 0$ when $u_2 \rightarrow -\infty$, Eq. (1.2) is a consequence of the Lebesgue theorem. \square

1.5 Affine (or Car) processes

In term structure applications, we will often consider *affine* processes (Definitions 1.1 and 1.2). These processes are indeed such that their multi-horizon Laplace transform are simple to compute (Lemma 1.5 and Proposition 1.5), which is key to compute bond prices.

1.5.1 Car processes of order one

Here is the definition of a *compound auto-regressive (Car)* process of order one:

Definition 1.1 (Affine process of order 1). A multivariate process w_{t+1} is Affine of order 1 [or $Car(1)$] if

$$\varphi_t(u) = \mathbb{E}_t[\exp(u'w_{t+1})] = \exp[a(u)'w_t + b(u)]$$

for some functions $a(\cdot)$ and $b(\cdot)$. These functions are univariate if w_{t+1} (and therefore u) is scalar.

Note that $a(\cdot)$ and $b(\cdot)$ may be deterministic functions of time (e.g., Chikhani and Renne (2023)).

A first key example is that of the Gaussian auto-regressive processes:

Example 1.6 (Univariate AR(1) Gaussian process). If $w_{t+1}|\underline{w}_t \sim \mathcal{N}(\nu + \rho w_t, \sigma^2)$, then:

$$\varphi_t(u) = \exp\left(u\rho w_t + u\nu + u^2\frac{\sigma^2}{2}\right) = \exp[a(u)'w_t + b(u)],$$

$$\text{with } \begin{cases} a(u) &= u\rho \\ b(u) &= u\nu + u^2\frac{\sigma^2}{2}. \end{cases}$$

Example 1.7 (Gaussian VAR). If $w_{t+1}|\underline{w}_t \sim \mathcal{N}(\mu + \Phi w_t, \Sigma)$, then:

$$\varphi_t(u) = \exp\left(u'(\mu + \Phi w_t) + \frac{1}{2}u'\Sigma u\right) = \exp[a(u)'w_t + b(u)],$$

$$\text{with } \begin{cases} a(u) &= \Phi'u \\ b(u) &= u'\mu + \frac{1}{2}u'\Sigma u = u'\mu + \frac{1}{2}(u \otimes u)'vec(\Sigma). \end{cases}$$

Example 1.8 (Quadratic Gaussian process). Consider vector $w_t = (x'_t, \text{vec}(x_t x'_t))'$, where x_t is a n -dimensional vector following a Gaussian $\text{VAR}(1)$, i.e.

$$x_{t+1} | w_t \sim \mathcal{N}(\mu + \Phi x_t, \Sigma).$$

Proposition 1.2 shows that if $u = (v, V)$ where $v \in \mathbb{R}^n$ and V a square symmetric matrix of size n , we have:

$$\begin{aligned} \varphi_t(u) &= \mathbb{E}_t \{ \exp [(v', \text{vec}(V)') \times w_{t+1}] \} \\ &= \exp \{ a_1(v, V)' x_t + \text{vec}(a_2(v, V))' \text{vec}(x'_t x_t) + b(v, V) \}, \end{aligned}$$

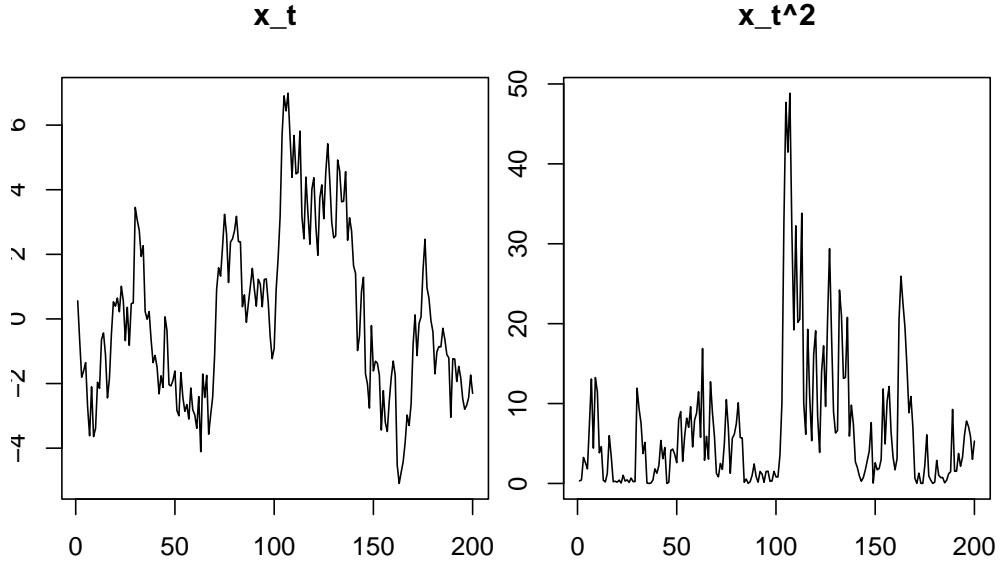
where:

$$\begin{aligned} a_2(u) &= \Phi' V (I_n - 2\Sigma V)^{-1} \Phi \\ a_1(u) &= \Phi' [(I_n - 2V\Sigma)^{-1} (v + 2V\mu)] \\ b(u) &= u' (I_n - 2\Sigma V)^{-1} \left(\mu + \frac{1}{2} \Sigma v \right) + \\ &\quad \mu' V (I_n - 2\Sigma V)^{-1} \mu - \frac{1}{2} \log |I_n - 2\Sigma V|. \end{aligned}$$

Quadratic processes can be used to construct positive process. Indeed, one can determine linear combinations of the components of w_t ($\alpha' w_t$, say) that are such that $\alpha' w_t \geq 0$. For instance, if x_t is scalar, $\alpha' w_t = x_t^2$ if $\alpha = (0, 1)'$. This is illustrated by Figure 1.1.

```
T <- 200
phi <- .9; sigma <- 1
x.t <- 0; x <- NULL
for(t in 1:T){
  x.t <- phi*x.t + sigma*rnorm(1)
  x <- c(x, x.t)}
par(mfrow=c(1,2),plt=c(.1,.95,.15,.85))
plot(x,type="l",xlab="",ylab="",main="x_t")
plot(x^2,type="l",xlab="",ylab="",main="x_t^2")
```

Another example of nonnegative process is that of the auto-regressive Gamma process (Gourieroux and Jasiak, 2006) and its extension (Monfort et al., 2017).

Figure 1.1: Simulation of a quadratic processes x_t .

Example 1.9 (Autoregressive gamma process, ARG(1)). An ARG process is defined as follows:

$$\frac{w_{t+1}}{\mu} \sim \gamma(\nu + z_t) \quad \text{where} \quad z_t \sim \mathcal{P}\left(\frac{\rho w_t}{\mu}\right),$$

with $\nu, \mu, \rho > 0$. (Alternatively $z_t \sim \mathcal{P}(\beta w_t)$, with $\rho = \beta\mu$.)

Proposition 1.3 shows that we have $\varphi_t(u) = \exp[a(u)'w_t + b(u)]$ with

$$\begin{cases} a(u) &= \frac{\rho u}{1 - u\mu} \\ b(u) &= -\nu \log(1 - u\mu). \end{cases}$$

One can simulate ARG processes by using this web-interface (select the “ARG” panel).

It can be shown that:

$$\begin{cases} \mathbb{E}(w_{t+1} | w_t) &= \nu\mu + \rho w_t \\ \text{Var}(w_{t+1} | w_t) &= \nu\mu^2 + 2\mu\rho w_t. \end{cases}$$

and that:

$$w_{t+1} = \nu\mu + \rho w_t + \varepsilon_{t+1},$$

where ε_{t+1} is a martingale difference $\Rightarrow w_{t+1}$ is a weak $AR(1)$.

Monfort et al. (2017) propose the extended ARG process and the ARG_0 process. The latter is such that $\nu = 0$ and βw_t is replaced with $\alpha + \beta w_t$, i.e.:

$$\frac{w_{t+1}}{\mu} \sim \gamma(z_t), \quad z_t \sim \mathcal{P}(\alpha + \beta w_t). \quad (1.3)$$

It is easily seen that we then have:

$$\varphi_t(u) = \exp \left[\frac{\beta \mu u}{1 - u \mu} w_t + \frac{\alpha \mu u}{1 - u \mu} \right].$$

The ARG_0 process features a point mass at zero, with conditional probability $\exp(-\alpha - \beta w_t)$. Note that 0 is absorbing if $\alpha = 0$.

Figure 1.2 displays the simulated path of an ARG_0 process (since we set ν to zero). Note that function `simul.ARG` is included in the `TSMODELS` package.

```
library(TSMODELS)
W <- simul.ARG(300,mu=.5,nu=0,rho=.9,alpha=.1)
plot(W,type="l")
```

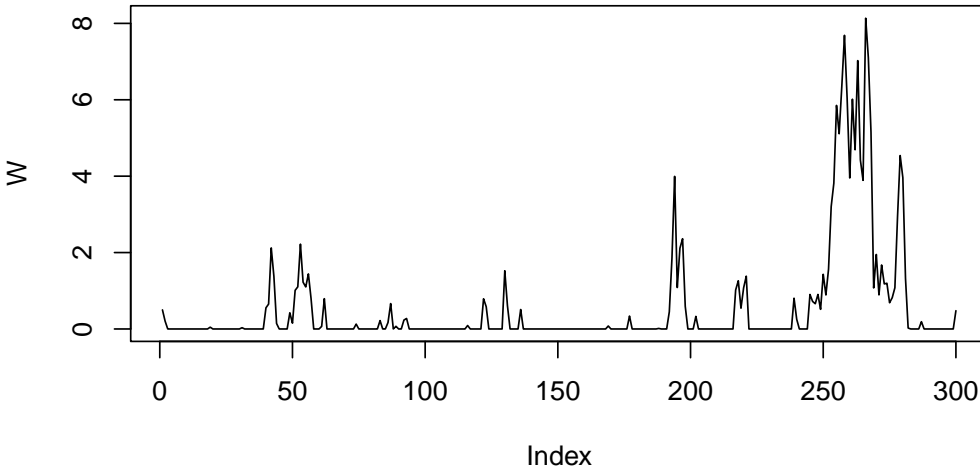


Figure 1.2: Simulation of an ARG0 processes.

Certain affine processes are valued in specific sets (e.g., integers). It is the case of compound Poisson processes:

Example 1.10 (Compound Poisson process). A compound Poisson process is defined as follows (with $\gamma > 0$, $0 < \pi < 1$, and $\lambda > 0$):

$$\frac{w_{t+1}}{\gamma} = z_{t+1} + \varepsilon_{t+1},$$

where z_{t+1} and ε_{t+1} conditionally independent, and where $z_{t+1} \sim \mathcal{B}\left(\frac{w_t}{\gamma}, \pi\right)$, with $\varepsilon_{t+1} \sim \mathcal{P}(\lambda)$.

This process is valued in $\{j\gamma, j \in \mathbb{N}, \gamma \in \mathbb{R}^+\}$ and we have:

$$\varphi_t(u) = \exp \left\{ \frac{w_t}{\gamma} \log[\pi \exp(u\gamma) + 1 - \pi] - \lambda[1 - \exp(u\gamma)] \right\},$$

i.e., $\varphi_t(u) = \exp(a(u)w_t + b(u))$ with

$$\begin{cases} a(u) &= \frac{1}{\gamma} \log[\pi \exp(u\gamma) + 1 - \pi], \\ b(u) &= -\lambda[1 - \exp(u\gamma)]. \end{cases}$$

We also have: $w_{t+1} = \pi w_t + \lambda\gamma + \eta_{t+1}$, where η_{t+1} is a martingale difference.

One can simulate such processes by using this web-interface (select the “Compound Poisson” panel). Figure 1.3 makes use of function `simul.compound.poisson` (in package `TSMODELS`) to simulate a compound Poisson process.

```
library(TSMODELS)
W <- simul.compound.poisson(100, Gamma=.5, Pi=0.5, lambda=.9)
plot(W, type="l")
```

1.5.2 Car processes of order p

Let us now define Car processes of order p :

Definition 1.2 (Affine process of order p). A multivariate process w_{t+1} is affine of order p [or $Car(p)$] if there exist functions $a_1(\cdot), \dots, a_p(\cdot)$, and $b(\cdot)$ such that:

$$\varphi_t(u) = \mathbb{E}_t[\exp(u'w_{t+1})] = \exp[a_1(u)'w_t + \dots + a_p(u)'w_{t+1-p} + b(u)].$$

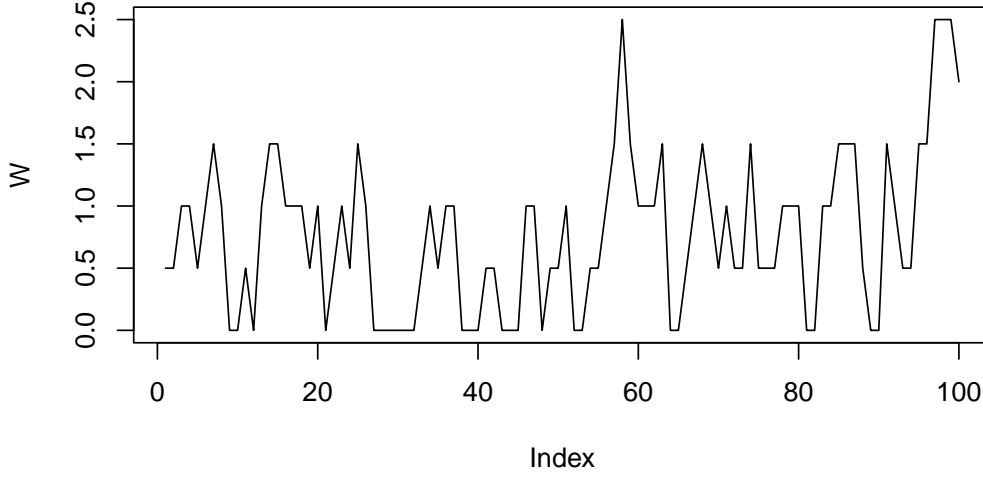


Figure 1.3: Simulation of a Compound Poisson process.

It can be seen that if w_t is $Car(p)$, then $W_t = [w'_t, w'_{t-1}, \dots, w'_{t-p+1}]'$ is $Car(1)$.¹ Therefore, without loss of generality we can assume $p = 1$.

The standard $Car(p)$ processes are auto-regressive processes of order p . These processes satisfy the definition of index affine processes:

Definition 1.3 (Univariate index affine process of order p). Let $\exp[a(u)w_t + b(u)]$ be the conditional Laplace transform of a univariate affine process of order 1, the process w_{t+1} is an *index-affine* process of order p if:

$$\varphi_t(u) = \mathbb{E}_t[\exp(uw_{t+1})] = \exp[a(u)(\beta_1 w_t + \dots + \beta_p w_{t+1-p}) + b(u)].$$

Examples 1.11 and 1.14 are two examples of index affine processes.

Example 1.11 (Gaussian AR(p) process). This example extends Example 1.6. Consider a Gaussian AR(p) process w_t ; that is:

$$w_{t+1} = \nu + \varphi_1 w_t + \dots + \varphi_p w_{t+1-p} + \sigma \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim i.i.d. \mathcal{N}(0, 1).$$

¹We have:

$$\begin{aligned} \mathbb{E}_t[\exp(u'W_{t+1})] &= \mathbb{E}_t[\exp(u'_1 w_{t+1} + u'_2 w_t + \dots + u'_p w_{t-p+2})] \\ &= \exp(u'_2 w_t + \dots + u'_p w_{t-p+2}) \mathbb{E}_t[\exp(u'_1 w_{t+1})] \\ &= \exp(u'_2 w_t + \dots + u'_p w_{t-p+2} + a'_1(u_1)w_t + \dots + a'_p(u_1)w_{t+1-p} + b(u_1)) \\ &= \exp[A(u)'W_t + B(u)], \end{aligned}$$

with $A(u)' = [u'_2 + a'_1(u_1), \dots, u'_p + a'_{p-1}(u_1), a'_p(u_1)]$ and $B(u) = b(u_1)$.

We have:

$$\varphi_t(u) = \exp \left[u\rho(\beta_1 w_t + \cdots + \beta_p w_{t+1-p}) + u\nu + u^2 \frac{\sigma^2}{2} \right]$$

with $\varphi_i = \rho\beta_i$.

Example 1.12 (ARG(p) process (positive)). This example extends Example 1.9. An ARG process of order p is defined as follows:

$$\frac{w_{t+1}}{\mu} \sim \gamma(\nu + z_t) \quad \text{where} \quad z_t \sim \mathcal{P}(\beta_1 w_t + \cdots + \beta_p w_{t+1-p}),$$

with $\nu, \mu, \beta_i > 0$. We have:

$$\varphi_t(u) = \exp \left[\frac{\rho u}{1 - u\mu} (\beta_1 w_t + \cdots + \beta_p w_{t+1-p}) - \nu \log(1 - u\mu) \right],$$

Process w_t admits the following AR(p) representation:

$$w_{t+1} = \nu\mu + \varphi_1 w_t + \cdots + \varphi_p w_{t+1-p} + \varepsilon_{t+1},$$

with $\varphi_i = \beta_i \mu$ and where ε_{t+1} is a martingale difference.

1.6 Markov chains

In this subsection, we show that the family of affine processes includes (some) regime-switching models. We consider a time-homogeneous Markov chain z_t , valued in the set of columns of Id_J , the identity matrix of dimension $J \times J$. The transition probabilities are denoted by $\pi(e_i, e_j)$, with $\pi(e_i, e_j) = \mathbb{P}(z_{t+1} = e_j | z_t = e_i)$. With these notations:

$$\mathbb{E}[\exp(v' z_{t+1}) | z_t = e_i, \underline{z_{t-1}}] = \sum_{j=1}^J \exp(v' e_j) \pi(e_i, e_j).$$

Hence, we have:

$$\varphi_t(v) = \exp[a_z(v)' z_t],$$

with

$$a_z(v) = \begin{bmatrix} \log \left(\sum_{j=1}^J \exp(v' e_j) \pi(e_1, e_j) \right) \\ \vdots \\ \log \left(\sum_{j=1}^J \exp(v' e_j) \pi(e_J, e_j) \right) \end{bmatrix}.$$

This proves that z_t is an affine process.

One can simulate a two-regime Markov chain by using this web-interface (select the “Markov-Switching” panel).

1.7 Wishart autoregressive (WAR) processes

WAR are *matrix processes*, valued in the space of $(L \times L)$ symmetric positive definite matrices.

Definition 1.4 (Wishart autoregressive (WAR) processes). Let W_{t+1} be a $WAR_L(K, M, \Omega)$ process. It is defined by:

$$\begin{aligned} & \mathbb{E}[\exp \text{Tr}(\Gamma W_{t+1}) | \underline{W}_t] \\ &= \exp \left\{ \text{Tr}[M' \Gamma (Id - 2\Omega \Gamma)^{-1} M W_t] - \frac{K}{2} \log[\det(Id - 2\Omega \Gamma)] \right\}, \end{aligned} \quad (1.4)$$

where Γ is a symmetric matrix,² K is a positive scalar, M is a $(L \times L)$ matrix, and Ω is a $(L \times L)$ symmetric positive definite matrix.

If K is an integer, Proposition 1.4 (in the appendix) shows that W_{t+1} can be obtained from:

$$\begin{cases} W_{t+1} &= \sum_{k=1}^K x_{k,t+1} x'_{k,t+1} \\ x_{k,t+1} &= M x_{k,t} + \varepsilon_{k,t+1}, \quad k \in \{1, \dots, K\}, \end{cases}$$

where $\varepsilon_{k,t+1} \sim i.i.d. \mathcal{N}(0, \Omega)$ (independent across k 's). The proposition also shows that we have:

$$\mathbb{E}(W_{t+1} | \underline{W}_t) = M W_t M' + K \Omega,$$

i.e. W_t follows a matrix weak AR(1) process.

²Indeed, $\text{Tr}(\Gamma W_{t+1})$ is equal to:

$$\sum_{i=1}^L (\Gamma W_{t+1})_{ii} = \sum_{i=1}^L \sum_{j=1}^L \gamma_{ij} W_{t+1,ij} = \sum_{i=1}^L \gamma_{ii} W_{t+1,ii} + \sum_{i < j}^L (\gamma_{ij} + \gamma_{ji}) W_{t+1,ij}.$$

In particular case, where $L = 1$ (univariate case), we have that:

$$\mathbb{E}[\exp(uW_{t+1})|\underline{W}_t] = \exp\left[\frac{um^2}{1-2\omega u}W_t - \frac{K}{2}\log(1-2\omega u)\right].$$

Hence, when $L = 1$, the Wishart process boils down to an $ARG(1)$ process (Example 1.9) with $\rho = m^2$, $\mu = 2\omega$, $\nu = \frac{K}{2}$.

1.8 Building affine processes

1.8.1 Univariate affine processes with stochastic parameters

Some univariate affine processes can be extended if they satisfy certain conditions. Specifically, consider a univariate affine process whose conditional L.T. is of the form:

$$\mathbb{E}_t \exp(uy_{t+1}) = \exp[a_0(u)y_t + b_0(u)\delta], \quad (1.5)$$

where $\delta = (\delta_1, \dots, \delta_m)' \in \mathcal{D}$. This process can be generalized by making δ stochastic (while staying in the affine family). More precisely assume that:

$$\mathbb{E}[\exp(uy_{t+1})|\underline{y}_t, \underline{z}_{t+1}] = \exp[a_0(u)y_t + b_0(u)'\Lambda z_{t+1}],$$

where Λ is a $(m \times k)$ matrix, with $\Lambda z_{t+1} \in \mathcal{D}$. In this case, if:

$$\mathbb{E}[\exp(v'z_{t+1})|\underline{y}_t, \underline{z}_t] = \exp[a_1(v)'z_t + b_1(v)],$$

then $w_{t+1} = (y_{t+1}, z'_{t+1})'$ is affine.³

³Indeed, we have:

$$\begin{aligned} & \mathbb{E}[\exp(uy_{t+1} + v'z_{t+1})|\underline{y}_t, \underline{z}_t] \\ &= \mathbb{E}\{\exp(v'z_{t+1})\mathbb{E}[\exp(uy_{t+1})|\underline{y}_t, \underline{z}_{t+1}]\mid \underline{y}_t, \underline{z}_t\} \\ &= \mathbb{E}\{\exp[a_0(u)y_t + b_0(u)'\Lambda z_{t+1} + v'z_{t+1}]\mid \underline{y}_t, \underline{z}_t\} \\ &= \exp\{a_0(u)y_t + a_1[\Lambda'b_0(u) + v]'z_t + b_1[\Lambda'b_0(u) + v]\}. \end{aligned}$$

Example 1.13 (Gaussian AR(p)). Using the notation of Example 1.11, it comes that an AR(p) processes satisfies Eq. (1.5) with $b_0(u) = \left(u, \frac{u^2}{2}\right)'$ and $\delta = (\nu, \sigma^2)' \in \mathcal{D} = \mathbb{R} \times \mathbb{R}^+$. In that case, δ (the vector of conditional mean and variance) can be replaced by ...

- $\begin{pmatrix} z_{1,t+1} \\ z_{2,t+1} \end{pmatrix}$, where $z_{1,t+1}$ and $z_{2,t+1}$ are independent AR(1) (see Example 1.6) and ARG(1) (see Example 1.9) processes, respectively.
- $\begin{pmatrix} \lambda'_1 & 0 \\ 0 & \lambda'_2 \end{pmatrix} \begin{pmatrix} z_{1,t+1} \\ z_{2,t+1} \end{pmatrix}$, where $z_{1,t+1}$ and $z_{2,t+1}$ are independent Markov chains.
- $\begin{pmatrix} \lambda'_1 \\ \lambda'_2 \end{pmatrix} z_{t+1}$, where z_{t+1} is a Markov chain.

Example 1.14 (ARG(p) model).

- $b_0(u) = -\nu \log(1 - u\mu)$, $\delta = \nu$.
- $\nu (\geq 0)$ can be specified for instance as a Markov chain or an ARG.

1.8.2 Multivariate affine processes

One can construct multivariate affine processes by employing the so-called recursive approach. Let us illustrate this by considering the bivariate case. (The multivariate generalization is straightforward.) Consider

$w_t = \begin{pmatrix} w_{1,t} \\ w_{2,t} \end{pmatrix}$, and assume that we have:

$$\begin{aligned} & \mathbb{E}[\exp(u_1 w_{1,t+1} | \underline{w_{1,t}}, \underline{w_{2,t}})] \\ &= \exp[a_{11}(u_1) w_{1,\textcolor{red}{t}} + a_{12}(u_1) w_{2,\textcolor{red}{t}} + b_1(u_1)], \end{aligned}$$

and:

$$\begin{aligned} & \mathbb{E}[\exp(u_2 w_{2,t+1} | \underline{w_{1,t+1}}, \underline{w_{2,t}})] \\ &= \exp[a_0(u_2) w_{1,\textcolor{red}{t+1}} + a_{21}(u_2) w_{1,\textcolor{red}{t}} + a_{22}(u_2) w_{2,\textcolor{red}{t}} + b_2(u_2)]. \end{aligned}$$

Then w_t is an affine process.⁴ The dynamics of the two components of w_t are of the form:

$$\begin{aligned} w_{1,t+1} &= \alpha_1 + \alpha_{11}w_{1,t} + \alpha_{12}w_{2,t} + \varepsilon_{1,t+1} \\ w_{2,t+1} &= \alpha_2 + \alpha_0w_{1,t+1} + \alpha_{21}w_{1,t} + \alpha_{22}w_{2,t} + \varepsilon_{2,t+1} \end{aligned}$$

Note that $\varepsilon_{1,t+1}$ and $\varepsilon_{2,t+1}$ are non-correlated martingale differences. In the general case, they are conditionally heteroskedastic. What precedes is at play in *VAR* model; Monfort et al. (2017) employ this approach to build vector auto-regressive gamma (VARG) processes.

1.8.3 Extending multivariate stochastic processes

Consider the same framework as in Section 1.8.1 when y_t is a n -dimensional vector. That is, replace Eq. (1.5) with:

$$\mathbb{E}_t \exp(u' y_{t+1}) = \exp[a_0(u)' y_t + b_0(u) \delta], \quad (1.6)$$

and further assume that δ is stochastic and depends on z_t , such that:

$$\mathbb{E}[\exp(uy_{t+1}) | \underline{y}_t, \underline{z}_{t+1}] = \exp[a_0(u)y_t + b_0(u)' \Lambda z_{t+1}],$$

where Λ is a $(m \times k)$ matrix, with $\Lambda z_{t+1} \in \mathcal{D}$. In this case, if:

$$\mathbb{E}[\exp(v' z_{t+1}) | \underline{y}_t, \underline{z}_t] = \exp[a_1(v)' z_t + b_1(v)],$$

then $w_{t+1} = (y_{t+1}, z'_{t+1})'$ is affine.

⁴Indeed, we have:

$$\begin{aligned} & \mathbb{E}[\exp(u_1 w_{1,t+1} + u_2 w_{2,t+1} | \underline{w}_{1,t}, \underline{w}_{2,t})] \\ &= \mathbb{E}\{\exp(u_1 w_{1,t+1}) \mathbb{E}[\exp(u_2 w_{2,t+1}) | \underline{w}_{1,t+1}, \underline{w}_{2,t}] | \underline{w}_{1,t}, \underline{w}_{2,t}\} \\ &= \mathbb{E}\{\exp[u_1 + a_0(u_2)w_{1,t+1} + a_{21}(u_2)w_{1,t} + a_{22}(u_2)w_{2,t} + b_2(u_2)] | \underline{w}_{1,t}, \underline{w}_{2,t}\} \\ &= \exp\{a_{11}[u_1 + a_0(u_2)]w_{1,t} + a_{12}[u_1 + a_0(u_2)]w_{2,t} + b_1[u_1 + a_0(u_2)] \\ & \quad + a_{21}(u_2)w_{1,t} + a_{22}(u_2)w_{2,t} + b_2(u_2)\}. \end{aligned}$$

Example 1.15 (Stochastic parameters Gaussian VAR(1)). This example extends Example 1.7. Using the same notations as in the latter example 1.7, we have

$$b_0(u) = \left(u', \frac{1}{2}(u \otimes u)' \right)' \quad \text{and} \quad \delta = (\mu', \text{vec}(\Sigma)')' \in \mathbb{R}^n \times \text{vec}(\mathcal{S}),$$

where \mathcal{S} is the set of symmetric positive semi-definite matrices. Vector δ can be replaced by:

$$\begin{pmatrix} z_{1,t+1} \\ z_{2,t+1} \end{pmatrix},$$

where

- $z_{1,t+1}$ is, for instance, a Gaussian VAR process.
- $z_{2,t+1}$ is
- obtained by applying the vec operator to a Wishart process,
- replaced by $\Lambda_2 z_{2,t+1}$, where Λ_2 is a $(n^2 \times J)$ matrix whose columns are $\text{vec}(\Sigma_j)$, $j \in \{1, \dots, J\}$, the Σ_j being $(n \times n)$ positive semi-definite,
- a standardized J -dimensional VARG process (multivariate extension of Example 1.9).

Example 1.16 (Regime-switching VAR(1)). One can also use this approach to construct (affine) regime-switching VAR processes (which is another extension of Example 1.7. For that, replace δ with

- $\begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \begin{pmatrix} z_{1,t+1} \\ z_{2,t+1} \end{pmatrix}$, where Λ_1 is a $(n \times J_1)$ matrix and $z_{1,t+1}$ is a Markov chain valued in the set of selection vectors of size J_1 (see Subsection 1.6), Λ_2 is the same matrix as in Example 1.15 and $z_{2,t+1}$ is a Markov chain valued in the set of selection vectors of size J_2 .
- or $\begin{pmatrix} \Lambda_1 \\ \Lambda_2 \end{pmatrix} z_{t+1}$, where Λ_1 and Λ_2 are the same matrices as above with $J_1 = J_2 = J$, and z_{t+1} is a Markov chain valued in the set of selection vectors of size J .

1.8.4 Extended affine processes

Some processes are not affine, but may be sub-components of an affine process. This can be useful to compute their conditional moments and multi-

horizon Laplace transform (as one can use the formulas presented above for that, using the enlarged—affine—vector).

Let us formally define an extended affine process:

Definition 1.5 (Extended Affine Processes). A process $w_{1,t}$ is extended affine if there exists a process $w_{2,t} = g(\underline{w_{1,t}})$ such that $(w'_{1,t}, w'_{2,t})'$ is affine (of order 1).

For an extended affine processes, $\varphi_{1,t}(u) = \mathbb{E}[\exp(u'w_{1,t+1})|\underline{w_{1,t}}]$ can be obtained from:

$$\begin{aligned}\varphi_t(u_1, u_2) &= \mathbb{E}[\exp(u'_1 w_{1,t+1} + u'_2 w_{2,t+1})|\underline{w_{1,t}}, \underline{w_{2,t}}] \\ &= \exp[a'_1(u_1, u_2)w_{1,t} + a'_2(u_1, u_2)w_{2,t} + b(u_1, u_2)]\end{aligned}$$

by:

$$\varphi_{1,t}(u) = \varphi_t(u, 0) = \exp[a'_1(u, 0)w_{1,t} + a'_2(u, 0)g(\underline{w_{1,t}}) + b(u, 0)].$$

In particular $w_{1,t}$ may be non-Markovian.

Similarly the multi-horizon Laplace transform (see Section 1.9)

$$\mathbb{E}[\exp(\gamma'_1 w_{1,t+1} + \dots + \gamma'_h w_{1,t+h})|\underline{w_{1,t}}]$$

can be obtained from the knowledge of the extended multi-horizon Laplace transform:

$$\begin{aligned}&\mathbb{E}_t[\exp(\{\gamma'_{1,1} w_{1,t+1} + \gamma'_{2,1} w_{2,t+1}\} + \dots + \{\gamma'_{1,h} w_{1,t+h} + \gamma'_{2,h} w_{2,t+h}\})] \\ &= \exp[A'_{1,t,h}(\gamma_1^h, \gamma_2^h)w_{1,t} + A'_{2,t,h}(\gamma_1^h, \gamma_2^h)w_{2,t} + B_{t,h}(\gamma_1^h, \gamma_2^h)],\end{aligned}$$

(with $\gamma_1^h = (\gamma'_{1,1}, \dots, \gamma'_{1,h})'$, and $\gamma_2^h = (\gamma'_{2,1}, \dots, \gamma'_{2,h})'$). We indeed have:

$$\begin{aligned}&\mathbb{E}[\exp(\gamma'_1 w_{1,t+1} + \dots + \gamma'_h w_{1,t+h})|\underline{w_{1,t}}] \\ &= \exp[A'_{1,t,h}(\gamma^h, 0)w_{1,t} + A'_{2,t,h}(\gamma^h, 0)g(\underline{w_{1,t}}) + B_{t,h}(\gamma^h, 0)],\end{aligned}$$

with $\gamma^h = (\gamma'_1, \dots, \gamma'_h)'$.

Example 1.17 (Affine process of order p). If $\{w_{1,t}\}$ is affine of order $p > 1$, then $(w_{1,t}, \dots, w_{1,t-p+1})$ is affine of order 1, but $\{w_{1,t}\}$ is not affine. That is, in that case, $w_{2,t} = (w'_{1,t-1}, \dots, w'_{1,t-p+1})'$.

This is a kind of extreme case since $w_{2,t}$ belongs to the information at $t-1$, which implies $a_2(u_1, u_2) = u_2$.

Example 1.18 (Gaussian ARMA process). Consider an $ARMA(1, 1)$ process

$$w_{1,t} - \varphi w_{1,t-1} = \varepsilon_t - \theta \varepsilon_{t-1},$$

with $|\varphi| < 1$, $|\theta| < 1$, and $\varepsilon_t \sim i.i.d. \mathcal{N}(0, \sigma^2)$.

$w_{1,t}$ is not Markovian. Now, take $w_{2,t} = \varepsilon_t = (1 - \theta L)^{-1}(1 - \varphi L)w_{1,t}$. We have:

$$\begin{pmatrix} w_{1,t+1} \\ w_{2,t+1} \end{pmatrix} = \begin{pmatrix} \varphi & -\theta \\ 0 & 0 \end{pmatrix} \begin{pmatrix} w_{1,t} \\ w_{2,t} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \varepsilon_{t+1}.$$

Hence $(w_{1,t}, w_{2,t})'$ is Gaussian $VAR(1)$, and, therefore, it is affine of order 1.

This is easily extended to $ARMA(p, q)$ and $VARMA(p, q)$ processes.

Example 1.19 (GARCH type process). Consider process $w_{1,t}$, defined by:

$$w_{1,t+1} = \mu + \varphi w_{1,t} + \sigma_{t+1} \varepsilon_{t+1},$$

where $|\varphi| < 1$ and $\varepsilon_t \sim i.i.d. \mathcal{N}(0, 1)$, and

$$\sigma_{t+1}^2 = \omega + \alpha \varepsilon_t^2 + \beta \sigma_t^2,$$

where $0 < \beta < 1$.

Consider $w_{2,t} = \sigma_{t+1}^2$ (which is a non-linear function of $\underline{w_{1,t}}$). Proposition 1.7 shows that:

$$\begin{aligned} & \mathbb{E} \left[\exp(u_1 w_{1,t+1} + u_2 w_{2,t+1}) | \underline{w_{1,t}} \right] \\ &= \exp \left[u_1 \mu + u_2 \omega - \frac{1}{2} \log(1 - 2u_2 \alpha) \right. \\ & \quad \left. + u_1 \varphi w_{1,t} + (u_2 \beta + \frac{u_1^2}{2(1 - 2u_2 \alpha)}) w_{2,t} \right], \end{aligned}$$

which is exponential affine in $(w_{1,t}, w_{2,t})$.

1.9 Multi-horizon Laplace transform

1.9.1 Recursive computation and direct pricing implications

In this subsection, we show that multi-horizon Laplace transforms of affine processes can be calculated recursively. Various examples will show how this can be exploited to price long-dated financial instruments.

Let us consider a multivariate process w_t , affine of order one. (As explained in Subsection 1.5.2, this includes the case of the order p case.) For the sake of generality, we consider the case where functions $a(\cdot)$, $b(\cdot)$ are possibly deterministic functions of time, denoted in this case $a_{t+1}(\cdot)$ and $b_{t+1}(\cdot)$:

$$\mathbb{E}_t \exp[(u' w_{t+1})] = \exp[a'_{t+1}(u) w_t + b_{t+1}(u)].$$

The multi-horizon Laplace transform associated with date t and horizon h is defined by:

$$\varphi_{t,h}(\gamma_1, \dots, \gamma_h) = \mathbb{E}_t[\exp(\gamma'_1 w_{t+1} + \dots + \gamma'_h w_{t+h})]. \quad (1.7)$$

Lemma 1.5 (in the appendix) shows that we have:

$$\varphi_{t,h}(\gamma_1, \dots, \gamma_h) = \exp(A'_{t,h} w_t + B_{t,h}),$$

where $A_{t,h} = A_{t,h}^h$ and $B_{t,h} = B_{t,h}^h$, the $A_{t,i}^h, B_{t,i}^h$ $i = 1, \dots, h$, being given recursively by:

$$\begin{cases} A_{t,i}^h &= a_{t+h+1-i}(\gamma_{h+1-i} + A_{t,i-1}^h), \\ B_{t,i}^h &= b_{t+h+1-i}(\gamma_{h+1-i} + A_{t,i-1}^h) + B_{t,i-1}^h, \\ A_{t,0}^h &= 0, B_{t,0}^h = 0. \end{cases}$$

If the functions a_t and b_t do not depend on t , these recursive formulas do not depend on t , and we get $\varphi_{t,h}(\gamma_1, \dots, \gamma_h)$, for any t , with only one recursion for each h .

Moreover, if the functions a_t and b_t do not depend on t , and if different sequences $(\gamma_1^h, \dots, \gamma_h^h)$, $h = 1, \dots, H$ (say) satisfy $\gamma_{h+1-i}^h = u_i$, for $i = 1, \dots, h$, and for any $h \leq H$, that is if we want to compute (*reverse-order* case):

$$\varphi_{t,h}(u_h, \dots, u_1) = \mathbb{E}_t[\exp(u'_h w_{\textcolor{red}{t}+1} + \dots + u'_1 w_{\textcolor{red}{t}+h})], \quad h = 1, \dots, H, \quad (1.8)$$

then Proposition 1.5 (in the appendix) shows that we can compute the $\varphi_{t,h}(u_h, \dots, u_1)$ for any t and any $h \leq H$ with only one recursion. That is $\varphi_{t,h}(u_h, \dots, u_1) = \exp(A'_h w_t + B_h)$ with:

$$\begin{cases} A_h &= a(u_h + A_{h-1}), \\ B_h &= b(u_h + A_{h-1}) + B_{h-1}, \\ A_0 &= 0, \quad B_0 = 0. \end{cases}$$

As mentioned above, what precedes has useful implications to price long-dated financial instruments such as nominal and real bonds (Examples 1.20 and 1.22, respectively), or futures (Example 1.23).

Example 1.20 (Nominal interest rates). Let $B(t, h)$ denote the date- t price of a nominal zero-coupon bond of maturity h . We Have:

$$B(t, h) = \mathbb{E}_t^{\mathbb{Q}} \exp(-r_t - \dots - r_{t+h-1}), \quad (1.9)$$

where r_t is the nominal short rate between t and $t+1$ (observed at t), and the associated (continuously-compounded) yield-to-maturity is given by:

$$R(t, h) = -\frac{1}{h} \log B(t, h), \quad h = 1, \dots, H. \quad (1.10)$$

If $r_t = \omega' w_t$ (say), then:

$$B(t, h) = \exp(-r_t) \mathbb{E}_t^{\mathbb{Q}} \exp(-\omega' w_{t+1} - \dots - \omega' w_{t+h-1}).$$

One can then price this bond by directly employing Eq. (1.8), with $u_1 = 0$ and $u_i = -\omega$, $i = 2, \dots, H$. The price $B(t, h)$ is exponential affine in w_t , the associated yield-to-maturity $R(t, h) = -1/h \log B(t, h)$ is affine in w_t .

Example 1.21 (No-arbitrage Nelson-Siegel model). In this example, we employ the results of Example 1.20 in the context described by Christensen et al. (2009). Specifically, we consider a three factor model following a Gaussian VAR (see Example 1.7):

$$w_t = \begin{bmatrix} X_{1,t} \\ X_{2,t} \\ X_{3,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 - \lambda & \lambda \\ 0 & 0 & 1 - \lambda \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \\ X_{3,t-1} \end{bmatrix} + \begin{bmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix},$$

where $\begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix} \sim i.i.d. \mathcal{N}(0, Id).$

The nominal short-term rate is given by $r_t = X_{1,t} + X_{2,t}$. In that case, we can use the results of Example 1.20 with $\omega = (-1, -1, 0)'$. The following lines of code do that:

```
library(TSModels)
lambda <- .05
Phi <- diag(c(1,1-lambda,1-lambda));Phi[2,3] <- lambda
Sigma <- .0005 * diag(3)
psi.parameterization=list(mu=matrix(0,3,1),Phi=Phi,Sigma=Sigma)
u1 <- matrix(0,3,1)
u2 <- matrix(c(-1,-1,0),ncol=1)
H <- 20
AB <- reverse.MHLT(psi.GaussianVAR,u1 = u1,u2 = u2,H = H,
                    psi.parameterization = psi.parameterization)
AB$A[1:2,,] <- AB$A[1:2,,] - 1 # add terms corresponding to exp(-r_t)
a.yield <- - AB$A / array((1:H) %x% rep(1,3),c(3,1,H))
b.yield <- - AB$B / array((1:H) %x% rep(1,3),c(1,1,H))
plot(a.yield[1,,],type="l",lwd=2,ylim=c(0,1),
     xlab="Maturity",ylab="Factor loadings")
lines(a.yield[2,,],col="red",lwd=2,lty=2)
lines(a.yield[3,,],col="blue",lwd=2,lty=3)
```

In the previous example, note the use of function `reverse.MHLT` (in package `TSModels`), that notably takes a L.T. as an argument (`psi`). In the previous example, we consider a Gaussian VAR, and we therefore assign `psi.GaussianVAR` to `psi`. We then need to provide function `reverse.MHLT` with the arguments of the `psi` function. These arguments are provided in the form of a list (input `psi.parameterization`).

Example 1.22 (Real interest rates). Denote by q_t the price index on date t

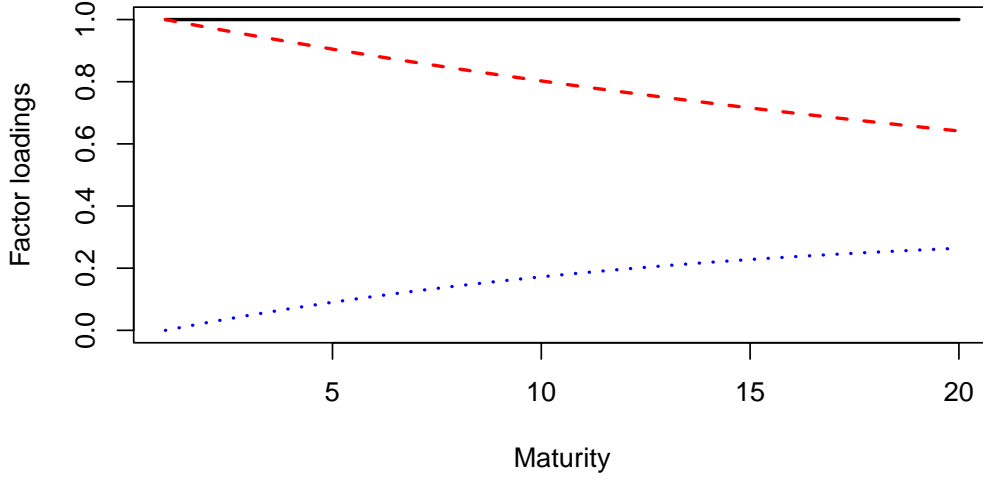


Figure 1.4: Factor loadings in the context of a no-arbitrage Nelson-Siegel model (Christensen, Diebold and Rudebusch, 2009). The first factor (black solid line) is a level factor. The second and third factors (red dashed line and blue dotted line, respectively) are slope factors.

and by $\pi_{t+1} = \log \frac{q_{t+1}}{q_t}$ the inflation rate on date $t + 1$. We have:

$$\bar{R}(t, h) = -\frac{1}{h} \log \bar{B}(t, h), \quad h = 1, \dots, H$$

$$\begin{aligned} \bar{B}(t, h) &= \mathbb{E}_t^{\mathbb{Q}} \exp(-r_t - \dots - r_{t+h-1} + \pi_{t+1} + \dots + \pi_{t+h}), \\ &= \exp(-r_t) \times \\ &\quad \mathbb{E}_t^{\mathbb{Q}} \exp(-r_{t+1} - \dots - r_{t+h-1} + \pi_{t+1} + \dots + \pi_{t+h}) \end{aligned}$$

If $r_t = \omega' w_t$ and $\pi_t = \bar{\omega}' w_t$, then $\bar{B}(t, h)$ is given by:

$$\exp(-r_t) \mathbb{E}_t^{\mathbb{Q}} \exp[(\bar{\omega} - \omega)' w_{t+1} + \dots + (\bar{\omega} - \omega)' w_{t+h-1} + \bar{\omega}' w_{t+h}]$$

One can then price this bond by directly employing Eq. (1.8), with $u_1 = \bar{\omega}$ and $u_i = \bar{\omega} - \omega$, $i = 2, \dots, H$.

Example 1.23 (Futures). Denote by $F(t, h)$ the date- t price of a future of maturity h (see Section XXX). That is $F(t, h) = \mathbb{E}_t^{\mathbb{Q}}(S_{t+h})$, $h = 1, \dots, H$, where S_t is the date- t price of the underlying asset.

- If $w_t = (\log S_t, x'_t)'$ then $F(t, h) = \mathbb{E}_t^{\mathbb{Q}} \exp(e'_1 w_{t+h})$. This can be calculated by using Eq. (1.8) with $u_1 = e_1$, and $u_i = 0$, for $i = 2, \dots, H$.
- If $w_t = (y_t, x'_t)'$ with $y_t = \log \frac{S_t}{S_{t-1}}$, then $F(t, h) = S_t \mathbb{E}_t^{\mathbb{Q}} \exp(e'_1 w_{t+1} + \dots + e'_1 w_{t+h})$. This can be calculated by using Eq. (1.8) with $u_i = e'_1$, $i = 1, \dots, H$.

1.9.2 Exponential payoff

Consider an asset providing the payoff $\exp(\nu' w_{t+h})$ on date $t + h$. Its price is given by:

$$P(t, h; \nu) = \mathbb{E}_t^{\mathbb{Q}} [\exp(-r_t - \dots - r_{t+h-1}) \exp(\nu' w_{t+h})].$$

If $r_t = \omega' w_t$, we have:

$$P(t, h; \nu) = \exp(-r_t) \mathbb{E}_t^{\mathbb{Q}} (\exp[-\omega' w_{t+1} - \dots - \omega' w_{t+h-1} + \nu' w_{t+h}]),$$

which can be calculated by Eq. (1.8), with $u_1 = \nu$ and $u_i = -\omega$ for $i = 2, \dots, H$.

What precedes can be extended to the case where the payoff (settled on date $t + h$) is of the form:

$$(\nu'_1 w_{t+h}) \exp(\nu'_2 w_{t+h}).$$

Indeed, we have

$$\left[\frac{\partial \exp[(s\nu_1 + \nu_2)' w_{t+h}]}{\partial s} \right]_{s=0} = (\nu'_1 w_{t+h}) \exp(\nu'_2 w_{t+h}).$$

Therefore:

$$\begin{aligned} & \mathbb{E}_t^{\mathbb{Q}} [\exp(-r_t - \dots - r_{t+h-1}) (\nu'_1 w_{t+h}) \exp(\nu'_2 w_{t+h})] \\ &= \left[\frac{\partial P(t, h; s\nu_1 + \nu_2)}{\partial s} \right]_{s=0}. \end{aligned} \tag{1.11}$$

This method is easily extended to price payoffs of the form $(\nu'_1 w_{t+h})^k \exp(\nu'_2 w_{t+h})$, with $k \in \mathbb{N}$.

1.10 VAR representation and conditional moments

An important property of affine processes is that their dynamics can be written as a vector-autoregressive process. This is useful to compute conditional moments of the process.

Proposition 1.1 (VAR representation of an affine process' dynamics). *If w_t is the affine process whose Laplace transform is defined in Def. 1.1, then its dynamics admits the following vectorial autoregressive representation:*

$$w_{t+1} = \mu + \Phi w_t + \Sigma^{\frac{1}{2}}(w_t) \varepsilon_{t+1}, \quad (1.12)$$

where ε_{t+1} is a difference of martingale sequence whose conditional covariance matrix is the identity matrix and where μ , Φ and $\Sigma(w_t) = \Sigma^{\frac{1}{2}}(w_t) \Sigma^{\frac{1}{2}}(w_t)'$ satisfy:

$$\mu = \left[\frac{\partial}{\partial u} b(u) \right]_{u=0}, \quad \Phi = \left[\frac{\partial}{\partial u} a(u)' \right]_{u=0} \quad (1.13)$$

$$\Sigma(w_t) = \left[\frac{\partial}{\partial u \partial u'} b(u) \right]_{u=0} + \left[\frac{\partial}{\partial u \partial u'} a(u)' w_t \right]_{u=0}. \quad (1.14)$$

Proof. When w_t is affine, its (conditional) cumulant generating function is of the form $\psi(u) = a(u)' w_t + b(u)$. The result directly follows from the formulas given in Section 1.3. \square

Proposition 1.6 (in the appendix) shows that the conditional means and variances of w_t are given by:

$$\mathbb{E}_t(w_{t+h}) = (I - \Phi)^{-1} (I - \Phi^h) \mu + \Phi^h w_t \quad (1.15)$$

$$\begin{aligned} \mathbb{V}ar_t(w_{t+h}) &= \Sigma(\mathbb{E}_t(w_{t+h-1})) + \Phi \Sigma(\mathbb{E}_t(w_{t+h-2})) \Phi' + \\ &\dots + \Phi^{h-1} \Sigma(w_t) \Phi^{h-1}'. \end{aligned} \quad (1.16)$$

Eq. (1.16) notably implies that $\mathbb{V}ar_t(w_{t+h})$ is an affine function of w_t . Indeed $\Sigma(\cdot)$ is an affine function, and the conditional expectations $\mathbb{E}_t(w_{t+h})$ are affine in w_t , as shown by Eq. (1.15).

The unconditional means and variances are given by:

$$\begin{cases} \mathbb{E}(w_t) &= (I - \Phi)^{-1} \mu \\ \text{vec}[\mathbb{V}ar(w_t)] &= (I_{n^2} - \Phi \otimes \Phi)^{-1} \text{vec}(\Sigma[(I - \Phi)^{-1} \mu]). \end{cases} \quad (1.17)$$

1.11 Truncated Laplace transforms of affine processes

In this section, we show how one can employ Fourier transforms to compute truncated conditional moments of affine processes. For that, let us introduce the following notation:

$$w_{t+1,T} = (w'_{t+1}, w'_{t+2}, \dots, w'_T)'$$

with w_t affine n -dimensional process.

We want to compute:

$$\tilde{\varphi}_t(u; v, \gamma) = \mathbb{E}_t[\exp(u' w_{t+1,T}) \mathbf{1}_{\{v' w_{t+1,T} < \gamma\}}].$$

Consider the complex untruncated conditional Laplace transform:

$$\varphi_t(z) = \mathbb{E}_t[\exp(z' w_{t+1,T})], \quad z \in \mathbb{C}^{nT},$$

computed using the same recursive algorithm as in the real case (see Section 1.9).

Duffie et al. (2000) have shown that we have (see also Proposition 1.8 in the appendix):

$$\tilde{\varphi}_t(u; v, \gamma) = \frac{\varphi_t(u)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\text{Im}[\varphi_t(u + ivx) \exp(-i\gamma x)]}{x} dx. \quad (1.18)$$

where Im means imaginary part.

Note that the integral in Eq. (1.18) is one dimensional (whatever the dimension of w_t). As shown in the following example, this can be exploited to price options.

Example 1.24 (Option pricing). Pricing calls and puts amounts to conditional expectations of the type (with $k > 0$):

$$\begin{aligned} & \mathbb{E}_t([\exp(u'_1 w_{t+1,T}) - k \exp(u'_2 w_{t+1,T})]^+) \\ &= \mathbb{E}_t([\exp(u'_1 w_{t+1,T}) - k \exp(u'_2 w_{t+1,T})] \mathbf{1}_{\{[\exp(u_1 - u_2)' w_{t+1,T}] > k\}}) \\ &= \tilde{\varphi}_t(u_1; u_2 - u_1, -\log k) - k \tilde{\varphi}_t(u_2; u_2 - u_1, -\log k). \end{aligned}$$

Example 1.25 (Exogenous short rate). Consider an asset whose date- t price is p_t . Denote its geometric asset return by y_t , i.e., $y_t = \log(p_t/p_{t-1})$. Consider an option written on this asset, with a strike equal kp_t .

If interest rates are deterministic, the option price, for a maturity h , is given by:

$$p_t \exp(-r_t - \dots - r_{t+h-1}) \mathbb{E}_t^{\mathbb{Q}} [\exp u'_1 w_{t+1,t+h} - k]^+$$

with $u_1 = e \otimes e_1$, where e is the h -dimensional vector with components equal to 1, and e_1 is the n -vector selecting the 1st component (y_t being the 1st component of w_t , say).

Example 1.26 (Endogenous short rate). Consider the same context as in Example 1.25, but with a stochastic (endogenous) short-term rate. For instance, assume that $r_{t+1} = \omega_0 + \omega'_1 w_t$. The option price then is:

$$\begin{aligned} & p_t \mathbb{E}_t^{\mathbb{Q}} [\exp(-\omega_0 - \omega'_1 w_t - \dots - \omega_0 - \omega'_1 w_{t+h-1}) [\exp(u'_1 w_{t+1,t+h}) - k]^+] \\ &= p_t \exp(-h\omega_0 - \omega'_1 w_t) \mathbb{E}_t^{\mathbb{Q}} \left([\exp(\tilde{u}'_1 w_{t+1,t+h}) - k \exp(u_2 w_{t+1,t+h})]^+ \right), \end{aligned}$$

with $\tilde{u}'_1 = u_1 + u_2$, [$u_1 = e \otimes e_1$ as before], and $u_2 = (-\omega'_1, \dots, -\omega'_1, 0)'$.

Example 1.27 (Numerical example: Conditional cumulated distribution function (c.d.f.)). Let us use the model used in Example 1.21. Suppose we want to compute the conditional distribution of the average interest rate over the next H periods, i.e., $\frac{1}{H}(r_{t+1} + \dots + r_{t+H})$. Hence, we want to compute $\mathbb{E}_t[\mathbf{1}_{\{v'w_{t+1,T} < \gamma\}}]$ with $v'w_{t+1,T} = \frac{1}{H}(r_{t+1} + \dots + r_{t+H})$.

```
H <- 10
X <- matrix(c(0.01,.02,0),3,1)
x <- exp(seq(-10,10,length.out=1000))
u1 <- matrix(c(1/H,1/H,0),3,1) %*% matrix(1i*x,nrow=1); u2 <- u1
AB <- reverse.MHLT(psi.GaussianVAR,u1 = u1,u2 = u2,H = H,
                    psi.parameterization = psi.parameterization)
s1 <- matrix(exp(t(X) %*% AB$A[, ,H] + AB$B[, ,H]),ncol=1)
dx <- matrix(x-c(0,x[1:length(x)-1]),length(x),1)
gamma <- seq(-.2,.3,length.out=1000)
fx <- outer(x,gamma,function(r,c){Im(s1[,1]*exp(-1i*r*c))/r})*dx[,1]
f <- 1/2 - 1/pi * apply(fx,2,sum)
plot(gamma,f,type="l",xlab="",lwd=2)
```

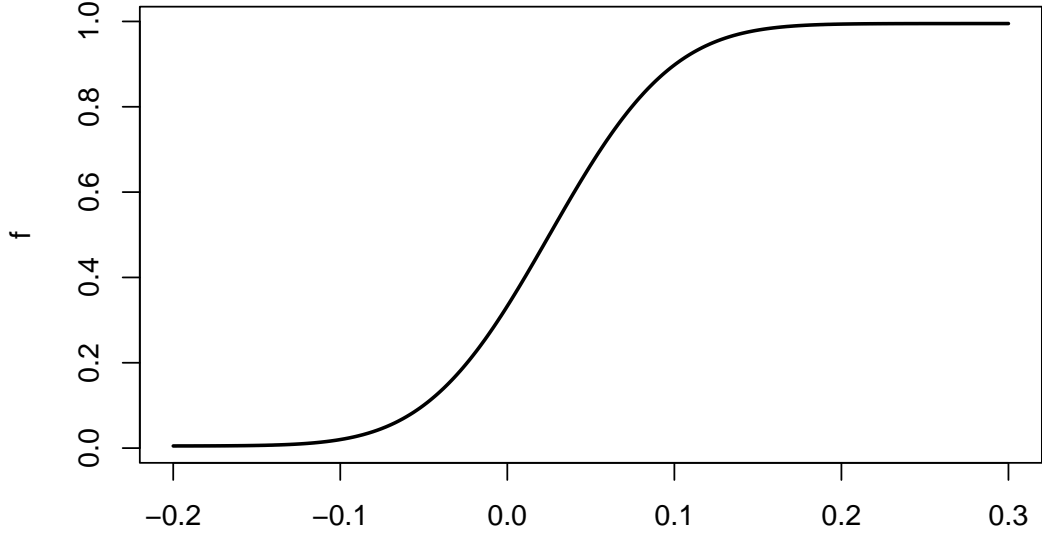


Figure 1.5: Conditional cumulated distribution function (c.d.f.) of $\frac{1}{H}(r_{t+1} + \dots + r_{t+H})$.

1.12 Appendices

Lemma 1.3. *If $\mu \in \mathbb{R}^L$ and Q is a $(L \times L)$ matrix symmetric positive definite, then:*

$$\int_{\mathbb{R}^L} \exp(-u'Qu + \mu'u) du = \frac{\pi^{L/2}}{(\det Q)^{1/2}} \exp\left(\frac{1}{4}\mu'Q^{-1}\mu\right).$$

Proof. The integral is:

$$\begin{aligned} & \int_{\mathbb{R}^L} \exp\left[-(u - \frac{1}{2}Q^{-1}\mu)'Q(u - \frac{1}{2}Q^{-1}\mu)\right] \exp\left(\frac{1}{4}\mu'Q^{-1}\mu\right) du \\ &= \frac{\pi^{L/2}}{(\det Q)^{1/2}} \exp\left(\frac{1}{4}\mu'Q^{-1}\mu\right) \end{aligned}$$

[using the formula for the unit mass of $\mathcal{N}(0.5Q^{-1}\mu, (2Q)^{-1})$]. \square

Lemma 1.4. *If $\varepsilon_{t+1} \sim \mathcal{N}(0, Id)$, we have*

$$\mathbb{E}_t(\exp[\lambda'\varepsilon_{t+1} + \varepsilon'_{t+1}V\varepsilon_{t+1}]) = \frac{1}{[\det(I - 2V)]^{1/2}} \exp\left[\frac{1}{2}\lambda'(I - 2V)^{-1}\lambda\right].$$

Proof. We have

$$\mathbb{E}_t \exp(\lambda' \varepsilon_{t+1} + \varepsilon'_{t+1} V \varepsilon_{t+1}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \exp \left[-u' \left(\frac{1}{2} I - V \right) u + \lambda' u \right] du$$

From Lemma 1.3, if $u \in \mathbb{R}^n$, then

$$\int_{\mathbb{R}^n} \exp(-u' Q u + \mu' u) du = \frac{\pi^{n/2}}{(\det Q)^{1/2}} \exp \left(\frac{1}{4} \mu' Q^{-1} \mu \right).$$

Therefore:

$$\begin{aligned} & \mathbb{E}_t \exp(\lambda' \varepsilon_{t+1} + \varepsilon'_{t+1} V \varepsilon_{t+1}) \\ &= \frac{1}{2^{n/2} \left[\det \left(\frac{1}{2} I - V \right) \right]^{1/2}} \exp \left[\frac{1}{4} \lambda' \left(\frac{1}{2} I - V \right)^{-1} \lambda \right]. \end{aligned}$$

□

Proposition 1.2 (Quadratic Gaussian process). *Consider vector $w_t = (x'_t, \text{vec}(x_t x'_t))'$, where x_t is a n -dimensional vector following a Gaussian VAR(1), i.e.*

$$x_{t+1} | \underline{w}_t \sim \mathcal{N}(\mu + \Phi x_t, \Sigma).$$

If $u = (v, V)$ where $v \in \mathbb{R}^n$ and V a square symmetric matrix of size n , we have:

$$\begin{aligned} \varphi_t(u) &= \mathbb{E}_t \{ \exp [(v', \text{vec}(V)') \times w_{t+1}] \} \\ &= \exp \{ a_1(v, V)' x_t + \text{vec}(a_2(v, V))' \text{vec}(x'_t x_t) + b(v, V) \}, \end{aligned}$$

where:

$$\begin{aligned} a_2(u) &= \Phi' V (I_n - 2\Sigma V)^{-1} \Phi \\ a_1(u) &= \Phi' [(I_n - 2V\Sigma)^{-1} (v + 2V\mu)] \\ b(u) &= u' (I_n - 2\Sigma V)^{-1} \left(\mu + \frac{1}{2} \Sigma v \right) + \\ &\quad \mu' V (I_n - 2\Sigma V)^{-1} \mu - \frac{1}{2} \log |I_n - 2\Sigma V|. \end{aligned}$$

Proof. We have:

$$\begin{aligned}
& \mathbb{E}_t(\exp(v'x_{t+1} + \text{vec}(V)' \text{vec}(x_{t+1}x'_{t+1}))) \\
&= \mathbb{E}_t[\exp(v'(\mu + \Phi x_t + \Sigma^{1/2}\varepsilon_{t+1}) + \\
&\quad \text{vec}(V)' \text{vec}((\mu + \Phi x_t + \Sigma^{1/2}\varepsilon_{t+1})(\mu + \Phi x_t + \Sigma^{1/2}\varepsilon_{t+1})'))] \\
&= \exp[v'(\mu + \Phi x_t) + \text{vec}(V)' \text{vec}\{(\mu + \Phi x_t)(\mu + \Phi x_t)'\}] \times \\
&\quad \mathbb{E}_t[\exp(v'\Sigma^{1/2}\varepsilon_{t+1} + \underbrace{2 \text{vec}(V)' \text{vec}\{(\mu + \Phi x_t)(\varepsilon'_{t+1}\Sigma^{1/2'})\}}_{=(\mu + \Phi x_t)'V\Sigma^{1/2}\varepsilon_{t+1}}) + \\
&\quad \underbrace{\text{vec}(V)' \text{vec}\{(\Sigma^{1/2}\varepsilon_{t+1})(\Sigma^{1/2}\varepsilon_{t+1})'\}}_{=\varepsilon'_{t+1}\Sigma^{1/2'}V\Sigma^{1/2}\varepsilon_{t+1}})]
\end{aligned}$$

Lemma 1.4 can be used to compute the previous conditional expectation, with $\lambda = \Sigma^{1/2'}(v + 2V'(\mu + \Phi x_t))$. Some algebra then leads to the result. \square

Proposition 1.3. *Consider the following auto-regressive gamma process:*

$$\frac{w_{t+1}}{\mu} \sim \gamma(\nu + z_t) \quad \text{where} \quad z_t \sim \mathcal{P}\left(\frac{\rho w_t}{\mu}\right),$$

with $\nu, \mu, \rho > 0$. (Alternatively $z_t \sim \mathcal{P}(\beta w_t)$, with $\rho = \beta\mu$.)

We have: $\varphi_t(u) = \exp\left[\frac{\rho u}{1 - u\mu}w_t - \nu \log(1 - u\mu)\right]$, for $u < \frac{1}{\mu}$.

Proof. Given \underline{w}_t , we have $z_t \sim \mathcal{P}\left(\frac{\rho w_t}{\mu}\right)$. We have:

$$\begin{aligned}
\mathbb{E}[\exp(uw_{t+1})|\underline{w}_t] &= \mathbb{E}\left\{\mathbb{E}\left[\exp\left(u\mu\frac{w_{t+1}}{\mu}\right)|\underline{w}_t, z_t\right]|\underline{w}_t\right\} \\
&= \mathbb{E}[(1 - u\mu)^{-(\nu + z_t)}|\underline{w}_t] \\
&= (1 - u\mu)^{-\nu} \mathbb{E}\{\exp[-z_t \log(1 - u\mu)]|\underline{w}_t\} \\
&= (1 - u\mu)^{-\nu} \exp\left\{\frac{\rho w_t}{\mu}[\exp(-\log(1 - u\mu))] - \frac{\rho w_t}{\mu}\right\} \\
&= \exp\left[\frac{\rho u w_t}{1 - u\mu} - \nu \log(1 - u\mu)\right],
\end{aligned}$$

using the fact that the L.T. of $\gamma(\nu)$ is $(1 - u)^{-\nu}$ and that the L.T. of $\mathcal{P}(\lambda)$ is $\exp[\lambda(\exp(u) - 1)]$. \square

Proposition 1.4 (Dynamics of a WAR process). *If K is an integer, W_{t+1} can be obtained from:*

$$\begin{cases} W_{t+1} &= \sum_{k=1}^K x_{k,t+1} x'_{k,t+1} \\ x_{k,t+1} &= Mx_{k,t} + \varepsilon_{k,t+1}, \quad k \in \{1, \dots, K\}, \end{cases}$$

where $\varepsilon_{k,t+1} \sim i.i.d.\mathcal{N}(0, \Omega)$ (independent across k 's). In particular, we have:

$$\mathbb{E}(W_{t+1} | \underline{W}_t) = MW_t M' + K\Omega,$$

i.e. W_t follows a matrix weak AR(1) process.

Proof. For $K = 1$, $W_{t+1} = x_{t+1} x'_{t+1}$, $x_{t+1} = Mx_t + \Omega^{1/2}u_{t+1}$ and $u_{t+1} \sim i.i.d.\mathcal{N}(0, Id_L)$. We have:

$$\mathbb{E}[\exp(\text{Tr} \Gamma W_{t+1}) | \underline{w}_t] = \mathbb{E}\{\mathbb{E}[\exp(\text{Tr} \Gamma x_{t+1} x'_{t+1}) | \underline{x}_t] | \underline{w}_t\}$$

and:

$$\begin{aligned} & \mathbb{E}[\exp(\text{Tr} \Gamma x_{t+1} x'_{t+1}) | \underline{x}_t] = \mathbb{E}[\exp(x'_{t+1} \Gamma x_{t+1} | \underline{x}_t)] \\ &= \mathbb{E}[\exp((Mx_t + \Omega^{1/2}u_{t+1})' \Gamma (Mx_t + \Omega^{1/2}u_{t+1}) / x_t)] \\ &= \exp(x'_t M' \Gamma M x_t) \mathbb{E}[\exp(2x'_t M' \Gamma \Omega^{1/2}u_{t+1} + u'_{t+1} \Omega^{1/2} \Gamma \Omega^{1/2} u_{t+1}) / x_t] \\ &= \frac{\exp(x'_t M' \Gamma M x_t)}{(2\pi)^{L/2}} \times \\ & \quad \int_{\mathbb{R}^L} \exp \left[2x'_t M' \Gamma \Omega^{1/2} u_{t+1} - u'_{t+1} \left(\frac{1}{2} Id_L - \Omega^{1/2} \Gamma \Omega^{1/2} \right) u_{t+1} \right] du_{t+1}. \end{aligned}$$

Using Lemma 1.3 with $\mu' = 2x'_t M' \Gamma \Omega^{1/2}$, $Q = \frac{1}{2} Id_L - \Omega^{1/2} \Gamma \Omega^{1/2}$ and after some algebra, the RHS becomes:

$$\frac{\exp[x'_t M' \Gamma (Id_L - 2\Omega \Gamma)^{-1} M x_t]}{\det[Id_L - 2\Omega^{1/2} \Gamma \Omega^{1/2}]} = \frac{\exp \text{Tr}[M' \Gamma (Id_L - 2\Omega^{-1}) M W_t]}{\det[Id_L - 2\Omega \Gamma]^{1/2}},$$

which depends on x_t through W_t , and gives the result for $K = 1$; the result for any K integer follows. \square

Lemma 1.5. *We have:*

$$\varphi_{t,h}(\gamma_1, \dots, \gamma_h) = \exp(A'_{t,h} w_t + B_{t,h}),$$

where $A_{t,h} = A_{t,h}^h$ and $B_{t,h} = B_{t,h}^h$, the $A_{t,i}^h, B_{t,i}^h$ $i = 1, \dots, h$, being given recursively by:

$$(i) \begin{cases} A_{t,i}^h &= a_{t+h+1-i}(\gamma_{h+1-i} + A_{t,i-1}^h), \\ B_{t,i}^h &= b_{t+h+1-i}(\gamma_{h+1-i} + A_{t,i-1}^h) + B_{t,i-1}^h, \\ A_{t,0}^h &= 0, B_{t,0}^h = 0. \end{cases}$$

Proof. For any $j = 1, \dots, h$ we have:

$$\varphi_{t,h}(\gamma_1, \dots, \gamma_h) = \mathbb{E}_t[\exp(\gamma'_1 w_{t+1} + \dots \gamma'_j w_{t+j} + A_{t,h-j}^{h'} w_{t+j} + B_{t,h-j}^h)]$$

where:

$$(ii) \begin{cases} A_{t,h-j+1}^h &= a_{t+j}(\gamma_j + A_{t,h-j}^h), \\ B_{t,h-j+1}^h &= b_{t+j}(\gamma_j + A_{t,h-j}^h) + B_{t,h-j}^h, \\ A_{t,0}^h &= 0, B_{t,0}^h = 0. \end{cases}$$

Since this is true for $j = h$, and if this is true for j , we get:

$$\begin{aligned} \varphi_{t,h}(\gamma_1, \dots, \gamma_h) &= \mathbb{E}_t[\exp(\gamma'_1 w_{t+1} + \dots + \gamma'_{j-1} w_{t+j-1} + a'_{t+j}(\gamma_j + A_{t,h-j}^h) w_{t+j-1} \\ &\quad + b_{t+j}(\gamma_j + A_{t,h-j}^h) + B_{t,h-j}^h)], \end{aligned}$$

and, therefore, this is true for $j - 1$, with $A_{t,h-j+1}^h$ and $B_{t,h-j+1}^h$ given by formulas (ii) above.

For $j = 1$ we get:

$$\begin{aligned} \varphi_{t,h}(\gamma_1, \dots, \gamma_h) &= \mathbb{E}_t \exp(\gamma'_1 w_{t+1} + A_{t,h-1}^{h'} w_{t+1} + B_{t,h-1}^h) \\ &= \exp(A'_{t,h} w_t + B_{t,h}), \end{aligned}$$

Finally note that if we put $h - j + 1 = i$, formulas (ii) become (i). \square

Proposition 1.5 (Reverse-order multi-horizon Laplace transform). *If the functions a_t and b_t do not depend on t , and if different sequences $(\gamma_1^h, \dots, \gamma_h^h), h = 1, \dots, H$ (say) satisfy $\gamma_{h+1-i}^h = u_i$, for $i = 1, \dots, h$, and for any $h \leq H$, that is if we want to compute (“reverse order” case):*

$$\varphi_{t,h}(u_h, \dots, u_1) = \mathbb{E}_t[\exp(u'_h w_{t+1} + \dots + u'_1 w_{t+h})], \quad h = 1, \dots, H,$$

then we can compute the $\varphi_{t,h}(u_h, \dots, u_1)$ for any t and any $h \leq H$, with only one recursion, i.e. $\varphi_{t,h}(u_h, \dots, u_1) = \exp(A'_h w_t + B_h)$ with:

$$\begin{cases} A_h &= a(u_h + A_{h-1}), \\ B_h &= b(u_h + A_{h-1}) + B_{h-1}, \\ A_0 &= 0, \quad B_0 = 0. \end{cases} \quad (1.19)$$

Proof. According to Lemma 1.5, we have, in this case:

$$\begin{cases} A_i^h &= a(u_i + A_{i-1}^h), \\ B_i^h &= b(u_i + A_{i-1}^h) + B_{i-1}^h, \\ A_0^h &= 0, \quad B_0^h = 0. \end{cases}$$

The previous sequences do not depend on h and are given by Eq. (1.19). \square

Proposition 1.6 (Conditional means and variances of an affine process). *Consider an affine process w_t . Using the notation of Proposition 1.1, we have:*

$$\mathbb{E}_t(w_{t+h}) = (I - \Phi)^{-1}(I - \Phi^h)\mu + \Phi^h w_t \quad (1.20)$$

$$\begin{aligned} \text{Var}_t(w_{t+h}) &= \Sigma(\mathbb{E}_t(w_{t+h-1})) + \Phi \Sigma(\mathbb{E}_t(w_{t+h-2}))\Phi' + \\ &\quad \dots + \Phi^{h-1} \Sigma(w_t) \Phi^{h-1}'. \end{aligned} \quad (1.21)$$

Eq. (1.21) notably shows that $\text{Var}_t(w_{t+h})$ is an affine function of w_t . Indeed $\Sigma(\cdot)$ is an affine function, and the conditional expectations $\mathbb{E}_t(w_{t+h})$ are affine in w_t , as shown by Eq. (1.20).

The unconditional mean and variance of w_t are given by:

$$\begin{cases} \mathbb{E}(w_t) &= (I - \Phi)^{-1}\mu \\ \text{vec}[\text{Var}(w_t)] &= (I_{n^2} - \Phi \otimes \Phi)^{-1} \text{vec}(\Sigma[(I - \Phi)^{-1}\mu]). \end{cases} \quad (1.22)$$

Proof. Eq. (1.20) is easily deduced from Eq. (1.12), using that $\mathbb{E}_t(\varepsilon_{t+k}) = 0$ for $k > 0$.

As regards Eq. (1.21):

$$\text{Var}_t(w_{t+h}) = \text{Var}_t\left(\Sigma(w_{t+h-1})^{\frac{1}{2}}\varepsilon_{t+h} + \dots + \Phi^{h-1}\Sigma(w_t)^{\frac{1}{2}}\varepsilon_{t+1}\right).$$

The conditional expectation at t of all the terms of the sum is equal to zero since, for $i \geq 1$:

$$\begin{aligned} \mathbb{E}_t\left[\Sigma(w_{t+i-1})^{\frac{1}{2}}\varepsilon_{t+i}\right] &= \mathbb{E}_t[\underbrace{\mathbb{E}_{t+i-1}\{\Sigma(w_{t+i-1})^{\frac{1}{2}}\varepsilon_{t+i}\}}_{=\Sigma(w_{t+i-1})^{\frac{1}{2}}\mathbb{E}_{t+i-1}\{\varepsilon_{t+i}\}=0}], \end{aligned}$$

and $\forall i < j$,

$$\text{Cov}_t\left[\Sigma(w_{t+i-1})^{\frac{1}{2}}\varepsilon_{t+i}, \Sigma(w_{t+j-1})^{\frac{1}{2}}\varepsilon_{t+j}\right] = \mathbb{E}_t\left[\Sigma(w_{t+i-1})^{\frac{1}{2}}\varepsilon_{t+i}\varepsilon'_{t+j}\Sigma'(w_{t+j-1})^{\frac{1}{2}}\right],$$

which can be seen to be equal to zero by conditioning on the information available on date $t + j - 1$.

Using the same conditioning, we obtain that:

$$\begin{aligned}
\mathbb{V}ar_t \left[\Phi^{h-j} \Sigma(w_{t+j-1})^{\frac{1}{2}} \varepsilon_{t+j} \right] &= \mathbb{E}_t \left[\Phi^{h-j} \Sigma(w_{t+j-1})^{\frac{1}{2}} \varepsilon_{t+j} \varepsilon'_{t+j} \Sigma'(w_{t+j-1})^{\frac{1}{2}} \Phi^{h-j'} \right] \\
&= \mathbb{E}_t \left[\Phi^{h-j} \Sigma(w_{t+j-1})^{\frac{1}{2}} \mathbb{E}_{t+j-1}(\varepsilon_{t+j} \varepsilon'_{t+j}) \Sigma'(w_{t+j-1})^{\frac{1}{2}} \Phi^{h-j'} \right] \\
&= \Phi^{h-j} \mathbb{E}_t[\Sigma(w_{t+j-1})] \Phi^{h-j'} \\
&= \Phi^{h-j} \Sigma(\mathbb{E}_t[w_{t+j-1}]) \Phi^{h-j'},
\end{aligned}$$

where the last equality results from the fact that the fact that $\Sigma(\cdot)$ is affine (see Eq. (1.14)). \square

Proposition 1.7 (Affine property of the GARCH-type process). *The process $w_t = (w_{1,t}, w_{2,t})$ defined by:*

$$\begin{cases} w_{1,t+1} &= \mu + \varphi w_{1,t} + \sigma_{t+1} \varepsilon_{t+1} \mid \varphi < 1 \\ \sigma_{t+1}^2 &= \omega + \alpha \varepsilon_t^2 + \beta \sigma_t^2 \mid 0 < \beta < 1, \alpha > 0, \omega > 0 \\ w_{2,t} &= \sigma_{t+1}^2, \quad \varepsilon_t \sim i.i.d. \mathcal{N}(0, 1) \end{cases}$$

is affine.

Proof. Note that $w_{2,t}$ is function of $\underline{w_{1,t}}$

$$\begin{aligned}
&\mathbb{E}[\exp(u_1 w_{1,t+1} + u_2 w_{2,t+1}) \mid \underline{w_{1,t}}] \\
&= \exp(u_1 \mu + u_1 \varphi w_{1,t} + u_2 \omega + u_2 \beta w_{2,t}) \mathbb{E}[\exp(u_1 \sigma_{t+1} \varepsilon_{t+1} + u_2 \alpha \varepsilon_{t+1}^2) \mid \underline{w_{1,t}}]
\end{aligned}$$

and, using Lemma 1.4:

$$\begin{aligned}
&\mathbb{E}[\exp(u_1 w_{1,t+1} + u_2 w_{2,t+1}) \mid \underline{w_{1,t}}] \\
&= \exp(u_1 \mu + u_1 \varphi w_{1,t} + u_2 \omega + u_2 \beta w_{2,t}) \exp \left[-\frac{1}{2} \log(1 - 2u_2 \alpha) + \frac{u_1^2 w_{2,t}}{2(1 - 2u_2 \alpha)} \right] \\
&= \exp \left[u_1 \mu + u_2 \omega - \frac{1}{2} \log(1 - 2u_2 \alpha) + u_1 \varphi w_{1,t} + \left(u_2 \beta + \frac{u_1^2}{2(1 - 2u_2 \alpha)} \right) w_{2,t} \right],
\end{aligned}$$

which is exponential affine in $(w_{1,t}, w_{2,t})$. \square

Proposition 1.8 (Computation of truncated conditional moments). *If $\varphi(z) = \mathbb{E}[\exp(z'w)]$, we have:*

$$\mathbb{E}[\exp(u'w) \mathbf{1}_{(v'w < \gamma)}] = \frac{\varphi(u)}{2} - \frac{1}{\pi} \int_0^\infty \frac{\mathcal{I}m[\varphi(u + ivx) \exp(-i\gamma x)]}{x} dx. \quad (1.23)$$

Proof. We want to compute $\tilde{\varphi}_t(u; v, \gamma) = \mathbb{E}_t[\exp(u'w) \mathbf{1}_{(v'w < \gamma)}]$. Let us first note that, for given u and v , $\tilde{\varphi}_t(u; v, \gamma)$ is a positive increasing bounded function of γ and therefore can be seen as the c.d.f. of a positive finite measure on \mathbb{R} , the Fourier transform of which is:

$$\int_{\mathbb{R}} \exp(i\gamma x) d\tilde{\varphi}(u; v, \gamma) = \mathbb{E} \int_{\mathbb{R}} \exp(i\gamma x) d\tilde{\varphi}_w(u; v, \gamma),$$

where, for given w , $\tilde{\varphi}_w(u; v, \gamma)$ is the c.d.f. of the mass point $\exp(u'w)$ at $v'w$. We then get:

$$\begin{aligned} \int_{\mathbb{R}} \exp(i\gamma x) d\tilde{\varphi}(u; v, \gamma) &= \mathbb{E}[\exp(ixv'w) \exp(u'w)] \\ &= \mathbb{E}[\exp(u + ivx)'w] \\ &= \varphi(u + ivx). \end{aligned}$$

Let us now compute $A(x_0, \lambda)$ for any real number λ , with:

$$\begin{aligned} &A(x_0, \lambda) \\ &= \frac{1}{2\pi} \int_{-x_0}^{x_0} \frac{\exp(i\lambda x) \varphi(u - ivx) - \exp(-i\lambda) \varphi(u + ivx)}{ix} dx \\ &= \frac{1}{2\pi} \int_{-x_0}^{x_0} \left[\int_{\mathbb{R}} \frac{\exp[-ix(\gamma - \lambda)] - \exp[ix(\gamma - \lambda)]}{ix} d\tilde{\varphi}(u; v, \gamma) \right] dx \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \left[\int_{-x_0}^{x_0} \frac{\exp[-ix(\gamma - \lambda)] - \exp[ix(\gamma - \lambda)]}{ix} dx \right] d\tilde{\varphi}(u; v, \gamma). \end{aligned}$$

Now :

$$\begin{aligned} &\frac{1}{2\pi} \int_{-x_0}^{x_0} \frac{\exp[-ix(\gamma - \lambda)] - \exp[ix(\gamma - \lambda)]}{ix} dx \\ &= \frac{-\text{sign}(\gamma - \lambda)}{\pi} \int_{-x_0}^{x_0} \frac{\sin(x |\gamma - \lambda|)}{x} dx \end{aligned}$$

which tends to $-sign(\gamma - \lambda)$ when $x_0 \rightarrow \infty$ (where $sign(\omega) = 1$ if $\omega > 0$, $sign(\omega) = 0$ if $\omega = 0$, $sign(\omega) = -1$ if $\omega < 0$). Therefore:

$$A(\infty, \lambda) = - \int_{\mathbb{R}} sign(\gamma - \lambda) d\tilde{\varphi}(u; v, \gamma) = -\mathbb{E} \int_{\mathbb{R}} sign(\gamma - \lambda) d\tilde{\varphi}_w(u; \theta, \gamma),$$

where $\tilde{\varphi}_w(u; v, \gamma)$ is the c.d.f. of the mass point $\exp(u'w)$ at $v'w$ and

$$\int_{\mathbb{R}} sign(\gamma - \lambda) d\tilde{\varphi}_w(u; v, \gamma) = \begin{cases} \exp(u'w) & \text{if } \lambda < v'w \\ 0 & \text{if } \lambda = v'w \\ -\exp(u'w) & \text{if } \lambda > v'w. \end{cases}$$

Therefore, we have:

$$\begin{aligned} A(\infty, \lambda) &= -\mathbb{E}[\exp(u'w)(1 - \mathbf{1}_{(v'w < \lambda)}) - \exp(u'w)\mathbf{1}_{(v'w < \lambda)}] \\ &= -\varphi(u) + 2\tilde{\varphi}(u; v, \lambda) \end{aligned}$$

and, further,:

$$\tilde{\varphi}(u; , v, \gamma) = \frac{\varphi(u)}{2} + \frac{1}{2}A(\infty, \gamma),$$

where

$$\begin{aligned} \frac{1}{2}A(\infty, \gamma) &= \frac{1}{4\pi} \int_{-\infty}^{\infty} \frac{\exp(i\gamma x)\varphi(u - ivx) - \exp(-i\gamma x)\varphi(u + ivx)}{ix} dx \\ &= \frac{1}{2\pi} \int_o^{\infty} \frac{\exp(i\gamma x)\varphi(u - ivx) - \exp(-i\gamma x)\varphi(u + ivx)}{ix} dx \\ &= -\frac{1}{\pi} \int_o^{\infty} \frac{\mathcal{I}m[\exp(-i\gamma x)\varphi(u + ivx)]}{x} dx, \end{aligned}$$

which leads to Eq. (1.23). □

Chapter 2

Pricing and risk-neutral dynamics

In brief: bla bla bla bla.

2.1 Consumption-based Capital Asset Pricing Model (CCAPM) and stochastic discount factor (SDF)

Consider an economy featuring a single good, whose date- t price is q_t . There is a representative agent with an external income R_t at t , and a portfolio of assets, with an allocation vector α_{t-1} (decided at $t-1$). The vector of date- t prices is p_t . At any date $t+j$, $j = 0, 1, \dots$, the agent will face the budget constraint:

$$q_{t+j}C_{t+j} + \alpha'_{t+j}p_{t+j} = R_{t+j} + \alpha'_{t+j-1}p_{t+j},$$

where C_{t+j} is her consumption on date $t+j$. The representative agent maximizes her expected utility time-separable preferences $\mathbb{E}_t \sum_{j=0}^{\infty} \delta^j U(C_{t+j})$ subject to the budget constraints at $t+j$, $j = 0, 1, \dots$, where U is the utility function and δ is time discount factor or subjective discount factor. Note that the latter is different from SDF.

Replacing C_{t+j} with $[R_{t+j} - (\alpha'_{t+j} - \alpha'_{t+j-1})p_{t+j}]/q_{t+j}$, the objective function becomes:

$$\mathbb{E}_t \sum_{j=0}^{\infty} \delta^j U[R_{t+j}/q_{t+j} - (\alpha'_{t+j} - \alpha'_{t+j-1})p_{t+j}/q_{t+j}].$$

The vector of allocation α_t appears in the first two terms:

$$U[R_t/q_t - (\alpha'_t - \alpha'_{t-1})p_t/q_t] + \delta \mathbb{E}_t U[R_{t+1}/q_{t+1} - (\alpha'_{t+1} - \alpha'_t)p_{t+1}/q_{t+1}].$$

As a result, the first order condition associated with vector α_t reads

$$\frac{p_t}{q_t} \frac{dU(C_t)}{dC} = \delta \mathbb{E}_t \left[\frac{p_{t+1}}{q_{t+1}} \frac{dU(C_{t+1})}{dC} \right],$$

or

$$p_t = \mathbb{E}_t(p_{t+1} \mathcal{M}_{t,t+1}), \quad (2.1)$$

where $\mathcal{M}_{t,t+1}$ is a strictly positive scalar called **stochastic discount factor (SDF)** between dates t and $t+1$. We have:

$$\mathcal{M}_{t,t+1} \equiv \delta \frac{q_t}{q_{t+1}} \frac{\frac{dU(C_{t+1})}{dC}}{\frac{dU(C_t)}{dC}}.$$

Example 2.1 (Power utility function). A standard case is the one of the power utility $U(C) = \frac{C^{1-\gamma}-1}{1-\gamma}$, where $\gamma > 0$ is the coefficient of relative risk aversion.

We have $U'(C) = C^{-\gamma} > 0$ and $U''(C) = -\gamma C^{-\gamma-1} < 0$. As a result, the stochastic discount factor is

$$\begin{aligned} & \mathcal{M}_{t,t+1} \\ &= \frac{q_t}{q_{t+1}} \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \\ &= \exp(\log \delta + \log q_t + \gamma \log C_t - \log q_{t+1} - \gamma \log C_{t+1}). \end{aligned} \quad (2.2)$$

The vector of prices p_t satisfies:

$$p_t = \delta q_t C_t^\gamma \mathbb{E}_t \left(\frac{C_{t+1}^{-\gamma}}{q_{t+1}} p_{t+1} \right),$$

or (Euler equation):

$$\mathbb{E}_t \left[\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \frac{q_t}{q_{t+1}} \frac{p_{t+1}}{p_t} - 1 \right] = 0.$$

According to Eq. (2.1), for any asset j :

$$p_{j,t} = \mathbb{E}_t(\mathcal{M}_{t,t+1}p_{j,t+1}). \quad (2.3)$$

Using that $p_{j,t+1} = \mathbb{E}_t(\mathcal{M}_{t+1,t+2}p_{j,t+2})$, we get:

$$\begin{aligned} p_{j,t} &= \mathbb{E}_t[\mathbb{E}_{t+1}(p_{j,t+2}\mathcal{M}_{t+1,t+2})\mathcal{M}_{t,t+1}] \\ &= \mathbb{E}_t(\mathcal{M}_{t,t+1}\mathcal{M}_{t+1,t+2}p_{j,t+2}). \end{aligned}$$

This can be generalized as follows:

$$p_{j,t} = \mathbb{E}_t[\mathcal{M}_{t,t+1} \dots \mathcal{M}_{t+h-1,t+h}p_{j,t+h}], \quad \forall h.$$

2.2 Recursive utilities

XXXX

2.3 SDF: Absence of Arbitrage Approach

Consider a period of interest $\mathcal{T} = \{0, 1, 2, \dots, T^*\}$. As in Chapter 1, vector w_t constitutes the new information in the economy at t . The historical, or physical, dynamics of w_t , $f(w_t)$, is defined by $f(w_{t+1}|w_t)$. The physical probability is denoted by \mathbb{P} . $\bar{L}_{2t}, t \in \mathcal{T}$, is the (Hilbert) space of square integrate functions $g(\underline{w}_t)$, and we have $L_{2t} \subset L_{2s}, t < s$.

2.3.1 Existence and unicity of the SDF

Hypothesis 2.1 (Price existence and uniqueness). For any \underline{w}_t , there exists a unique $p_t[g(\underline{w}_s)]$, function of \underline{w}_t , price at t of a payoff $g(\underline{w}_s)$ delivered at $s, \forall t \leq s$.

Hypothesis 2.2 (Linearity and continuity). For all $t < s$, $\underline{w}_t, g_1, g_2$, we have

$$\bullet \quad p_t[\lambda_1 g_1(\underline{w}_s) + \lambda_2 g_2(\underline{w}_s)] = \lambda_1 p_t[g_1(\underline{w}_s)] + \lambda_2 p_t[g_2(\underline{w}_s)],$$

- If $g_n(\underline{w}_s) \xrightarrow[n \rightarrow \infty]{L_{2s}} 0$, then $p_t[g_n(\underline{w}_s)] \xrightarrow[n \rightarrow \infty]{} 0$.

Hypothesis 2.3 (Absence of Arbitrage Opportunity (AAO)). At any t , it is impossible to constitute a portfolio of future payoffs, possibly modified at subsequent dates, such that:

- the price of the portfolio at t is zero,
- payoffs at subsequent dates are ≥ 0 ,
- there is at least one subsequent date s such that the net payoff at s is strictly positive with a non zero conditional probability at t .

Theorem 2.1 (Riesz representation theorem). *Under Assumptions 2.1 and 2.2, for all \underline{w}_t , and $s > t$, there exists $\mathcal{M}_{t,s}(\underline{w}_s) \in L_{2s}$, unique such that, $\forall g(\underline{w}_s) \in L_{2s}$,*

$$p_t[g(\underline{w}_s)] = \mathbb{E}[\mathcal{M}_{t,s}(\underline{w}_s)g(\underline{w}_s)|\underline{w}_t].$$

In particular the price at t of a zero coupon bond maturing at s is $\mathbb{E}(\mathcal{M}_{t,s}|\underline{w}_t)$.

Proposition 2.1 (Positivity of \mathcal{M}). *If Assumption 2.3 is satisfied, then for all t and s , $\mathbb{P}(\mathcal{M}_{t,s} > 0|\underline{w}_t) = 1$.*

Proof. \Leftarrow is obvious. If \Rightarrow was not true, the payoff $\mathbf{1}_{\{\mathcal{M}_{t,s} \leq 0\}}$, at s , would be such that: $\mathbb{P}[\mathbf{1}_{\{\mathcal{M}_{t,s} \leq 0\}} = 1|\underline{w}_t] > 0$ and $p_t[\mathbf{1}_{\{\mathcal{M}_{t,s} \leq 0\}}] = \mathbb{E}_t[\mathcal{M}_{t,s}\mathbf{1}_{\{\mathcal{M}_{t,s} \leq 0\}}] \leq 0$. \square

Proposition 2.2 (Time consistency). *For all $t < r < s$, we have $\mathcal{M}_{t,s} = \mathcal{M}_{t,r}\mathcal{M}_{r,s}$, which implies:*

- $\mathcal{M}_{t,s} = \mathcal{M}_{t,t+1}\mathcal{M}_{t+1,t+2} \dots \mathcal{M}_{s-1,s}$
- $\mathcal{M}_{0,t} = \prod_{j=0}^{t-1} \mathcal{M}_{j,j+1}$ ($\mathcal{M}_{0,t}$ is called **pricing kernel**).

Proof. Using Lemma 2.1 we have:

$$\begin{aligned} p_t(g_s) &= \mathbb{E}(\mathcal{M}_{t,s}g_s|\underline{w}_t) = \mathbb{E}(\mathcal{M}_{t,r}p_r(g_s)|\underline{w}_t) \\ &= \mathbb{E}[\mathcal{M}_{t,r}\mathbb{E}(\mathcal{M}_{r,s}g_s|\underline{w}_r)|\underline{w}_t] = \mathbb{E}(\mathcal{M}_{t,r}\mathcal{M}_{r,s}g_s|\underline{w}_t), \forall g, \forall \underline{w}_t \end{aligned}$$

and, therefore, $\mathcal{M}_{t,s} = \mathcal{M}_{t,r}\mathcal{M}_{r,s}$. \square

Lemma 2.1. *For any payoff g_s at s , $p_t(g_s) = p_t[p_r(g_s)]$.*

Proof. If this was not true, we could construct a sequence of portfolios with a strictly positive payoff at s with zero payoff at any other future date and with price zero at t , contradicting Assumption 2.3. Indeed, assuming, for instance, $p_t(g_s) > p_t[p_r(g_s)]$, the payoff at s is defined by the following strategy: (i) at t : buy $p_r(g_s)$, (short) sell g_s , buy $\frac{p_t(g_s) - p_t[p_r(g_s)]}{\mathbb{E}(\mathcal{M}_{t,s}|\underline{w}_t)}$ zero-coupon bonds maturing at s , at global price zero, (ii) at r : buy g_s and sell $p_r(g_s)$, generating a zero net payoff, (iii) at s , the net payoff is: $g_s - g_s + \frac{p_t(g_s) - p_t[p_r(g_s)]}{\mathbb{E}(\mathcal{M}_{t,s}|\underline{w}_t)} > 0$. \square

Consider an asset whose payoff, on date s , is $g(\underline{w}_s)$. We have, $\forall t < s$:

$$\boxed{p_t[g(\underline{w}_s)] = \mathbb{E}_t[\mathcal{M}_{t,t+1} \dots \mathcal{M}_{s-1,s} g(\underline{w}_s)]}. \quad (2.4)$$

In particular, since $\mathcal{M}_{t,t+1}$ contains 1, the price at t of a zero-coupon with residual maturity one is given by:

$$B(t, 1) := \mathbb{E}_t[\mathcal{M}_{t,t+1}].$$

Denoting by r_t the continuously-compounded interest rate, defined through $B(t, 1) = \exp(-r_t)$, we get

$$r_t = -\log \mathbb{E}_t[\mathcal{M}_{t,t+1}].$$

Definition 2.1 (Bank account). The bank account process R_t is defined by $R_t \equiv \exp(r_0 + \dots + r_{t-1}) = \frac{1}{\mathbb{E}_0[\mathcal{M}_{0,1}] \times \dots \times \mathbb{E}_{t-1}[\mathcal{M}_{t-1,t}]}$.

R_t is the price of an investment initiated on date 0, when it was worth one dollar, and invested on each date at the risk-free rate (for one period).

For any price process p_t , we have $p_t = \mathbb{E}_t(\mathcal{M}_{t,s} p_s)$ (with $s > t$), or $\mathcal{M}_{0,t} p_t = \mathbb{E}_t(\mathcal{M}_{0,s} p_s)$. That is, $\mathcal{M}_{0,t} p_t$ is a martingale. In particular $\mathcal{M}_{0,t} R_t$ is a martingale.

2.3.2 Exponential Affine SDF

A specific (tractable) case is that of exponential affine SDF. Assume that

$$\mathcal{M}_{t,t+1}(\underline{w}_{t+1}) = \exp[\alpha_t(\underline{w}_t)' \underline{w}_{t+1} + \beta_t(\underline{w}_t)]$$

where α_t defines the *prices of risk* or *sensitivity* vector. Using $\mathbb{E}_t[\mathcal{M}_{t,t+1}] = \exp(-r_t) = \exp[\psi_t(\alpha_t) + \beta_t]$, we get:

$$\mathcal{M}_{t,t+1} = \exp[-r_t + \alpha'_t w_{t+1} - \psi_t(\alpha_t)]. \quad (2.5)$$

Example 2.2 (CCAPM/Power utility case). In the CCAPM-power-utility case (see Example 2.1), we have (Eq. (2.2)):

$$\mathcal{M}_{t,t+1} = \exp(\log \delta + \log q_t + \gamma \log C_t - \log q_{t+1} - \gamma \log C_{t+1}),$$

where q_t is the price of the consumption good, C_t is the quantity consumed at t and δ is the intertemporal discount rate.

Hence, in that case, $\mathcal{M}_{t,t+1}$ is exponential affine in $w_{t+1} = (\log q_{t+1}, \log C_{t+1})'$ (and its first lag).

2.4 The risk-neutral (R.N.) dynamics

The historical Dynamics is characterized by $f(w_{T^*})$, or by the sequence of conditional p.d.f. $f_{t+1}(w_{t+1}|\underline{w}_t)$, or $f_{t+1}(w_{t+1})$, with respect to (w.r.t.) some measure μ .

We define the conditional risk-neutral p.d.f. w.r.t. the conditional historical probability. For that, we employ the Radon-Nikodym derivative $d_{t+1}^{\mathbb{Q}}(w_{t+1}|\underline{w}_t)$:¹

$$d_{t+1}^{\mathbb{Q}}(w_{t+1}|\underline{w}_t) = \frac{\mathcal{M}_{t,t+1}(\underline{w}_{t+1})}{\mathbb{E}[\mathcal{M}_{t,t+1}(\underline{w}_{t+1})|\underline{w}_t]},$$

or

$$d_{t+1}^{\mathbb{Q}}(w_{t+1}) = \frac{\mathcal{M}_{t,t+1}}{\mathbb{E}_t(\mathcal{M}_{t,t+1})} = \exp(r_t) \mathcal{M}_{t,t+1}.$$

In this context, the risk neutral conditional p.d.f. is:

$$\begin{aligned} f_{t+1}^{\mathbb{Q}}(w_{t+1}) &= f_{t+1}(w_{t+1}) d_{t+1}^{\mathbb{Q}}(w_{t+1}) \\ &= f_{t+1}(w_{t+1}) \mathcal{M}_{t,t+1}(\underline{w}_{t+1}) \exp[r_t(\underline{w}_t)]. \end{aligned} \quad (2.6)$$

¹Of course, the conditional historical p.d.f. with respect to the conditional risk-neutral (R.N.) p.d.f. is: $d_{t+1}(w_{t+1}) = \frac{1}{d_{t+1}^{\mathbb{Q}}(w_{t+1})}$ or $d_{t+1}(w_{t+1}) = \frac{\exp(-r_t)}{\mathcal{M}_{t,t+1}}$.

The p.d.f. of \mathbb{Q} w.r.t. the historical dynamics \mathbb{P} is:

$$\xi_{T^*} = \frac{d\mathbb{Q}}{d\mathbb{P}} = \Pi_{t=0}^{T^*-1} d_{t+1}^{\mathbb{Q}}(w_{t+1}) = \Pi_{t=0}^{T^*-1} \exp(r_t) \mathcal{M}_{t,t+1},$$

and the p.d.f. of the R.N. distribution of \underline{w}_t , w.r.t. the corresponding historical distribution is:

$$\xi_t = \Pi_{\tau=1}^{t-1} d_{\tau+1}^{\mathbb{Q}}(w_{\tau+1}) = \mathbb{E}_t \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) = \mathbb{E}_t \xi_{T^*}.$$

Therefore, ξ_t is a \mathbb{P} -martingale.²

Consider the date- t price of a payoff $g(\underline{w}_s)$ at time $s > t$. An equivalent form of the pricing formula (2.4) is:

$$\begin{aligned} p_t[g(\underline{w}_s)] &= \mathbb{E}_t[\mathcal{M}_{t,t+1} \dots \mathcal{M}_{s-1,s} g(\underline{w}_s)] \\ &= \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{s-1}) g(\underline{w}_s)], \end{aligned}$$

or, with simpler notations:

$$p_t = \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{s-1}) p_s] = \mathbb{E}_t^{\mathbb{Q}} \left(\frac{R_t}{R_s} p_s \right),$$

where R_t is the *bank account*.

We also have $p_t/R_t = \mathbb{E}_t^{\mathbb{Q}}(p_s/R_s)$, that is, p_t/R_t is a \mathbb{Q} -martingale. In particular $p_t = \exp(-r_t) \mathbb{E}_t^{\mathbb{Q}}(p_{t+1})$, or, using the arithmetic return of any payoff $(p_{t+1} - p_t)/p_t$, and the arithmetic return of the riskless asset $r_{A,t+1} = \exp(r_t) - 1$, we get:

$$\mathbb{E}_t^{\mathbb{Q}} \left(\frac{p_{t+1} - p_t}{p_t} \right) = r_{A,t}.$$

Moreover the excess arithmetic return process $(p_{t+1} - p_t)/p_t - r_{A,t}$ is a \mathbb{Q} -martingale difference and, therefore, \mathbb{Q} -serially uncorrelated.

²Indeed:

$$\mathbb{E}_t \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) = \Pi_{\tau=1}^{t-1} d_{\tau+1}^{\mathbb{Q}}(w_{\tau+1}) \mathbb{E}_t \left(d_{t+1}^{\mathbb{Q}}(w_{t+1}) \dots d_{T^*}^{\mathbb{Q}}(w_{T^*}) \right).$$

Let us consider the case of an exponential affine SDF $\mathcal{M}_{t,t+1} = \exp(\alpha'_t w_{t+1} + \beta_t)$:

$$d_{t+1}^{\mathbb{Q}}(w_{t+1}) = \frac{\mathcal{M}_{t,t+1}}{\mathbb{E}_t(\mathcal{M}_{t,t+1})} = \frac{\exp(\alpha'_t w_{t+1} + \beta_t)}{\exp[\psi_t(\alpha_t) + \beta_t]} = \exp[\alpha'_t w_{t+1} - \psi_t(\alpha_t)]$$

We then have that $d_{t+1}^{\mathbb{Q}}(w_{t+1})$ is also exponential affine. Moreover:

$$f_{t+1}^{\mathbb{Q}}(w_{t+1}) = \frac{f_{t+1}(w_{t+1}) \exp(\alpha'_t w_{t+1})}{\varphi_t(\alpha_t)}.$$

The previous equation shows that $f_{t+1}^{\mathbb{Q}}$ is the Esscher transform of f_{t+1} evaluated at α_t .

Let us now consider the Laplace transform of the conditional R.N. probability, $\varphi_t^{\mathbb{Q}}(u|\underline{w}_t)$, also denoted by $\varphi_t^{\mathbb{Q}}(u)$. We have:

$$\begin{aligned} \varphi_t^{\mathbb{Q}}(u) &= \mathbb{E}_t^{\mathbb{Q}} \exp(u' w_{t+1}) \\ &= \mathbb{E}_t \exp[(u + \alpha_t)' w_{t+1} - \psi_t(\alpha_t)] \\ &= \exp[\psi_t(u + \alpha_t) - \psi_t(\alpha_t)] = \frac{\varphi_t(u + \alpha_t)}{\varphi_t(\alpha_t)}. \end{aligned}$$

Hence:

$$\boxed{\psi_t^{\mathbb{Q}}(u) = \psi_t(u + \alpha_t) - \psi_t(\alpha_t)}. \quad (2.7)$$

We check that, if $\alpha_t = 0$, $\psi_t^{\mathbb{Q}} = \psi_t$ (since $\psi_t(0) = 0$).

Moreover, putting $u = -\alpha_t$ in the expression of $\psi_t^{\mathbb{Q}}(u)$ we get $\psi_t^{\mathbb{Q}}(-\alpha_t) = -\psi_t(\alpha_t)$, and, replacing u by $u - \alpha_t$, we get:

$$\boxed{\psi_t(u) = \psi_t^{\mathbb{Q}}(u - \alpha_t) - \psi_t^{\mathbb{Q}}(-\alpha_t)}.$$

Also:

$$\begin{cases} d_{t+1}(w_{t+1}) &= \exp[-\alpha'_t(w_{t+1}) - \psi_t^{\mathbb{Q}}(-\alpha_t)] \\ d_{t+1}^{\mathbb{Q}}(w_{t+1}) &= \exp[\alpha'_t(w_{t+1}) + \psi_t^{\mathbb{Q}}(-\alpha_t)]. \end{cases}$$

2.5 Typology of Econometric Asset Pricing Models

Definition 2.2 (Econometric Asset Pricing Model (EAPM)). An Econometric Asset Pricing Model (EAPM) is defined by the following functions:

- $r_t(\underline{w}_t)$,
- $f(w_{t+1}|\underline{w}_t)$ [or $\psi_t(u)$],
- $\mathcal{M}_{t,t+1}(\underline{w}_{t+1})$,
- $f^{\mathbb{Q}}(w_{t+1}|\underline{w}_t)$ [or $\psi_t^{\mathbb{Q}}(u)$].

The previous functions have to be specified and parameterized. They are linked by:

$$f^{\mathbb{Q}}(w_{t+1}|\underline{w}_t) = f(w_{t+1}|\underline{w}_t)\mathcal{M}_{t,t+1}(\underline{w}_{t+1})\exp[r_t(\underline{w}_t)].$$

In the following, we present three ways of specifying an EAPM:

1. the direct modelling,
2. the R.N.-constrained direct modelling (or mixed modelling),
3. the back modelling.

We focus on the case where $\mathcal{M}_{t,t+1}$ is exponential affine, as in Eq. (2.5):

$$\mathcal{M}_{t,t+1}(\underline{w}_{t+1}) = \exp \left\{ -r_t(\underline{w}_t) + \alpha'_t(\underline{w}_t)w_{t+1} - \psi_t[\alpha_t(w_t)] \right\}.$$

Once the short-term rate function $r_t(\underline{w}_t)$ is specified, we have to specify ψ_t , α_t , and $\psi_t^{\mathbb{Q}}$, that are linked by Eq. (2.7).

In all approaches, we have to specify the status of the short rate. The short rate r_t is a function of \underline{w}_t , this function may be known or unknown by the econometrician. It is known in two cases: (a) r_t is exogenous ($r_t(\underline{w}_t)$ does not depend on \underline{w}_t) or (b) r_t is a component of w_t . By contrast, if the function $r_t(\underline{w}_t)$ is unknown, it has to be specified parametrically:

$$\left\{ r_t(\underline{w}_t, \tilde{\theta}), \tilde{\theta} \in \tilde{\Theta} \right\},$$

where $r_t(\bullet, \bullet)$ is a known function.

Let us now detail the three steps on which each of the three ways of defining an EAPM is based.

2.5.1 The Direct Modelling

- **Step 1 – Specification of the historical dynamics.** We choose a parametric family for the conditional historical Log-Laplace transform $\psi_t(u|\underline{w}_t): \{\psi_t(u|\underline{w}_t; \theta_1), \theta_1 \in \Theta_1\}$.
- **Step 2 – Specification of the SDF.** Considering the affine specification of as Eq. (2.5), that is:

$$\mathcal{M}_{t,t+1}(\underline{w}_{t+1}) = \exp \left\{ -r_t(\underline{w}_t, \tilde{\theta}) + \alpha'_t(\underline{w}_t) w_{t+1} - \psi_t[\alpha_t(\underline{w}_t)|\underline{w}_t; \theta_1] \right\},$$

we need to specify functions $r_t(\underline{w}_t, \tilde{\theta})$ and $\alpha_t(\underline{w}_t)$. Assume that $\alpha_t(\underline{w}_t)$ belongs to a parametric family: $\{\alpha_t(\underline{w}_t; \theta_2), \theta_2 \in \Theta_2\}$. We then have:

$$\mathcal{M}_{t,t+1}(\underline{w}_{t+1}, \theta) = \exp \left\{ -r_t(\underline{w}_t, \tilde{\theta}) + \alpha'_t(\underline{w}_t, \theta_2) w_{t+1} - \psi_t[\alpha_t(\underline{w}_t, \theta_2)|\underline{w}_t; \theta_1] \right\},$$

where $\theta = (\tilde{\theta}', \theta'_1, \theta'_2)' \in \tilde{\Theta} \times \Theta_1 \times \Theta_2 = \Theta$.

- **Step 3 – Internal consistency conditions (ICC).** For any payoff $g(\underline{w}_s)$ settled at $s > t$, with price $p(\underline{w}_t)$ at t which is a known function of \underline{w}_t , we must have:

$$p(\underline{w}_t) = \mathbb{E} \left\{ \mathcal{M}_{t,t+1}(\theta) \dots \mathcal{M}_{s-1,s}(\theta) g(\underline{w}_s) | \underline{w}_t, \theta_1 \right\} \forall \underline{w}_t, \theta.$$

These ICC pricing conditions may imply strong constraints on θ . For instance, when components of w_t are returns of some assets: if $w_{1,t} = \log(p_{1,t}/p_{1,t-1})$, then we must have $\mathbb{E}_t[\mathcal{M}_{t,t+1} \exp(e'_1 w_{t+1})] = 1$ (Euler equation). Or, in the case of interest rates with various maturities: if $w_{1,t} = -1/h \log B(t, h)$, then we must have $e'_1 w_t = -1/h \log \mathbb{E}_t(\mathcal{M}_{t,t+1} \times \dots \times \mathcal{M}_{t+h-1,t+h})$.

The previous three steps imply the specification of the R.N. dynamics (according to Eq. (2.7)):

$$\psi^{\mathbb{Q}}(u|\underline{w}_t, \theta_1, \theta_2) = \psi_t[u + \alpha_t(\underline{w}_t, \theta_2)|\underline{w}_t, \theta_1] - \psi_t[\alpha_t(\underline{w}_t, \theta_2)|\underline{w}_t, \theta_1].$$

2.5.2 The R.N.-constrained direct modelling (or mixed modelling)

- **Step 1 – Specification of the physical dynamics.** We select a family $\{\psi_t(u|\underline{w}_t, \theta_1), \theta_1 \in \Theta_1\}$.

- **Step 2 – Specification of the risk-neutral dynamics.** We select a family $\{\psi_t^{\mathbb{Q}}(u|\underline{w}_t, \theta^*), \theta^* \in \Theta^*\}$ and, possibly, $\{r_t(\underline{w}_t, \tilde{\theta}), \tilde{\theta} \in \tilde{\Theta}\}$.
- **Step 3 – Internal Consistency Conditions (ICC).** Once the parameterization $(\tilde{\theta}, \theta_1, \theta^*) \in \tilde{\Theta} \times \Theta_1^* \times \Theta^*$ is defined, ICCs may be imposed. For instance, if $w_{1,t} = \log(p_{1,t}/p_{1,t-1})$, then we must have $\exp(-r_t)\mathbb{E}_t^{\mathbb{Q}} \exp(e'_1 w_{t+1}) = 1$. Or if $w_{1,t} = B(t, h)$, then $e'_1 w_t = \mathbb{E}_t^{\mathbb{Q}} \exp(-r_t - \dots - r_{t+h-1})$.

The SDF is a by-product. If we want an exponential affine SDF, for any pair $(\psi_t^{\mathbb{Q}}, \psi_t)$ belonging to these families, there must exist a unique function $\alpha_t(\underline{w}_t)$ denoted by $\alpha_t(w_t; \theta_1, \theta^*)$, and satisfying:

$$\psi_t^{\mathbb{Q}}(u|\underline{w}_t) = \psi_t \left[u + \alpha_t(w_t)|\underline{w}_t \right] - \psi_t \left[\alpha_t(\underline{w}_t)|\underline{w}_t \right].$$

2.5.3 The Back Modelling (based on three steps)

- **Step 1 – Specification of the R.N. dynamics,** and possibly of $r_t(\underline{w}_t)$: $\psi_t^{\mathbb{Q}}(u|\underline{w}_t; \theta_1^*)$.
- **Step 2 – Internal consistency conditions (ICC),** if relevant, are taken into account:

$$p(\underline{w}_t) = \mathbb{E}_t^{\mathbb{Q}} \left[\exp(-r_t(\underline{w}_t, \tilde{\theta}) - \dots - r_{s-1}(\underline{w}_s, \tilde{\theta})) g(\underline{w}_s) | \underline{w}_t, \theta_1^* \right], \\ \forall \underline{w}_t, \tilde{\theta}, \theta_1^*.$$

- **Step 3 – Choice of the specification of the prices of risk.** One chooses function $\alpha_t(w_t)$ without any constraint; this amounts to defining the family $\{\alpha_t(\underline{w}_t, \theta_2^*), \theta_2^* \in \Theta_2^*\}$.

The historical dynamics is obtained as a by-product. Indeed:

$$\psi_t(u|\underline{w}_t; \theta_1^*, \theta_2^*) = \psi_t^{\mathbb{Q}} \left[u - \alpha_t(\underline{w}_t, \theta_2^*) | \underline{w}_t; \theta_1^* \right] - \psi_t^{\mathbb{Q}} \left[-\alpha_t(\underline{w}_t, \theta_2^*) | \underline{w}_t, \theta_1^* \right].$$

Chapter 3

The term structure or risk-free yields

3.1 Introduction

Risk-free yields are the yields-to-maturity associated with bonds that carry no default and/or liquidity risks. Bonds issued by sovereign entities with top credit quality are usually considered to be risk-free.

An important share of the term-structure literature pertains to the modelling of risk-free yields. Some models involve macro factors in w_t (Ang and Piazzesi, 2003); some do not (Duffie and Singleton, 1997). The latter are sometimes called *yield-only* models.

The basic pricing formula of a risk-free zero coupon bond is (see Eq. (1.9) in Example 1.20):

$$B(t, h) = \exp(-r_t) \mathbb{E}_t^{\mathbb{Q}} \exp(-r_{t+1} - \dots - r_{t+h-1}) \quad (3.1)$$

$$R(t, h) = -\frac{1}{h} \log B(t, h). \quad (3.2)$$

Term structure models are often used to extract **term premiums** from observed yields-to-maturity. Term premiums are those components of yields that would not exist if investors were not risk-averse.

If agents were not risk averse, i.e., under the **Expectation Hypothesis (EH)**, we would have $\mathcal{M}_{t,t+1} = \exp(-r_t)$. We would then have “ $\mathbb{P} \equiv \mathbb{Q}$ ”.

and $B(t, h)$ would become:

$$\exp(-r_t) \mathbb{E}_t \exp(-r_{t+1} - \dots - r_{t+h-1}).. \quad (3.3)$$

And the maturity- h yield-to-maturity would then be:

$$\begin{aligned} R^{EH}(t, h) &= -\frac{1}{h} \log (\mathbb{E}_t \exp(-r_t - \dots - r_{t+h-1})) \\ &\approx \frac{1}{h} \mathbb{E}_t (r_t + \dots + r_{t+h-1}). \end{aligned} \quad (3.4)$$

The term premium is given by:

$$\begin{aligned} TP_{t,h} &= \underbrace{-\frac{1}{h} \log \mathbb{E}_t^{\mathbb{Q}} \exp(-r_{t+1} - \dots - r_{t+h-1})}_{=R(t,h)} - \\ &\quad \underbrace{-\frac{1}{h} \log \mathbb{E}_t \exp(-r_{t+1} - \dots - r_{t+h-1})}_{=R^{EH}(t,h)}. \end{aligned} \quad (3.5)$$

What is the economic meaning of the term premium? Under EH, investors are willing to buy a maturity- h bond as long as its expected return is—up to Jensen’s inequality—equal to the average of future short-term rates. (Hence the definition of $R^{EH}(t, h)$, see Eq. (3.4).) When $TP_{t,h} > 0$, investors are willing to buy the maturity- h bond only if its return is, on average, higher than expected future short-term rates; this corresponds to a situation where investors consider that long-term bonds tend to lose value in *bad states of the world* (i.e., states of high marginal utility).

Risk premium: According to Eq. (2.3), the price of any asset j satisfies:

$$p_{jt} = \mathbb{E}_t(\mathcal{M}_{t,t+1} p_{j,t+1}).$$

The previous equation rewrites:

$$p_{jt} = \text{Cov}_t(\mathcal{M}_{t,t+1}, p_{j,t+1}) + \mathbb{E}_t(\mathcal{M}_{t,t+1}) \mathbb{E}_t(p_{j,t+1})$$

or

$$p_{jt} = \underbrace{\exp(-r_t) \mathbb{E}_t(p_{j,t+1})}_{=p_{jt}^{EH}} + \underbrace{\text{Cov}_t(\mathcal{M}_{t,t+1}, p_{j,t+1})}_{\text{Risk premium}}. \quad (3.6)$$

If investors were not risk-averse, then we would have $p_{jt} = p_{jt}^{EH}$. The S.D.F. is high (resp. low) in bad (resp. good) states of the world (states of high marginal utility in the equilibrium approach). Hence, we have $p_{jt} < p_{jt}^{EH}$ if asset j tends to pay less in bad states of the world (i.e., if $\text{Cov}_t(\mathcal{M}_{t,t+1}, p_{j,t+1}) < 0$).

3.2 Swap rates and forward rates

3.2.1 Swap rates

Bonds issued by top-rated (Aaa/AAA) countries are often considered to be risk-free. Because of call-margins mechanisms, swap rates are also used as risk-free benchmarks (Duffie and Stein, 2015).

Definition 3.1 (Interest Rate Swap (IRS)). In an Interest Rate Swap (IRS), a *fixed-rate payer* agrees to provide the *fixed-rate receiver* with a sequence of cash flows that are determined at the negotiation date of the swap, and at predetermined dates. These cash flows constitute the *fixed leg* of the swap. Conversely, the fixed-rate receiver provides the fixed-rate payer with cash-flows that depend on future values of a reference rate; this is the *floating rate* of the swap.

More precisely, the dates of payment are of the form $t + \tau, t + 2\tau, \dots, t + n\tau$, where τ is a period expressed in years (typically 1/2 or 1/4) and n is the number of payments. The maturity, or *tenor*, of the swap contract is $h = n\tau$.

The payoffs of the fixed leg are τS , where S is the annualized payment (or *swap rate*). On date $t + j\tau$, the payoff of the floating leg is $\tau L(t + (j-1)\tau, \tau)$, where L , the annualized linear rate is given by:

$$L(t + (j-1)\tau, \tau) = \frac{1 - B(t + (j-1)\tau, \tau)}{\tau B(t + (j-1)\tau, \tau)},$$

where $B(t + (j-1)\tau, \tau)$ is the price, at date $t + (j-1)\tau$ of a bond of maturity τ .

On the negotiation date, the values of the fixed and floating legs are identical, so that the value of the swap is zero.

At t , the price of the fixed leg is:

$$\sum_{j=1}^n \tau S B(t, j\tau).$$

Let us turn to the price of the floating leg. The payoff at date $t + j\tau$, that is $\frac{1-B(t+(j-1)\tau, \tau)}{B(t+(j-1)\tau, \tau)}$ is known on date $t + (j-1)\tau$, so its price at date $t + (j-1)\tau$ is $1 - B(t + (j-1)\tau, \tau)$, and its price at t therefore is $B(t, (j-1)\tau) - B(t, j\tau)$. Summing over $j = 1, \dots, n$, the date- t price of the floating leg is $1 - B(t, n\tau)$ (independent of the payment dates).

Since the price of the contract is zero at date t (by definition of the swap), we must have:

$$\sum_{j=1}^n \tau S B(t, j\tau) = 1 - B(t, n\tau) \Rightarrow S = \frac{1 - B(t, n\tau)}{\tau \sum_{j=1}^n B(t, j\tau)},$$

or

$$S(t, h) = \frac{1 - B(t, h)}{\tau \sum_{j=1}^{h/\tau} B(t, j\tau)}.$$

Note that all the terms appearing in the previous formula are available in closed-form in the context of an affine model (see Example 1.20).

3.2.2 Forward rates

Definition 3.2 (Forward Rate Agreement (FRA)). An interest rate forward contract is a contract in which the rate to be paid or received on a specific obligation for a set period, beginning in the future, is set at contract initiation.

Denote by $f(t, h_1, h_2)$ the forward interest rate, set on date t , for the period between $t + h_1$ and $t + h_2$. Let us relate $f(t, h_1, h_2)$ to bond rates. For that, consider two strategies (decided on date t):

1. Buy a zero-coupon bond of maturity h_2 (price $B(t, h_2)$) and sell zero-coupon bonds of maturity h_1 for the same amount (yielding a payoff of $B(t, h_2)/B(t, h_1)$ on date $t + h_1$).

2. Enter a forward rate agreement between dates $t+h_1$ and $t+h_2$, whereby you receive 1 on date $t+h_2$.

These two strategies deliver the same payoffs on date t (the payoff is zero) and on date $t+h_2$ (the payoff is 1). By absence of arbitrage, the payoffs on date $t+h_1$ have to be the same. Therefore

$$\begin{aligned} \exp(-(h_2 - h_1)f(t, h_1, h_2)) &= B(t, h_2)/B(t, h_1) \\ \Rightarrow f(t, h_1, h_2) &= \frac{1}{h_2 - h_1}(\log[B(t, h_1)] - \log[B(t, h_2)]), \end{aligned}$$

which gives:

$$\boxed{f(t, h_1, h_2) = \frac{h_2 R(t, h_2) - h_1 R(t, h_1)}{h_2 - h_1}.} \quad (3.7)$$

In an affine model (where $R(t, h)$ is an affine function of the state vector w_t), forward rates are linear in the state vector w_t .

3.3 The Affine Case

3.3.1 Affine yields

In this subsection, we consider the case where the state vector w_t is affine under both \mathbb{P} and \mathbb{Q} . If the nominal short-term rate is affine in w_t , i.e., if $r_t = \omega_0 + \omega'_1 w_t$, then:

$$\begin{aligned} B(t, h) &= \mathbb{E}_t^{\mathbb{Q}} \exp(-r_t - \dots - r_{t+h-1}) \\ &= \exp(-h\omega_0 - \omega'_1 w_t) \mathbb{E}_t^{\mathbb{Q}} \exp(-\omega'_1 w_{t+1} - \dots - \omega'_1 w_{t+h-1}). \end{aligned}$$

The (blue) expectation is easily computed using the recursive equations of Lemma 1.5 (see Example 1.20), leading to:

$$R(t, h) = -\frac{1}{h} \log B(t, h) = A'_h w_t + B_h. \quad (3.8)$$

It is easily seen that we can also get:

$$R^{EH}(t, h) = A_h^{EH'} w_t + B_h^{EH}. \quad (3.9)$$

Moreover, if inflation is also affine in w_t , i.e., if $\pi_t = \bar{\omega}_0 + \bar{\omega}_1' w_t$, then real yields are given by:

$$\bar{B}(t, h) = \mathbb{E}_t^{\mathbb{Q}} \exp(-r_t - \dots - r_{t+h-1} + \pi_{t+1} + \dots + \pi_{t+h})$$

(see Example 1.22) which also leads to:

$$\bar{R}(t, h) = -\frac{1}{h} \log \bar{B}(t, h) = \bar{A}_h' w_t + \bar{B}_h. \quad (3.10)$$

Eqs. (3.8) and (3.9) imply that term premiums are affine in w_t (see Eq. (3.5)). Specifically:

$$TP(t, h) = R(t, h) - E^{EH}(t, h) = B_h - B_h^{EH} + (A_h - A_h^{EH})' w_t.$$

Expected excess returns resulting from holding zero-coupon bonds are also affine in w_t . Indeed, holding a maturity- h zero-coupon bond for one period provides the following expected gross return:

$$\mathbb{E}_t \left(\frac{B(t+1, h-1)}{B(t, h)} \right) = \mathbb{E}_t (\exp(B_{h-1} - B_h + A_{h-1}' w_{t+1} - A_h' w_t)),$$

which is clearly exponential affine in w_t if w_t is an affine process. Therefore, the expected excess return, that is:

$$\log \mathbb{E}_t \left(\frac{B(t+1, h-1)}{B(t, h)} \right) - r_t$$

is also affine in w_t in this context. The fact that excess returns are affine in this context is exploited in the estimation approach proposed by Adrian et al. (2013).

Moreover, *conditional expectations* of future interest rates (real or nominal) and of term premiums are also affine in w_t . In particular:

$$\mathbb{E}_t[R(t+k, h)] = \mathbb{E}_t[A_h' w_{t+k} + B_h] = A_h' \mathbb{E}_t(w_{t+k}) + B_h, \quad (3.11)$$

and $\mathbb{E}_t(w_{t+k})$ is affine in w_t (see Eq. (1.15)). This can notably be used at the estimation stage, if one wants to fit survey data (see Section XXX).

Similarly, *conditional variances* of future interest rates (real or nominal) and of term premiums are affine in w_t . In particular:

$$\mathbb{V}ar_t[R(t+k, h)] = \mathbb{V}ar_t[A_h' w_{t+k} + B_h] = A_h' \mathbb{V}ar_t(w_{t+k}) A_h, \quad (3.12)$$

where the components of $\mathbb{V}ar_t(w_{t+k})$ (and therefore $\mathbb{V}ar_t[R(t+k, h)]$) is affine in w_t (see Eq. (1.16)). This can also be used at the estimation stage, if one wants to fit (proxies of) conditional variances (Monfort et al., 2017).

3.3.2 Maximum Sharpe ratio

In an affine model, the maximum Sharpe ratio is easily computed. This has been noted early by Duffee (2010) for the Gaussian model; Gouriéroux et al. (2021) and Pallara and Renne (2023) use it in more sophisticated affine models.

Let us derive the maximum Sharpe ratio in the context of a genral affine framework. Eq. (3.6) implies that

$$\underbrace{\mathbb{E}_t \left(\frac{p_{j,t+1}}{p_{j,t}} - \exp(r_t) \right)}_{=xs_{j,t+1}, \text{ excess return}} = -\exp(r_t) \text{Cov}_t \left(\mathcal{M}_{t,t+1}, \frac{p_{j,t+1}}{p_{j,t}} \right),$$

and, using $|\text{Cov}(X, Y)| \leq \sqrt{\mathbb{V}ar(X)\mathbb{V}ar(Y)}$, we get the Hansen and Jagannathan (1991) bound:

$$\underbrace{\frac{\mathbb{E}_t(xs_{j,t+1})}{\sqrt{\mathbb{V}ar_t(xs_{j,t+1})}}}_{\text{Sharpe ratio}} \leq \underbrace{\frac{\sqrt{\mathbb{V}ar_t(\mathcal{M}_{t,t+1})}}{\mathbb{E}_t(\mathcal{M}_{t,t+1})}}_{\text{Maximum Sharpe ratio}}. \quad (3.13)$$

If the S.D.F. is given by $\mathcal{M}_{t,t+1} = \exp[-r_t + \alpha'_t w_{t+1} - \psi_t(\alpha_t)]$ (Eq. (2.5)), and using that $\mathbb{E}_t(\mathcal{M}_{t,t+1}^2) = \exp(-2r_t + \psi_t(2\alpha_t) - 2\psi_t(\alpha_t))$ we get:

$$\text{Maximum Sharpe ratio} = \sqrt{\exp(\psi_t(2\alpha_t) - 2\psi_t(\alpha_t)) - 1}.$$

3.4 Gaussian Affine Term Structure Model

The Gaussian Affine Term Structure Model (GATSM) is a *Workhorse* model, widely used in academic and economic-policy circles. In a GATSM, w_t follows a Gaussian vector autoregressive model, and is therefore affine under \mathbb{P} . The S.D.F. is exponential affine in w_t , which implies that it is also affine under \mathbb{Q} (see Subsection XXX). Since the components of w_t are valued in \mathbb{R} , one can easily introduce macro-factors among the state variables.

Let us be more specific. The state vector w_t follows:

$$w_{t+1} = \mu + \Phi w_t + \Sigma^{1/2} \varepsilon_{t+1}, \text{ where } \varepsilon_t \sim i.i.d. \mathcal{N}(0, Id). \quad (3.14)$$

(The fact that we consider a VAR(1) process is without loss of generality since a VAR(p) admits a VAR(1) companion representation.)

This implies the following Laplace transform for w_t (see Example ??):

$$\psi_t(u) = \log \mathbb{E}_t(\exp(u'w_{t+1})|\underline{w}_t) = u'\mu + u'\Phi w_t + \frac{1}{2}u'\Sigma'u.$$

Using the notations of Eq. (2.5), the s.d.f. is defined as:

$$\mathcal{M}_{t,t+1} = \exp(-r_t + \alpha'_t w_{t+1} - \psi_t(\alpha_t)), \text{ where } \alpha_t = \alpha_0 + \alpha'_1 w_t.$$

In that case, using Eq. (2.7), we get:

$$\begin{aligned} \psi_t^{\mathbb{Q}}(u) &= \psi_t(u + \alpha_t) - \psi_t(\alpha_t) \\ &= (u + \alpha_t)'\mu + (u + \alpha_t)'\Phi w_t + \frac{1}{2}(u + \alpha_t)'\Sigma(u + \alpha_t) \\ &\quad - \left(\alpha'_t \mu + \alpha'_t \Phi w_t + \frac{1}{2} \alpha'_t \Sigma \alpha_t \right) \\ &= u'(\mu + \Sigma \alpha_0) + u'(\Phi + \Sigma \alpha'_1)w_t + \frac{1}{2}u'\Sigma'u. \end{aligned}$$

The \mathbb{Q} -dynamics of w_t is (from Example ??):

$$w_{t+1} = \mu + \Sigma \alpha_0 + (\Phi + \Sigma \alpha'_1)w_t + \Sigma^{1/2} \varepsilon_{t+1}^*, \text{ where } \varepsilon_t^* \sim i.i.d. \mathcal{N}^{\mathbb{Q}}(0, Id).$$

Hence, w_t also follows a VAR process under \mathbb{Q} since the previous equation rewrites:

$$w_{t+1} = \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} w_t + \Sigma^{1/2} \varepsilon_{t+1}^*,$$

where $\mu^{\mathbb{Q}} = \mu + \Sigma \alpha_0$ and $\Phi^{\mathbb{Q}} = \Phi + \Sigma \alpha'_1$.

With affine specifications of the nominal short term rate ($r_t = \omega_0 + \omega'_1 w_t$) and of the inflation rate ($\pi_t = \bar{\omega}_0 + \bar{\omega}'_1 w_t$), we obtain affine formulas for nominal and real yields of any maturity (Eqs. (3.8) and (3.10)).

Example 3.1 (Kim and Wright (2005)). This model is a three-factor *yield-only model* (no macro variables, except inflation in one variant of the model), where the short-term rate reads $r_t = \omega_0 + \omega_{1,1}w_{1,t} + \omega_{1,2}w_{2,t} + \omega_{1,3}w_{3,t}$.

The model estimated by Kalman filter (see Subsection,4.2; the state-space model (Def. 4.1) includes survey-based variables (see Subsection 4.5).



Figure 3.1: Kim and Wright (2005) outputs.

Outputs are regularly updated by the Federal Reserve Board.

Monthly data on the 6-month and 12-month-ahead forecasts of the three-month T-Bill yield from Blue Chip Financial Forecasts and semiannual data on the average expected three-month T-Bill yield from 6 to 11 years.

Example 3.2 (Ang and Piazzesi (2003)). Ang and Piazzesi (2003) propose one of the first paper mixing latent and macrovariables. The set up is also of the form Eq. (3.14), except that the VAR features several lags.¹ In their model, $w_t = [f_{1,t}^o, f_{2,t}^o, f_{1,t}^u, f_{2,t}^u, f_{3,t}^u]'$ where:

- $f_{1,t}^o$ is the first Principal Component of a set of 3 price indexes (growth rates)
- $f_{2,t}^o$ is the first Principal Component of a set of 4 real activity proxies (HELP, EMPLOY, IP, UE).
- $f_{i,t}^u$ are unobserved, or latent, factors.

The nominal short-term rate follows a Taylor rule. And latent factors are estimated via *inversion techniques* (Subsection 4.4).

¹Note that a VAR with p lags (i.e., a VAR(p)) admits a VAR(1) companion form.

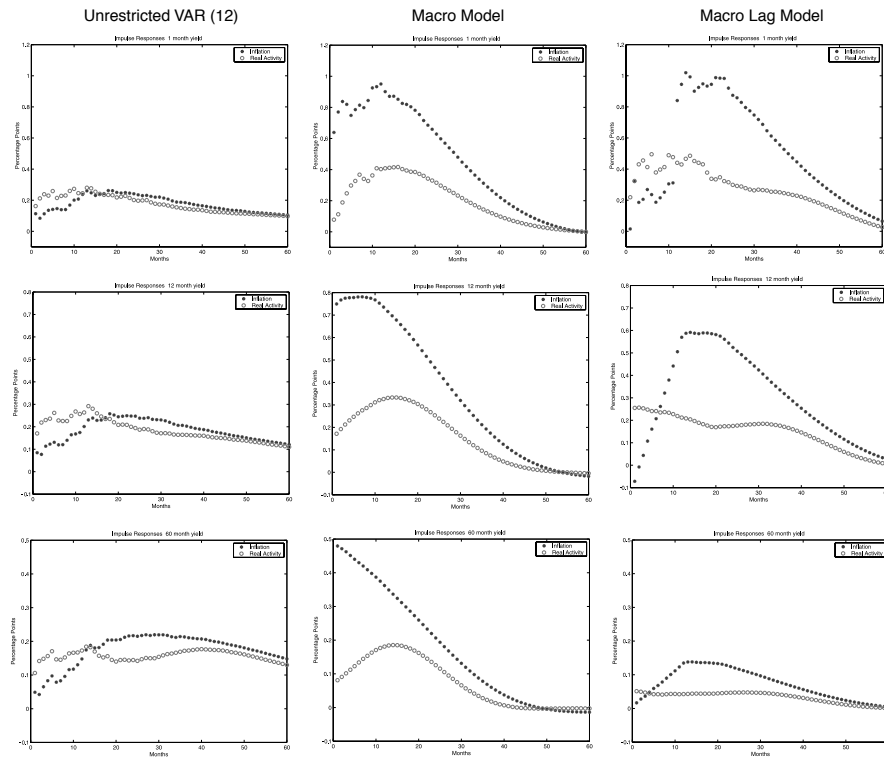


Figure 3.2: Source: Ang and Piazzesi (1998). Impulse response functions.

Example 3.3 (Joslin, Priebisch and Singleton (2014)). Joslin et al. (2014) first note that affine models stating that the short term rate is affine in macro factors imply that macro-factors are *spanned* by the yield curve: macro-factors should be perfectly explained by yields of different maturities. Further, they show that this is not the case in the data. (That is, regressing macro factors on yields provides R^2 that are far from one.)

They propose a model where macro factors are unspanned by the yield curve, but can still help predict yields. In their model, $w_t = [\mathcal{P}'_t, M'_t]'$, where \mathcal{P}_t are yield factors (\approx principal components) and M_t are macro factors. The model is as follows:

$$\begin{aligned} r_t &= \omega_0 + \omega'_\mathcal{P} \mathcal{P}_t \\ \begin{bmatrix} \mathcal{P}_t \\ M_t \end{bmatrix} &= \begin{bmatrix} \Phi_{\mathcal{P}\mathcal{P}} & \Phi_{\mathcal{P}M} \\ \Phi_{M\mathcal{P}} & \Phi_{MM} \end{bmatrix} \begin{bmatrix} \mathcal{P}_{t-1} \\ M_{t-1} \end{bmatrix} + \Sigma \varepsilon_t \\ \begin{bmatrix} \mathcal{P}_t \\ M_t \end{bmatrix} &= \mu + \begin{bmatrix} \Phi_{\mathcal{P}\mathcal{P}}^\mathbb{Q} & \textcolor{red}{0} \\ \Phi_{M\mathcal{P}}^\mathbb{Q} & \Phi_{MM}^\mathbb{Q} \end{bmatrix} \begin{bmatrix} \mathcal{P}_{t-1} \\ M_{t-1} \end{bmatrix} + \Sigma \varepsilon_t^\mathbb{Q}, \end{aligned}$$

where ε_t and $\varepsilon_t^\mathbb{Q}$ are $\mathcal{N}(0, Id)$ under \mathbb{P} and \mathbb{Q} , respectively.

M_t does Granger-cause \mathcal{P}_t under \mathbb{Q} and r_t is affine in \mathcal{P}_t (only).

In this context, yields $R(t, h)$ are affine in \mathcal{P}_t (only). However M_t does Granger-cause \mathcal{P}_t under \mathbb{P} , that is, macro-shocks affect the yield curve.

Example 3.4 (Ang, Boivin, Dong and Loo-Kung (2011)). Ang et al. (2011) propose a macro-finance model based on a quadratic framework. The short-term rate follows a Taylor rule with time-varying parameters:

$$r_t = \omega_0 + a_t g_t + b_t \pi_t,$$

where $x_t = (g_t, \pi_t, a_t, b_t)'$ follows a Gaussian VAR. This is the context described in Example 1.8. The previous equation shows that r_t is linear in $w_t = (x_t, \text{vec}(x_t x_t'))'$. Specifically:

$$r_t = \omega_0 + \omega'_1 w_t,$$

with $\omega_1 = [v, \text{vec}(V)]'$, where

$$v = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \\ 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \end{bmatrix}.$$

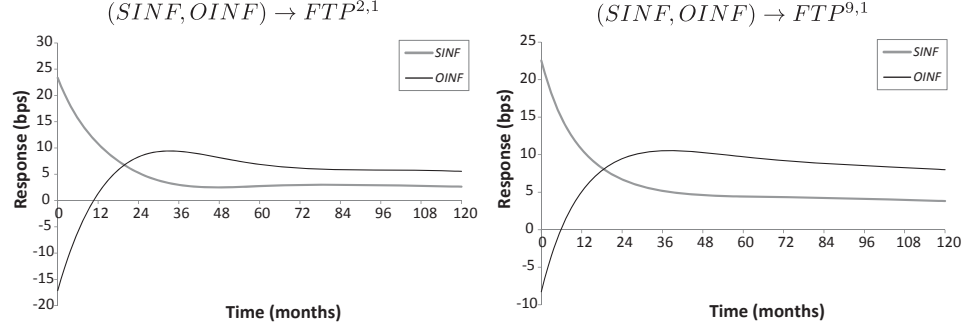


Figure 6. Response to inflation shock. Each panel plots the impulse responses of forward term premiums (FTPs) to shocks to spanned and unspanned inflation ($SINF$, $OINF$), implied by our preferred model with unspanned macro risks (\mathcal{M}_{us}). $SINF$ is the projection of INF onto the PCs; $OINF$ is the component of INF orthogonal to the PCs. $FTP^{p,1}$ is defined as the difference between the forward rate that one could lock in today for a one-year loan commencing in p years and the expectation for p years in the future of the one-year yield.

Figure 3.3: Source: Joslin, Priebsch, and Singleton (2014). Impulse response functions.

3.5 Non-Negative Affine Term Structure Model

In the presence of physical currency, absence of arbitrage opportunity and of storing cost of cash, nominal interest rates should be nonnegative. Many standard models (e.g. Gaussian ATSM) are non consistent with non-negative nominal yields. The period of extremely low interest rates challenged these models. Against this backdrop, approaches have been developed to accommodate zero (or effective) lower bounds. We provide two examples; only the second is an affine model.

3.5.1 The shadow-rate approach

The shadow-rate model is originally due to Black (1995). In this model, the short term rate is given by:

$$r_t = \max(s_t, \underline{r}), \quad (3.15)$$

where s_t is the shadow short-term interest rate and \underline{r} is the effective lower bound (≤ 0). While s_t can be real-valued, the short term rate is nonnegative

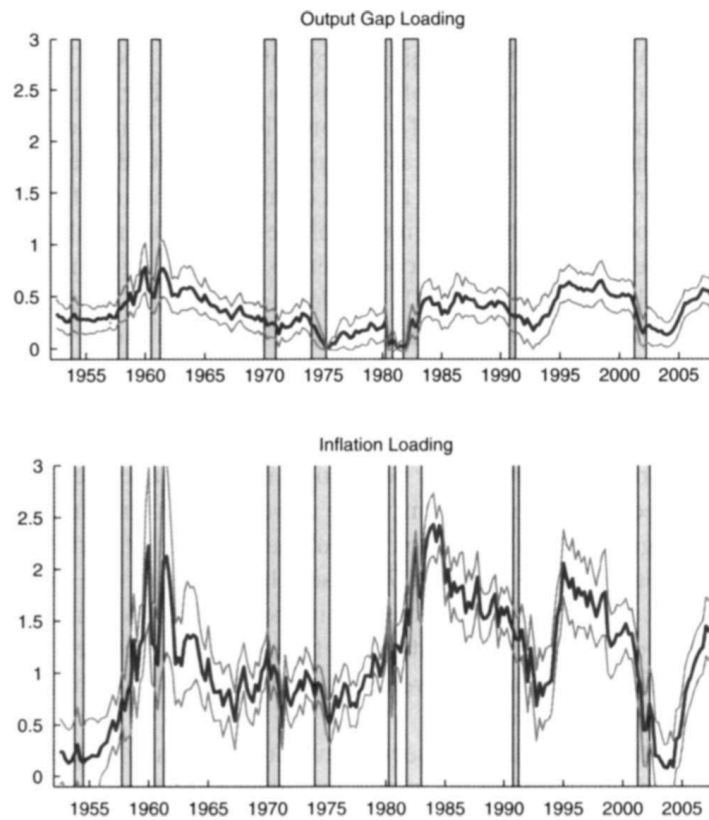


Figure 3.4: Source: Ang, Boivin, Dong, Loo-Kung (2011). Estimated factor loadings (a_t and b_t).

under (3.15). In shadow-rate models, the shadow rate s_t is usually a linear combination of a vector w_t that follows a Gaussian auto-regressive model. While s_t is a linear combination of components of an affine process, this is not the case for r_t . As a result, pricing formula are not available in closed-form. Approximation formula have been proposed by, e.g., Krippner (2013), Pribsch (2013), Wu and Xia (2016).

Let us describe the latter approach (Wu and Xia, 2016). As in Subsection 3.4, the S.D.F. is defined as:

$$\mathcal{M}_{t,t+1} = \exp(-r_t + \alpha'_t w_{t+1} - \psi_t(\alpha_t)), \text{ where } \alpha_t = \alpha_0 + \alpha'_1 w_t,$$

(this is Eq. (2.5)), but the short-term rate r_t is given by $r_t = \max(s_t, 0)$, with

$$s_t = \delta_0 + \delta'_1 w_t.$$

The approximation approach proposed by Wu and Xia (2016) is based on an approximation to the conditional expectations of forward rates. Using the results of Subsection 3.2.2, we have (Eq. (3.7)):

$$f_{n-1,n,t} = nR_{t,n} - (n-1)R_{t,n-1},$$

for $n > 0$ (and using $R_{t,0} = 0$). This implies that:

$$R_{t,h} = \frac{1}{h}(f_{t,0,1} + f_{t,1,2} + \dots + f_{t,h-1,h}).$$

The approximation of Wu and Xia (2016) consists in finding approximations of the forward rates $f_{t,n-1,n}$ (denoted by $\tilde{f}_{t,n-1,n}$, say) and to use them in the previous equation to get:

$$\boxed{R_{t,h} \approx \frac{1}{h}(\tilde{f}_{t,0,1} + \tilde{f}_{t,1,2} + \dots + \tilde{f}_{t,h-1,h}).} \quad (3.16)$$

Using that, for any random variable Z , we have $\log(\mathbb{E}[e^Z]) \approx \mathbb{E}[Z] + \frac{1}{2}\text{Var}[Z]$ (based on a second order Taylor expansion), Wu and Xia (2016) further show that:

$$f_{t,n,n+1} = -\log\left(\mathbb{E}_t^{\mathbb{Q}}\left(e^{-\sum_{j=0}^n r_{t+j}}\right)\right) + \log\left(\mathbb{E}_t^{\mathbb{Q}}\left(e^{-\sum_{j=0}^{n-1} r_{t+j}}\right)\right) \quad (3.17)$$

$$\approx \mathbb{E}_t^{\mathbb{Q}}[r_{t+n}] - \frac{1}{2}\left(\text{Var}_t^{\mathbb{Q}}\left(\sum_{j=0}^n r_{t+j}\right) - \text{Var}_t^{\mathbb{Q}}\left(\sum_{j=0}^{n-1} r_{t+j}\right)\right) \quad (3.18)$$

The expectation can be computed analytically:

$$\mathbb{E}_t^{\mathbb{Q}}[r_{t+n}] = \underline{r} + \sigma_n^{\mathbb{Q}} g\left(\frac{\bar{a}_n + b'_n X_t - \underline{r}}{\sigma_n^{\mathbb{Q}}}\right),$$

where $g(x) = x\Phi(x) - \phi(x)$, Φ and ϕ being the c.d.f. and p.d.f. of the standard normal distribution, respectively, and where

$$\begin{aligned}\bar{a}_n &= \delta_0 + \delta'_1 \left(\sum_{j=0}^{n-1} [\Phi^{\mathbb{Q}}]^j \right) \mu^{\mathbb{Q}} \\ b'_n &= \delta'_1 (\Phi^{\mathbb{Q}})^n.\end{aligned}$$

They also show that

$$\frac{1}{2} \left(\mathbb{V}ar_t^{\mathbb{Q}} \left(\sum_{j=0}^n r_{t+j} \right) - \mathbb{V}ar_t^{\mathbb{Q}} \left(\sum_{j=0}^{n-1} r_{t+j} \right) \right) \approx \Phi \left(\frac{\bar{a}_n + b'_n X_t - \underline{r}}{\sigma_n^{\mathbb{Q}}} \right) \times (\bar{a}_n - a_n),$$

where

$$a_n = \bar{a}_n - \frac{1}{2} \sigma_n^{\mathbb{Q}},$$

with

$$\sigma_n^{\mathbb{Q}} := \mathbb{V}ar_t^{\mathbb{Q}}(s_{t+n}) = \delta'_1 \left(\sum_{j=0}^{n-1} [\Phi^{\mathbb{Q}}]^j \right) \Sigma \Sigma' \left(\sum_{j=0}^{n-1} [\Phi^{\mathbb{Q}}]^j \right)' \delta_1.$$

They finally obtain:

$$\boxed{f_{t,n,n+1} \approx \tilde{f}_{t,n,n+1} = \underline{r} + \sigma_n^{\mathbb{Q}} g\left(\frac{a_n + b'_n X_t - \underline{r}}{\sigma_n^{\mathbb{Q}}}\right),}$$

which is used in (3.16) to obtain an approximation to $R_{t,h}$.

```
library(TSModels)
# Specify model:
n <- 2 # number of factors
rho <- matrix(0,n,n)
diag(rho) <- .97
mu <- matrix(0,n,1)
```

```

Sigma <- diag(n)
delta.0 <- 0; delta.1 <- rep(.01,n)
r.bar <- 0 # r = max(s,r.bar) [i.e., r.bar=0 in standard model]
Model <- list(rho = rho,mu = mu,Sigma = Sigma,
  delta.0 = delta.0,delta.1 = delta.1,r.bar = r.bar)
# Simulate model and compute shadow rate:
X <- simul.var(Model,nb.sim = 200) # simulated path
s <- delta.0 + X %*% delta.1
# Compute yields:
res <- compute.price.WX(Model,X,max.H=100)
# Prepare plots:
par(plt=c(.1,.95,.2,.75))
par(mfrow=c(2,1))
plot(s,type="l",xlab="time",ylab="",lwd=2,main="(a) Shadow rate")
t <- 50 #t <- which(s==min(s))
abline(v=t,col="dark grey",lwd=2,lty=3)
plot(res$vec.f[t,],type="l",xlab="maturity",ylab="",
  lwd=2,main="(b) yields and forward rates")
lines(res$vec.y[t,],col="red",lwd=2)
legend("topright",
  c("forward rates","yields to maturity"),lwd=c(2),lty=1,
  col=c("black","red"),bg = "white")

```

3.5.2 The auto-regressive gamma approach

Monfort et al. (2017) introduce an affine framework where the short-term rate can stay at zero for a prolonged period of time and with a stochastic lift-off probability.

Under \mathbb{P} and \mathbb{Q} , the state vector w_t follows a multi-variate auto-regressive gamma (VARG) process—a multivariate extension of Example 1.9. Conditionally on \underline{w}_t , the n components of w_{t+1} are independent and distributed as follows:

$$\frac{w_{i,t+1}}{\mu_i} \sim \gamma(\nu_i + z_{i,t}) \quad \text{where} \quad z_{i,t} \sim \mathcal{P}(\alpha_i + \beta_i' w_t). \quad (3.19)$$

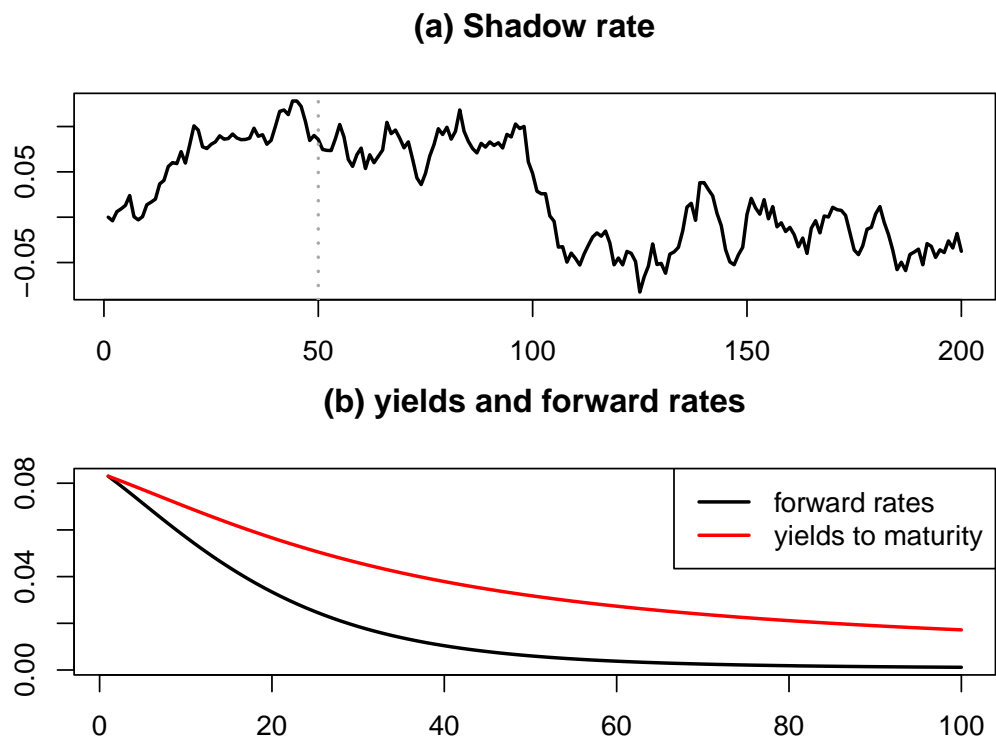


Figure 3.5: bla bla bla.

If $\mu = (\mu_1, \dots, \mu_n)'$, $\alpha = (\alpha_1, \dots, \alpha_n)'$, $\nu = (\nu_1, \dots, \nu_n)'$ and $\beta = (\beta_1, \dots, \beta_n)$, then

$$\begin{aligned} \varphi_t(u) = & \exp \left[\left(\frac{u \odot \mu}{1 - u \odot \mu} \right)' \beta' w_t \right. \\ & \left. + \alpha' \left(\frac{u \odot \mu}{1 - u \odot \mu} \right) - \nu' \log(1 - u \odot \mu) \right], \end{aligned}$$

where \odot denotes the element-by-element multiplication and, where, with abuse of notation, the division and log operators work element-by-element when applied to vectors.

In their baseline model, Monfort et al. (2017) use four factors. They set $\nu_1 = \nu_2 = 0$, implying that $w_{1,t}$ and $w_{2,t}$ can stay at zero (see Example 1.9). The short-term rate r_t is posited to be an affine combination of $w_{1,t}$ and $w_{2,t}$, that is:

$$r_t = \omega' w_t = \omega_1 w_{1,t} + \omega_2 w_{2,t},$$

hence, it can stay at zero.

Factors $w_{3,t}$ and $w_{4,t}$ Granger-cause $w_{1,t}$ and $w_{2,t}$, thereby causing r_t . As a result, for $h \geq 2$, $R(t, h)$ is a non-zero combination of the four components of w_t .

For the same reason, when $r_t = 0$, the lift-off probability depends on $w_{3,t}$ and $w_{4,t}$. The framework offers closed-form solutions for lift-off probabilities. Indeed, using Lemma ??:

$$\mathbb{P}_t(\alpha' w_{t+h} = 0) = \lim_{u \rightarrow -\infty} \varphi_{t,h}(0, \dots, 0, u\alpha),$$

where $\varphi_{t,h}$ is the multi-horizon Laplace transform defined in Eq. (1.7), which can be computed using Proposition 1.5. We have:

$$\begin{cases} \mathbb{P}_t(r_{t+h} > 0) = 1 - \lim_{u \rightarrow -\infty} \varphi_{t,h}(0, \dots, 0, u\omega) \\ \mathbb{P}_t(r_{t+1} = 0, \dots, r_{t+h} = 0) = \lim_{u \rightarrow -\infty} \varphi_{t,h}(u\omega, \dots, u\omega, u\omega) \equiv p_h \\ \mathbb{P}_t(r_{t+1} = 0, \dots, r_{t+h-1} = 0, r_{t+h} > 0) = p_{h-1} - p_h. \end{cases} \quad (3.20)$$

Other lift-off probabilities, of the type $\mathbb{P}_t[R(t+h, k) > \text{threshold}]$, can be derived from Eq. (1.18).

Monfort et al. (2017) estimate this model by means of Kalman filtering techniques (see Subsection ??). Observed variables include (levels of) yields, as well as survey-based forecasts of yields (see Subsection 4.5 and (e-GARCH-based) proxies of conditional variances (see Eq. (1.16)).

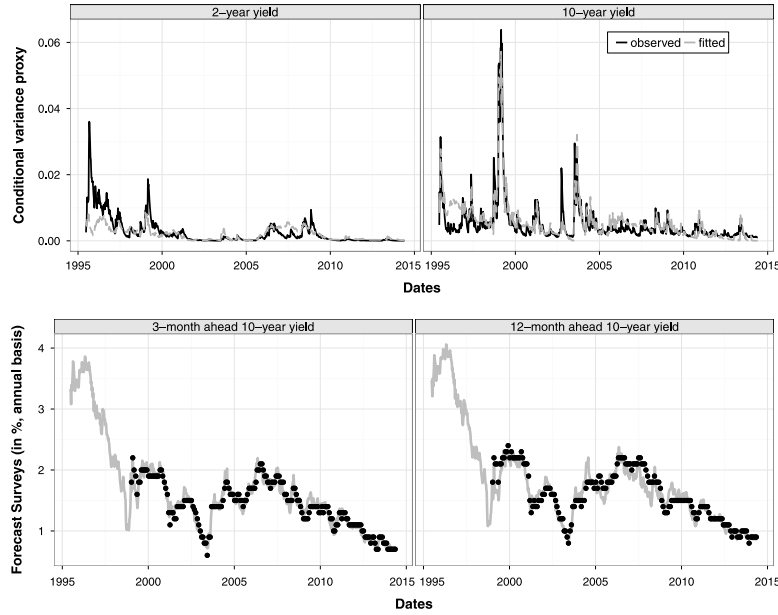


Fig. 7. Fitted conditional variance proxies and surveys. *Notes:* The top panel presents the two conditional variance proxies $V_t(h)$ estimated with an EGARCH(1,1) model on 2- and 10-year yields (left and right tiles) of weekly data from June 30, 1995 to May 30, 2014. The black solid lines are the observed variance proxies and the grey dotted lines are the model-implied (or fitted) equivalent. The bottom panel presents the survey of professional forecasters for the 10-year yield, 3- and 12-months ahead. Survey-based data are available at the monthly frequency from 1999 to the end of the sample. The black dots correspond to the observed data, and the grey solid lines are the fitted equivalent.

Figure 3.6: Source: Monfort et al. (2017). Model fit of conditional variances and surveys of professional forecasters.

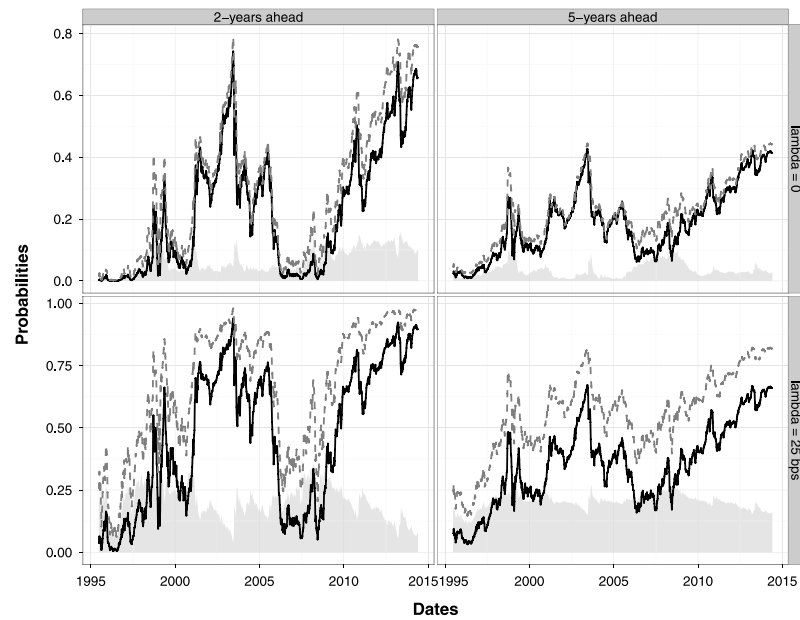


Fig. 9. Time-series of ZLB probabilities: $\mathbb{P}_t(r_{t+k} \leq \lambda)$ and $Q_t(r_{t+k} \leq \lambda)$. Notes: Probabilities are computed with weekly data from June 16, 1995 to May 30, 2014. The top panels present the probabilities of the short-rate hitting zero in two years (top-left panel) and 5 years (top-right panel). On bottom panels, we represent the probabilities of the short-rate being below 25 bps in 2 years (bottom-left panel) and 5 years (bottom-right panel). Black solid lines are the risk-neutral probabilities whereas grey dashed lines are the historical ones; grey-shaded areas are the difference between the two probabilities.

Figure 3.7: Source: Monfort et al. (2017). Lift-off probabilities.

Chapter 4

Estimation of affine asset-pricing models

4.1 State-Space Model

By nature, dynamic asset-pricing models are *state-space models*: The dynamics of all variables, gathered in vector y_t (yields, equity returns, macroeconomic variables, survey-based variables) are accounted for by state variables (w_t). The equations defining the relationship between the other variables and the state variables are called *measurement equations* (Eq. (4.1)). The equations defining the dynamics of the state variables are called *transition equations* (Eq. (4.2)).

In the case where w_t is an affine process (see Definition 1.1), the transition equations admit a VAR(1) representation (Proposition 1.1). In that case, the state-space model is said to be linear, as formalized by the following definition. This definition introduces, in particular, the notion of *measurement errors*.

Definition 4.1 (Linear State-Space Model). A linear state-space model writes as follows:

$$\underset{(m \times 1)}{y_t} = A + Bw_t + \eta_t \quad \text{with } \eta_t \sim i.i.d.\mathcal{N}(0, \Omega) \quad (4.1)$$

$$\underset{(n \times 1)}{w_t} = \mu + \Phi w_{t-1} + \Sigma^{\frac{1}{2}}(w_{t-1})\varepsilon_t, \quad (4.2)$$

where ε_t is a martingale difference sequence satisfying $\text{Var}_t(\varepsilon_{t+1}) = Id$. The components of η_t are measurement errors.

Note: Eq. (4.2) derives from Proposition 1.1 (Eq. (1.12)).

In practice, one can distinguish two situations: (a) all state variables (components of w_t) are observed and (b) some of these variables are latent. What do we mean by *model estimation* in case (b)? There are different situations:

- (b.i) We know the model parameters but we want to recover the latent factors—for instance to compute model-implied prices.
- (b.ii) We know neither the model parameters nor the latent variables, we want to estimate both of them.
- (b.iii) We know neither the model parameters nor the latent variables, we are just interested in the model parameters.

In case (a): One can resort to standard estimation techniques (GMM, Maximum Likelihood) to estimate model parameters. Take for instance the maximum-likelihood case, we have:

$$\begin{aligned} f(y_t, w_t | \underline{y_{t-1}}, \underline{w_{t-1}}) &= f(y_t | \underline{y_{t-1}}, \underline{w_t}) \times \underbrace{f(w_t | \underline{y_{t-1}}, \underline{w_{t-1}})}_{=f(w_t | \underline{w_{t-1}})} \\ &= \mathbf{n}(y_t; A + Bw_t, \Omega) f(w_t | \underline{w_{t-1}}), \end{aligned}$$

where $\mathbf{n}(x; \mu, \Omega)$ denotes the evaluation, at vector x , of the p.d.f. of the multivariate normal distribution $\mathcal{N}(\mu, \Omega)$. That is, if x is a m -dimensional vector:

$$\mathbf{n}(x; \mu, \Omega) = \frac{1}{\sqrt{(2\pi)^m |\Omega|}} \exp \left(-\frac{1}{2} \{x - \mu\}' \Omega^{-1} \{x - \mu\} \right). \quad (4.3)$$

Once this conditional p.d.f. is known, the total likelihood is given by (conditional on y_0 and w_0):

$$\prod_{t=1}^T f(y_t, w_t | \underline{y_{t-1}}, \underline{w_{t-1}}).$$

Of course, $f(w_t|\underline{w_{t-1}})$ depends on the process chosen for w_t . If it is complicated to compute, one can employ Pseudo Maximum Likelihood (Gourieroux et al., 1984). The p.d.f. may involve, for instance, an infinite sum, which is the case in the ARG case of Example 1.9. When w_t is affine, the Pseudo Maximum Likelihood approach consists in replacing $f(w_t|\underline{w_{t-1}})$ by:

$$\mathbf{n}(w_t; \mu + \Phi w_{t-1}, \Sigma(w_t)),$$

where μ , Φ and $\Sigma(w_t)$ are introduced in Eqs. (1.13) and (1.14) in Proposition 1.1.

In case (b.iii), one can estimate the model by Generalized Method of Moments (GMM), fitting sample moments computed using observed variables (prices, yields, returns). In the context of affine processes, conditional and unconditional moments of the state vector w_t are available in closed form, as shown by Eqs. (1.15), (1.16) and (1.17). If the model-implied moments are not available in closed-form, one may have to resort to the *Simulated Method of Moments (SMM)* (see, e.g., Gouriéroux and Monfort (1997) or Duffie and Singleton (1993)).

In cases (b.i) and (b.ii), one has to implement *filtering methods*, on which we focus on in the following.

4.2 Kalman-Filter-Based Approach

4.2.1 The Gaussian linear state-space case

Let us start with a particular case of state-space model (Def. 4.1) where ε_t is Gaussian and where $\Sigma^{\frac{1}{2}}$ does not depend on w_t , i.e. with a homoskedastic linear Gaussian state-space model.

Let's denote by θ the vector of parameters that defines the model. For a given θ and a sequence of observations $\{y_1, \dots, y_T\}$, the Kalman filter computes the distribution of w_t given $\{y_1, \dots, y_t\}$ (see Def. 4.2). This distribution is Gaussian, and obtained by a recursive algorithm. A byproduct of this algorithm is the likelihood function associated with θ and $\{y_1, \dots, y_T\}$. This opens the door to the estimation of θ by MLE, maximizing this function. In this sense, Kalman-filter techniques can address Objective (b.ii).

Let us first introduce the notion of *filtered* and *smoothed* estimates of a latent variable (or vector of variables) w_t :

Definition 4.2 (Filtered versus smoothed estimates). The filtering and smoothing problems consist in computing the following conditional moments:

$$\begin{aligned} \text{Filtering: } w_{t|t} &= \mathbb{E}(w_t | \underline{y}_t) \quad \text{and} \quad P_{t|t} = \mathbb{V}ar(w_t | \underline{y}_t) \\ \text{Smoothing: } w_{t|T} &= \mathbb{E}(w_t | \underline{y}_T) \quad \text{and} \quad P_{t|T} = \mathbb{V}ar(w_t | \underline{y}_T). \end{aligned}$$

The following proposition outlines the Kalman algorithm (see, e.g. Nelson and Kim (1999)).

Proposition 4.1 (Kalman filter and smoother). *If $\varepsilon_t \sim \mathcal{N}(0, I)$ in the state-space defined in Def. 4.1, then we have (filtering):*

$$w_t | y_1, \dots, y_t \sim \mathcal{N}(w_{t|t} | P_{t|t}),$$

where $w_{t|t}$ and $P_{t|t}$ result from the following recursive equations:

$\begin{aligned} w_{t t} &= w_{t t-1} + K_t \lambda_t \\ P_{t t} &= (I - K_t B) P_{t t-1} \end{aligned}$
<p>where (updating step)</p>
$\begin{aligned} \lambda_t &= y_t - A - B w_{t t-1} \quad (\text{forecast error}) \\ S_{t t-1} &= \mathbb{V}ar(y_t \underline{y}_{t-1}) = B P_{t t-1} B' + \Omega \\ K_t &= P_{t t-1} B' S_{t t-1}^{-1} \quad (\text{Kalman gain}) \end{aligned}$
<p>and where (forecasting step)</p>
$\begin{aligned} w_{t t-1} &= \mu + \Phi w_{t-1 t-1} \\ P_{t t-1} &= \Sigma + \Phi P_{t-1 t-1} \Phi' \quad (\Sigma = \Sigma^{\frac{1}{2}} \Sigma^{\frac{1}{2}}'). \end{aligned}$

The log likelihood is (recursively) computed as follows:

$$\begin{aligned} \log \mathcal{L}(\theta; \underline{y}_T) &= \frac{mT}{2} \log(2\pi) \\ &\quad - \frac{1}{2} \sum_{t=1}^T \left(\log |S_{t|t-1}(\theta)| + \lambda_t'(\theta) S_{t|t-1}^{-1}(\theta) \lambda_t(\theta) \right). \end{aligned} \tag{4.4}$$

Moreover, we have (smoothing):

$$w_t|y_1, \dots, y_T \sim \mathcal{N}(w_{t|T}|P_{t|T}),$$

where $w_{t|T}$ and $P_{t|T}$ result from the following recursive equations:

$$\begin{aligned} w_{t|T} &= w_{t|t} + F_t(w_{t+1|T} - w_{t+1|t}) \\ P_{t|T} &= P_{t|t} + F_t(P_{t+1|T} - P_{t+1|t})F_t' \\ \text{where} \\ F_t &= P_{t|t}\Phi_{t+1}'P_{t+1|t}^{-1}. \end{aligned}$$

The following figure illustrates the updating step of the Kalman algorithm:

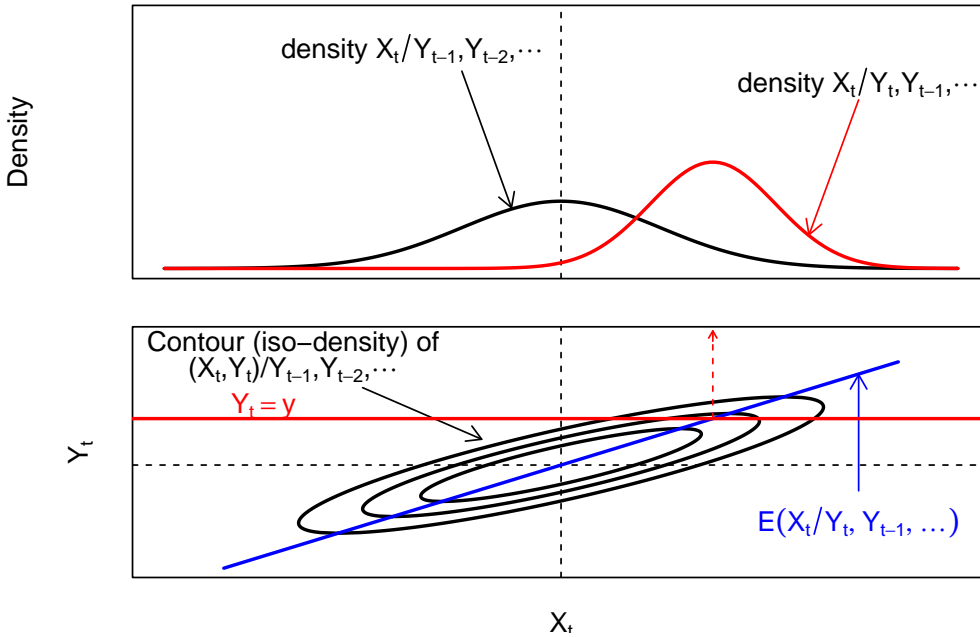


Figure 4.1: Updating in the Kalman filter.

The recursive equations of the Kalman filter need to be initialized. That is, one needs to provide initial values for $w_{0|0}$, $P_{0|0}$. Different possibilities have been proposed. One can for instance:

- Include the elements of $(w_{0|0}, P_{0|0})$ among the parameters to estimate;

- Set $w_{0|0}$ and $P_{0|0}$ to their unconditional values (using, e.g., Eq. (1.17));
- Set $w_{0|0}$ to a prior value and take either an arbitrary large value for $P_{0|0}$ if the prior value is uncertain (which depicts a situation of *diffuse prior*) or a small value for $P_{0|0}$ if we are confident in this prior value $w_{0|0}$.

4.2.2 Missing observations

In many application, one does not observe all the entries of y_t at every date. This arises for instance in situations where (i) measurement variables feature different frequencies, (ii) we have unreliable data for some period (that we prefer not to include among observations), (iii) some of the measurement variables are observed over a shorter time span.

These situations are easily addressed by Kalman filtering/smoothing (e.g., Chow and Lin (1971) or Harvey and Pierse (1984)). To accommodate missing observations in some of the y_t 's, one simply has to change the size of this vector (and of λ_t , $S_{t|t-1}$, A and B) for the relevant dates. Of course, the accuracy of $w_{t|t}$ will tend to be lower during periods where one or several (or all) the entries of y_t are unobserved. (This will be apparent in the resulting covariance matrix of $w_{t|t}$, namely $P_{t|t}$.) The log-likelihood computation (Eq. (4.4)) is still valid in this case; one simply has to adjust the number of observed variables at each iteration; that is, m then depends on time.

To illustrate, consider the following model:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-2} \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} w_t + D\eta_t \quad (4.5)$$

$$w_t = \phi w_{t-1} + \varepsilon_t, \quad (4.6)$$

where $\eta_t \sim \mathcal{N}(0, Id)$.

We simulate a 100-period sample, we remove observations of $y_{1,t}$ (respectively of $y_{2,t}$) between periods $t = 30$ and $t = 50$ (resp. between periods $t = 40$ and $t = 70$), and try to recover the states w_t .

```
library(TSModels) # Kalman filter procedure in there.
# Set model specifications:
alpha1 <- .5; alpha2 <- .95; Alpha <- diag(c(alpha1, alpha2))
```



```

d_11 <- 1;d_12 <- .5;d_21 <- .5;d_22 <- 2
D <- matrix(c(d_11,d_21,d_12,d_22),2,2)
gamma1 <- 1;gamma2 <- 2;Gamma <- matrix(c(gamma1,gamma2),2,1)
phi <- .8
# Simulate model:
T <- 100
Y <- NULL;X <- NULL
Alpha.Y_1 <- NULL
y <- c(0,0);x <- 0
for(i in 1:T){
  Alpha.Y_1 <- rbind(Alpha.Y_1,c(Alpha %%% y))
  y <- Alpha %%% y + Gamma * x + D %%% rnorm(2)
  x <- phi * x + rnorm(1)
  Y <- rbind(Y,t(y));X <- rbind(X,x)}
# Define matrices needed in the Kalman_filter procedures:
nu_t <- matrix(0,T,1)
H <- phi;G <- Gamma
mu_t <- Alpha.Y_1
N <- 1;M <- D
Sigma_0 <- 1/(1-phi^2) # unconditional variance of w
rho_0 <- 0 # unconditional mean of w
filter.res <- Kalman_filter(Y,nu_t,H,N,mu_t,G,M,Sigma_0,rho_0,indic_pos=0)
smoother.res <- Kalman_smoother(Y,nu_t,H,N,mu_t,G,M,Sigma_0,rho_0,indic_pos=0)

```

```

par(mfrow=c(2,1))
par(plt=c(.1,.95,.15,.8))

plot(X,type="l",main=expression(X[t]),lwd=2,xlab="",ylab="")
lines(filter.res$r,col="red",lwd=2)
lines(smoother.res$r,col="blue",lwd=2,lty=1)

legend("topright",
      c(expression(X[t]),
        expression(paste("Filtered values of ",X[t],sep="")),
        expression(paste("Smoothed values of ",X[t],sep=""))),
      lwd=c(2), # line width

```

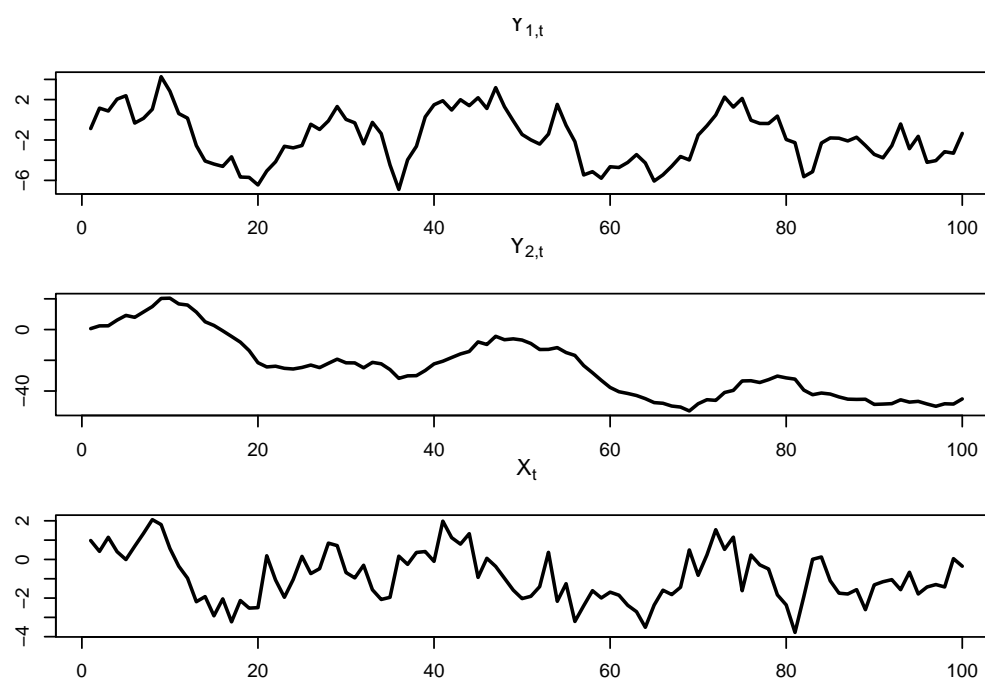


Figure 4.2: XXXX.

```

        lty=c(1,1,1),
        col=c("black","red","blue"),
        seg.len = 4,
        bg = "white"
    )

plot(sqrt(filter.res$Sigma_tt),type="l",main=expression(X[t]),
     lwd=2,xlab="",ylab="",col="red",ylim=c(min(sqrt(smoother.res$S_smooth)),
                                             max(sqrt(filter.res$Sigma_tt))))
lines(sqrt(smoother.res$S_smooth),lwd=2,col="blue")

legend("topright",
      c(expression(paste("Filtering standard error ",sqrt(Var(X[t*|"*t]))),sep="")),
        expression(paste("Smoothing standard error ",sqrt(Var(X[t*|"*T]))),sep="")),
      lwd=c(2), # line width
      lty=c(1,1,3),
      col=c("black","red","red"),
      seg.len = 4,
      bg = "white"
)

# Missing data issue: remove observations.

Y.modif <- Y
Y.modif[30:50,1] <- NaN
Y.modif[40:70,2] <- NaN

filter.res <- Kalman_filter(Y.modif,nu_t,H,N,mu_t,G,M,Sigma_0,rho_0,indic_pos=0)
smoother.res <- Kalman_smoother(Y.modif,nu_t,H,N,mu_t,G,M,Sigma_0,rho_0,indic_pos=0)

par(mfrow=c(3,1))
par(plt=c(.1,.95,.15,.85))

plot(Y.modif[,1],type="l",main=expression(Y[1*|"*t]),lwd=2,xlab="",ylab="")
plot(Y.modif[,2],type="l",main=expression(Y[2*|"*t]),lwd=2,xlab="",ylab="")
plot(X,type="l",main=expression(X[t]),lwd=2,xlab="",ylab="")

```

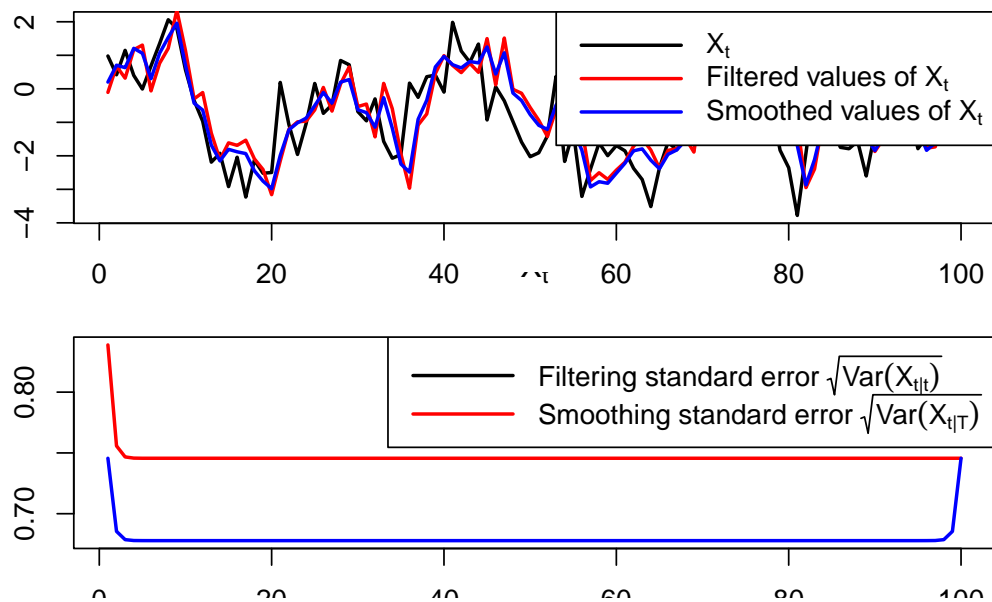


Figure 4.3: XXXX.

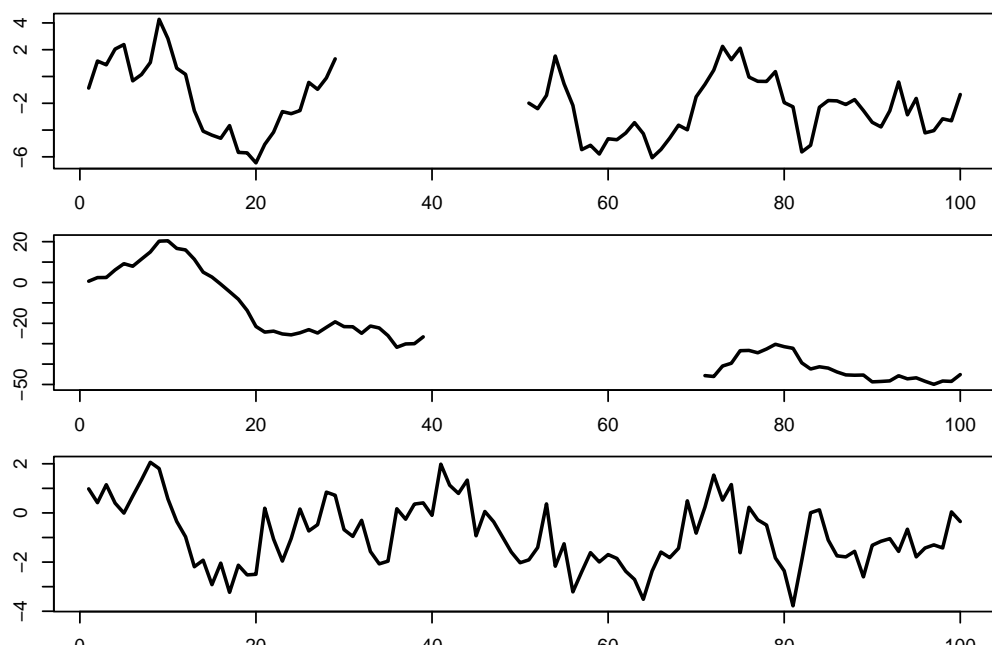


Figure 4.4: XXXX.

```

par(mfrow=c(3,1))
par(plt=c(.15,.95,.3,.95))

par(mfrow=c(3,1))
plot(Y.modif[,1],type="l",main=expression(y[1*","*t]),lwd=2,xlab="",ylab=expression(y
plot(Y.modif[,2],type="l",main=expression(y[2*","*t]),lwd=2,xlab="",ylab=expression(y
plot(X,type="l",main=expression(w[t]),lwd=2,xlab="",ylab=expression(w[t]))

```

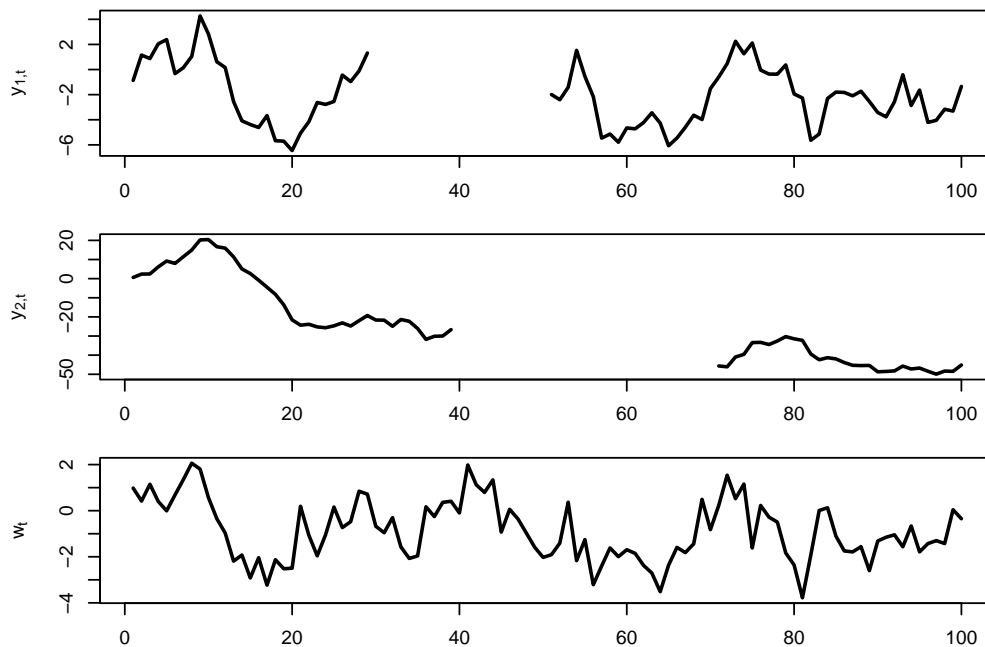


Figure 4.5: XXXX.

```

par(mfrow=c(2,1))
par(plt=c(.1,.95,.15,.8))

plot(X,type="l",main=expression(X[t]),lwd=2,xlab="",ylab="")
lines(filter.res$r,col="red",lwd=2)
lines(smooth.res$r,col="blue",lwd=2,lty=1)

legend("topright",

```

```

        c(expression(X[t]),
          expression(paste("Filtered values of ",X[t],sep="")),
          expression(paste("Smoothed values of ",X[t],sep=""))),
        lwd=c(2), # line width
        lty=c(1,1,1),
        col=c("black","red","blue"),
        seg.len = 4,
        bg = "white"
      )

plot(sqrt(filter.res$Sigma_tt),type="l",main=expression(X[t]),lwd=2,xlab="",y
      ylim=c(min(sqrt(smoother.res$S_smooth)),max(sqrt(filter.res$Sigma_tt)))
lines(sqrt(smoother.res$S_smooth),lwd=2,col="blue")

legend("topright",
      c(expression(paste("Filtering standard error ",sqrt(Var(X[t*|"*t])),s
        expression(paste("Smoothing standard error ",sqrt(Var(X[t*|"*T])),s
        lwd=c(2), # line width
        lty=c(1,1,3),
        col=c("black","red","red"),
        seg.len = 4,
        bg = "white"
      )

```

```

par(mfrow=c(2,1))
par(plt=c(.1,.95,.15,.8))

plot(X,type="l",main=expression(w[t]),lwd=2,xlab="",ylab="")
lines(filter.res$r,col="red",lwd=2)
lines(smoother.res$r,col="blue",lwd=2,lty=1)

legend("topright",
      c(expression(w[t]),
        expression(paste("Filtered values of ",w[t],sep="")),
        expression(paste("Smoothed values of ",w[t],sep=""))),
        lwd=c(2), # line width

```

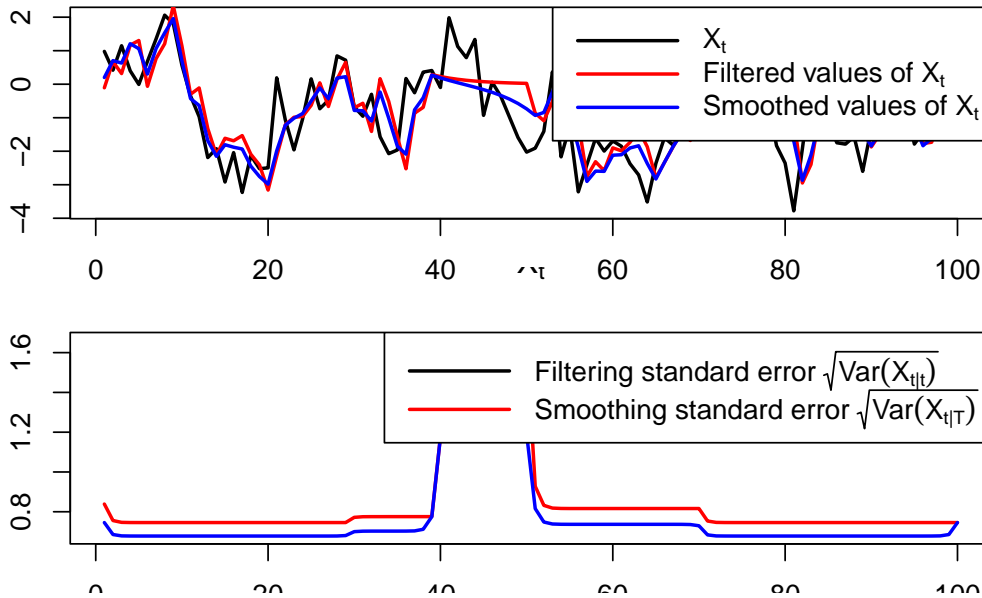


Figure 4.6: XXXX.

```

lty=c(1,1,1),
col=c("black","red","blue"), # gives the legend lines the correct color and w
seg.len = 4,
bg = "white"
)

plot(sqrt(filter.res$Sigma_tt),type="l",
     main = "Conditional standard deviation of filtered/smoothed estimates",
     lwd=2,xlab="",ylab="",col="red",
     ylim=c(min(sqrt(smooth.res$S_smooth)),1.4*max(sqrt(filter.res$Sigma_tt))))
lines(sqrt(smooth.res$S_smooth),lwd=2,col="blue")

legend("topleft",
      c(expression(paste("Filtering standard error ",sqrt(Var(w[t*"|"*t])),sep="")),
        expression(paste("Smoothing standard error ",sqrt(Var(w[t*"|"*T])),sep="))),
      lwd=c(2), # line width
      lty=c(1,1,3),

```

```
col=c("red","blue"),
seg.len = 4,
bg = "white"
)
```

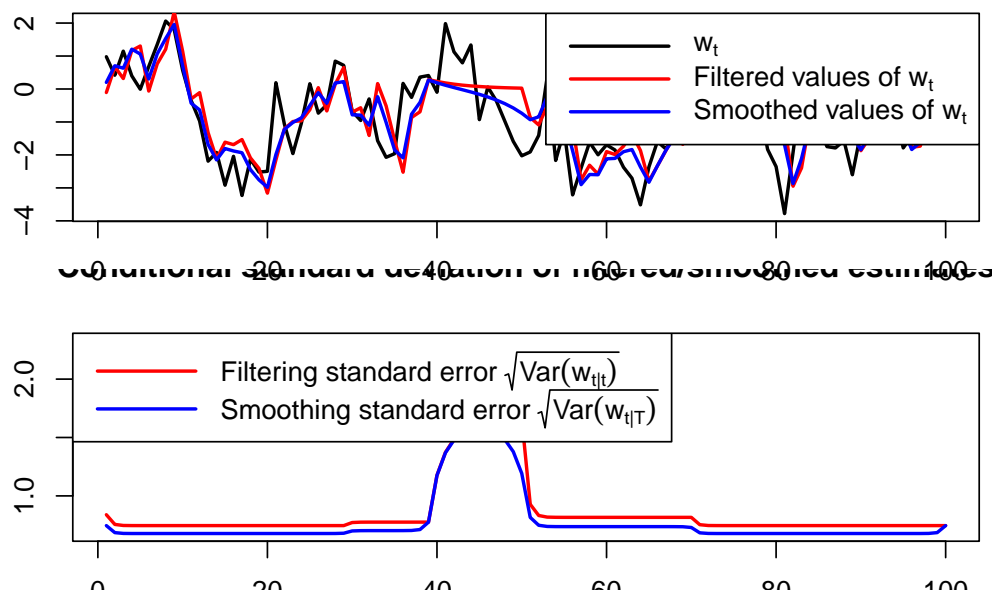


Figure 4.7: XXXX.

4.3 About non-constant conditional matrix Σ

- Proposition 4.1 is valid when ε_t is Gaussian and when $\Sigma_t^{\frac{1}{2}}$ does not depend on w_t .
- How to proceed in the general case?
- In the general case (but when w_t is an affine process), we have that $\Sigma(w_{t-1}) \equiv \text{Var}(w_{t+1}|w_t)$ is affine in w_t [Prop. 1.1].
- In order to deal with this, the Kalman filter algorithm can be modified.
- Specifically, in the prediction step (see Prop. 4.1), $P_{t|t-1}$ can be computed as:

$$P_{t|t-1} = \Sigma(w_{t-1|t-1}) + \Phi P_{t-1|t-1} \Phi',$$

i.e. we replace $\Sigma(w_{t-1})$ by $\Sigma(w_{t-1|t-1})$.

- Though the approach is then not necessarily optimal, it shows good empirical properties de Jong (2000) or Duan and Simonato (1999).
- In order to test for the validity of the approach in a specific context, one can resort to Monte-Carlo simulations [see e.g. Monfort, Pegoraro, Renne and Roussellet (2017)].

4.3.1 Non-linear models

- If w_t follows an affine process, it admits the VAR dynamics presented in Prop. 1.1 \Rightarrow linear transition equation.
- However, measurement equations may be non-linear (affine) functions of w_t .
- This is in particular the case if observed variables include Swaps rates (see Remark SWAPS), CDS rates (see Subsection ??, in particular Eq. (??)) or prices of tranche products (see Example ??).
- In this case, one can for instance resort to the Extended Kalman Filter (linearizing the measurement equations, see supplementary material) or the Quadratic Kalman Filter Monfort, Renne and Roussellet, 2015, second-order expansion of the measurement equations).

4.4 Inversion Technique

- The **inversion technique** has been introduced by Chen and Scott (1993).
- It is used e.g. by Ang and Piazzesi (2003) and Liu, Longstaff and Mandell (2006).
- Contrary to Kalman-type approaches, this approach is not recursive. Hence can be faster (especially for long sample – in time dimension).
- Recall that y_t and w_t are respectively of dimension m and n (see Eqs. (4.1) and (4.2) in Def. 4.1).

Hypothesis 4.1 (Perfectly-modelled variables).

- Assumption: n components of the m -dimensional vector y_t (with $n \leq m$) are perfectly modelled (no measurement errors in associated measurement equations).

- Without loss of generality, these perfectly-modelled variables are the first n components of y_t , that is:

$$y_t = \begin{pmatrix} \underbrace{y_{1,t}}_{(n \times 1)} \\ \underbrace{y_{2,t}}_{(m-n) \times 1} \end{pmatrix}.$$

- Under Assumption 4.1, the measurement equation (Eq. (4.1)) becomes:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} w_t + \begin{bmatrix} 0 \\ \eta_{2,t} \end{bmatrix},$$

where $\eta_{2,t} \sim \mathcal{N}(0, \Omega_2)$ (say). This notably implies

$$w_t = B_1^{-1}(y_{1,t} - A_1). \quad (4.7)$$

- Under this assumption and if the conditional distribution of w_t is available in closed form, then the (exact) likelihood of the model can then be computed.
- Proposition 4.2 (next slide) shows that the computation of $f_{Y_t|Y_{t-1}}(y_t; y_{t-1})$ involves three terms:
- The first term (in blue in (4.8)) stems from the conditional distribution $w_t|w_{t-1}$.
- The second term (in red in (4.8)) is associated with the measurement errors pertaining to $y_{2,t}$, that are the components of $\eta_{2,t}$.
- The third term (in brown in (4.8)) is the determinant of the Jacobian matrix associated with the linear transformation between w_t and $y_{1,t}$ (Eq. (4.7)), that is $|B_1|$.
- Once one knows how to compute $f_{Y_t|Y_{t-1}}(y_t; y_{t-1})$, the total likelihood is easily obtained since:

$$f_{Y_1, \dots, Y_T}(y_1, y_2, \dots, y_T) = f_{Y_1}(y_1) \prod_{t=2}^T f_{Y_t|Y_{t-1}}(y_t; y_{t-1}).$$

Proposition 4.2 (Log-likelihood in the inversion context). *In the context of a linear state-space model as defined in Def. 4.1, under Assumption 4.1, and if w_t is a Markovian process, we have:*

$$\begin{aligned} f_{Y_t|Y_{t-1}}(y_t; y_{t-1}) &= \underbrace{f_{w_t|w_{t-1}}(w(y_{1,t}); w(y_{t-1}))}_{\text{blue}} \times \\ &\quad \underbrace{\mathbf{n}(y_{2,t}; A_2 + B_2 w(y_{1,t}), \Omega_2)}_{\text{red}} \times |B_1|^{-1}. \end{aligned} \quad (4.8)$$

where $w(y_{1,t}) = B_1^{-1}(y_{1,t} - A_1)$ and where \mathbf{n} denotes the multivariate normal p.d.f. (Eq. (4.3)).

Proof. Since w_t is Markov, so is $y_{1,t}$ and since $y_{2,t} = A_2 + B_2 w(y_{1,t}) + \eta_{2,t}$, with $w(y_{1,t}) = B_1^{-1}(y_{1,t} - A_1)$, we have:

$$\begin{aligned} f(y_t|y_{t-1}) &= f_1(y_{1,t}|y_{1,t-1})f_2(y_{2,t}|y_{1,t}) \\ &= |B_1|^{-1}f_w(w(y_{1,t})|w(y_{1,t-1}))\mathbf{n}(y_{2,t}; A_2 + B_2 w(y_{1,t}), \Omega_2), \end{aligned}$$

where * $f_1(y_{1,t}|y_{1,t-1}) = |B_1|^{-1}f_w(w(y_{1,t})|w(y_{1,t-1}))$ comes from the fact that, if U and V are two random variables such that $V = g(U)$, where g is a bijective and differentiable function, then $f_V(v) = \left| \frac{\partial g^{-1}(v)}{\partial v'} \right| f_U(g^{-1}(v))$, * and $f_2(y_{2,t}|y_{1,t}) = \mathbf{n}(y_{2,t}; A_2 + B_2 w(y_{1,t}), \Omega_2)$ comes from the fact that $y_{2,t}|y_{1,t} \sim \mathbf{n}(A_2 + B_2 w(y_{1,t}), \Omega_2)$. \square

4.4.1 Dealing with Unobserved Regimes

- The (adjusted) Kalman and inversion techniques are not suited to the case where some of the components of w_t are valued in a discrete set. This is typically the case if w_t is of the form:

$$w_t = \begin{pmatrix} z_t \\ x_t \end{pmatrix},$$

where z_t is valued in $\{e_1, \dots, e_J\}$, e_j being the j^{th} column of Id_J .

- Let's assume that z_t is an exogenous and homogenous Markov chain whose dynamics is defined by the $\pi(e_i, e_j)$'s that are such that:

$$\pi(e_i, e_j) = \mathbb{P}(z_{t+1} = e_j | z_t = e_i). \quad (4.9)$$

- We denote by Π the matrix of transition probabilities, i.e. the (i, j) component of Π is $\pi(e_i, e_j)$.
- Assume further that we have:

$$x_t = m(z_t, x_{t-1}) + \varepsilon_t, \quad (4.10)$$

where $\mathbb{E}(\varepsilon_t | \underline{z_t}, \underline{x_{t-1}}) = 0$.

- The conditional distribution of ε_t w.r.t. (z_t, x_{t-1}) is denoted by $f_\varepsilon(\cdot; z_t, x_{t-1})$.

Hypothesis 4.2 (Measurement equations). The measurement equation is of the form:

$$y_t = Az_t + Bx_t + \eta_t, \quad \text{with } \eta_t \sim i.i.d. \mathbf{n}(0, \Omega). \quad (4.11)$$

Example 4.1 (Regime-Switching Gaussian VAR).

- Building on Example 1.15, we know that if

$$x_t = \mu z_t + \Phi x_{t-1} + \varepsilon_t, \quad (4.12)$$

where $\varepsilon_t | x_t, z_t \sim \mathcal{N}(0, \Sigma(z_t))$ and if z_t is an exogenous independent Markov chain, then $w_t = (x'_t, z'_t)'$ is affine.

- This context is consistent with the situation presented XXX. In particular, using the notations of Eq. (4.10), we have:

$$m(z_t, x_{t-1}) = \mu z_t + \Phi x_{t-1}.$$

- If r_t and the s.d.f. are respectively affine and exponential in w_t , then, in particular, yields are also affine in w_t , i.e. of the form $R(t, h) = A'_h z_t + B'_h x_t$ (see Eq. (3.8)).
- Therefore, if the components of y_t are yields of different maturities, the measurement equations are consistent with Assumption 4.2.
- How to estimate such a model when the regimes z_t are unobservable? Two distinct situations.
- **Case 1.** The x_t factors are observable.
- The probabilities of being in the different regimes on each date can be estimated by employing the Kitagawa-Hamilton filter.
- The Kitagawa-Hamilton filter (Appendix ??) can be employed, with $F_t = (y'_t, x'_t)'$ and:

$$\begin{aligned} f(F_t | z_t = e_j, \underline{F}_{t-1}) &= f(y_t | x_t, z_t = e_j, \underline{F}_{t-1}) f(x_t | z_t = e_j, \underline{F}_{t-1}) \\ &= \mathbf{n}(y_t; Az_t + Bx_t, \Omega) \times \\ &\quad f_\varepsilon(x_t - m(z_t, x_{t-1}); z_t, x_{t-1}), \end{aligned}$$

where, as in Slide ??, $\mathbf{n}(u; \mu, \Omega)$ denotes the evaluation, at vector u of the p.d.f. of the multivariate normal distribution $\mathcal{N}(\mu, \Omega)$.

- A by-product of the Kitagawa-Hamilton filter is the likelihood function associated with the dataset \Rightarrow The model parameterization can be estimated by MLE.
- **Case 2.** The x_t factors are not observable. Two sub-cases:
 - 2.i The components of y_t are not perfectly modelled (i.e. $\Omega \neq 0$, where Ω defined in Eq. (4.11)).
 - 2.ii n_x components of y_t are perfectly modelled (where n_x is the dimension of x_t).
- In Case (2.i), one has to resort to filters dealing with two types of uncertainty [with hidden discrete values (z_t) and continuously distributed latent variables (x_t)].

In particular Kim (1994)'s filter can be employed when the state-space model is of the form (4.11)-(4.12) Monfort and Renne (2014), Example ??]. * In Case (2.ii), one can resort to an inversion technique, complemented with the Kitagawa-Hamilton filter, to estimate the model (see following box).

Proposition 4.3 (Kitagawa-Hamilton filter). *Consider a q -dimensional vector of variables F_t and an exogenous homogenous Markov chain z_t . We make use of the following notations:*

- η_t is a J -dimensional vector whose j^{th} component is the p.d.f. of F_t conditional on $(z_t = e_j, \underline{F}_{t-1})$, i.e. $f(F_t|z_t = e_j, \underline{F}_{t-1})$
- ξ_t is a J -dimensional vector whose j^{th} component is $\mathbb{P}(z_t = e_j|\underline{F}_t)$.

The sequence ξ_t can then be computed recursively as follows:

$$\xi_t = \frac{(\Pi' \xi_{t-1}) \odot \eta_t}{\mathbf{1}'(\Pi' \xi_{t-1} \odot \eta_t)}, \quad (4.13)$$

where \odot denotes the element-by-element (Hadamard) product and where $\mathbf{1}$ denotes a J -dimensional vector of ones.

Moreover, the previous formulas also show how to compute the likelihood of the model since:

$$f(F_t|\underline{F}_{t-1}) = \mathbf{1}'(\Pi' \xi_{t-1} \odot \eta_t). \quad (4.14)$$

Proposition 4.4 (Kitagawa-Hamilton and inversion techniques). *The matrix of transition probabilities of \mathcal{Z}_t is of the form $\mathbf{1}_{n \times 1} \otimes \tilde{\Pi}$, with*

$$\tilde{\Pi} = \begin{bmatrix} \pi_{1,\bullet} & 0_{1 \times n} & \dots & & 0_{1 \times n} \\ 0_{1 \times n} & \pi_{2,\bullet} & 0_{1 \times n} & \dots & 0_{1 \times n} \\ & & \ddots & & \\ & & 0_{1 \times n} & \pi_{n-1,\bullet} & 0_{1 \times n} \\ 0_{1 \times n} & \dots & & 0_{1 \times n} & \pi_{n,\bullet} \end{bmatrix},$$

where $\pi_{i,\bullet}$ denotes the i^{th} row of Π (Π is defined on Slide XXX).

The last term appearing in Eq. (4.16) can be computed as follows:

$$\begin{aligned} & f \left(\begin{bmatrix} x(y_t, z(\mathcal{Z}_t)) \\ y_{2,t} \end{bmatrix} \middle| \mathcal{Z}_t, \underline{y_{t-1}} \right) \\ &= f(y_{2,t} | x_t = x(y_t, z(\mathcal{Z}_t)), \mathcal{Z}_t, \underline{y_{t-1}}) \times \\ & \quad f(x(y_t, z(\mathcal{Z}_t)) | \mathcal{Z}_t, \underline{y_{t-1}}) \\ &= \mathbf{n}(y_{2,t}; A_2 z_t + B_2 x_t, \Omega_2) \times \\ & \quad f_\varepsilon(\varepsilon_t | z_t = z(\mathcal{Z}_t), x_{t-1} = x(y_{t-1}, z_{-1}(y_{t-1}, \mathcal{Z}_t))), \end{aligned}$$

where \mathbf{n} is the p.d.f. of a multivariate normal distri, (Prop. 4.2, where $\varepsilon_t = x_t - m[z(\mathcal{Z}_t), x(y_{t-1}, z_{-1}(y_{t-1}, \mathcal{Z}_t))]$ (m defined in Eq. (4.10)).

4.4.2 Mixed use of Kitagawa-Hamilton and inversion techniques

- Without loss of generality, assume that the n_x first components of y_t are observed without error, i.e. Assumption 4.2 becomes

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} A_1 z_t \\ A_2 z_t \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ \eta_{2,t} \end{bmatrix},$$

where $\varepsilon_2 \sim \mathcal{N}(0, \Omega_2)$.

- Since $y_{1,t} = A_1 z_t + B_1 x_t$, we then have:

$$x_t \equiv x(y_t, z_t) = B_1^{-1}(y_{1,t} - A_1 z_t). \quad (4.15)$$

- In order to employ the Kitagawa-Hamilton filter (Appendix ??), one need to define the extended Markov chain:

$$\mathcal{Z}_t = z_{t-1} \otimes z_t,$$

whose matrix of transition probabilities is detailed in Appendix ??.

- Note that we have:

$$\begin{cases} z_t & \equiv z(\mathcal{Z}_t) & = (\mathbf{1}' \otimes Id_n) \mathcal{Z}_t \\ z_{t-1} & \equiv z_{-1}(\mathcal{Z}_t) & = (Id_n \otimes \mathbf{1}') \mathcal{Z}_t. \end{cases}$$

- The Kitagawa-Hamilton filter (Appendix ??) can then be employed, with $F_t = y_t$ and:

$$f(y_t | \mathcal{Z}_t, \underline{y_{t-1}}) = |B_1^{-1}| \times f\left(\begin{bmatrix} x(y_t, z(\mathcal{Z}_t)) \\ y_{2,t} \end{bmatrix} | \mathcal{Z}_t, \underline{y_{t-1}}\right), \quad (4.16)$$

where the computation of the last term is detailed in Appendix ??.

Example 4.2 (Interest-rate model with monetary policy-related regimes).

- Renne (2017): model where regimes have monetary-policy interpretations.
- Short-term rate = (daily) EONIA, euro-area overnight interbank rate:

$$\boxed{r_t = \underbrace{\bar{r}_t}_{\text{Target}} + \underbrace{x_t}_{\text{EONIA spread}}}$$

- Target rate \bar{r}_t has a step-like path $\bar{r}_t = \Delta' z_{r,t}$, where
- $\Delta = [0 \ 0.25 \ 0.50 \ \dots \ \bar{r}_{max}]'$ and
- $z_{r,t} = [0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0]'$
- EONIA spread (x_t) persistent and mean-reverting fluctuations (AR process).
- $z_t = z_{r,t} \otimes z_{m,t}$ where $z_{m,t}$ is the monetary-policy regime:
- Easing ($z_{m,t} = [1 \ 0 \ 0]$),
- Status Quo ($z_{m,t} = [0 \ 1 \ 0]$),
- Tightening ($z_{m,t} = [0 \ 0 \ 1]$).

Web-interface illustrating \bar{r}_t 's dynamic

- $z_{r,t}$ is observed, but not $z_{m,t}$.
- In the model, according to Example 4.1, we have:

$$R(t, h) = A'_h z_t + B_h x_t.$$

- Denote by \mathcal{A}_h the $(3 \times n_r)$ matrix such that $A_h = \text{vec}(\mathcal{A}_h)$. We then have $A'_h z_t = (\mathcal{A}_h z_{r,t})' z_{m,t}$ and, therefore:

$$R(t, h) = A'_{t,h} z_t + B_h x_t, \quad \text{where } A_{t,h} = \mathcal{A}_h z_{r,t}.$$

- The model is estimated by a mixe use of Kitagawa-Hamilton and inversion techniques, assuming that a linear combination of yields is modelled without errors (giving $x_t = x(y_t, z_{m,t}, z_{r,t})$).

4.5 A Typical Small-Sample Issue

Interest rates are particularly persistent variables.

Since affine models eventually lead to linear relationships between state variables and interest rates [Eq. (3.8)], some state variables are also necessarily highly persistent.

In small sample, maximum-likelihood estimates of the model parameters are likely to suffer from a downward bias, Michael D. Bauer and Wu (2012) or Jardet et al. (2013).

This relates to a well-known econometric problem:

- This small-sample downward bias has dramatic consequences in terms of term premium estimates.
- For the sake of illustration, let's consider the following process for the short-term interest rate under the physical measure (monthly frequency):

$$i_{t+1} = \bar{i} + \phi(i_t - \bar{i}) + \sigma \varepsilon_{t+1}, \quad \varepsilon_t \sim \mathcal{N}(0, 1).$$

and the following under the risk-neutral measure:

$$i_{t+1} = \bar{i}^* + \phi^*(i_t - \bar{i}^*) + \sigma \varepsilon_{t+1}, \quad \varepsilon_t \sim \mathcal{N}(0, 1).$$

with $\bar{i} = 3\%/12$, $\bar{i}^* = 6\%/12$, $\phi = 0.97$, $\phi^* = 0.99$, $\sigma = 0.2\%$.

- Assume the estimate of ϕ is downward biased ($\hat{\phi} = 0.9$). Influence on the 10-year term premium? [Figure ??, next slide]
- Kim and Orphanides (2005) and Kim and Orphanides (2005) have proposed a simple approach to deal with this problem.
- They add measurement equations to impose that the model-implied forecasts are –up to some measurement errors– close to survey-based ones.
- Kim and Orphanides use the Blue Chip Financial Forecasts.
- Alternative (publicly available) surveys: Philly Fed Survey of Professional Forecasters and ECB SPF.
- Importantly, this approach is tractable because model-implied forecasts of yields are affine in the state vector. [This is the case for any affine model: see Eq. (3.11).]
- Their model is able to fit both survey-based forecasts and market yields [Figure ??].

Chapter 5

Forward, futures, dividends, commodity pricing, and convenience yields

5.1 Forward Contracts

- w_t : information at t
- Forward contract : Agreement signed at t (to-day) to buy/sell an asset (a commodity) at a given delivery date $T > t$ at a price $\Phi_{t,T}$ (*delivery price* or *forward price*) decided at t .
- S_T : value of the asset (commodity) at T , (function of w_T).
- $S_T - \Phi_{t,T}$: payoff of the contract at T .
- r_t : interest rate between t and $t + 1$ (known at t), (function of w_t).
- $B(t_1, t_2 - t_1)$: date- t_1 price of a zero-coupon bond whose value is 1 on t_2 .

Proposition 5.1 (Forward price). *We have:*

$$\begin{aligned}\Phi_{t,T} &= \frac{\mathbb{E}_t(M_{t,t+1}, \dots, M_{T-1,T} S_T)}{B(t, T-t)} \\ &= \frac{\mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{T-1}) S_T]}{B(t, T-t)}.\end{aligned}$$

::: {proof} The price at t of $S_T - \Phi_{t,T}$ is zero. Therefore: $\mathbb{E}_t(M_{t,t+1}, \dots, M_{T-1,T} S_T) - \Phi_{t,T} \mathbb{E}_t(M_{t,t+1} \dots M_{T-1,T}) = 0$. The result follows. :::

- If the processes (r_t) and (S_t) are \mathbb{Q} -independent, then

$$\Phi_{t,T} = \mathbb{E}_t^{\mathbb{Q}}(S_T)$$

and $\Phi_{t,T}$ is a \mathbb{Q} -martingale.

- If the asset does not generate any payoff before T (no dividends), then

$$\Phi_{t,T} = \frac{S_t}{B(t, T-t)}.$$

- The price at s , with $t < s < T$, of a forward contract signed at t is:

$$(\Phi_{s,T} - \Phi_{t,T})B(s, T-s),$$

since the payoff $S_T - \Phi_{t,T}$ at T has this price at s .

- Notations associated with **dividends**:
- \tilde{S}_t : ex-dividend price at t ,
- S_t : cum-dividend price at t ,
- $S_t = \tilde{S}_t \exp(\delta_t)$, δ_t : dividend yield (or rate) known at t .
- We get:

$$\begin{aligned} \tilde{S}_{T-1} &= \mathbb{E}_{T-1}(M_{T-1,T} S_T) \\ \text{or} \\ S_{T-1} &= \mathbb{E}_{T-1}[M_{T-1,T} \exp(\delta_{T-1}) S_T], \end{aligned}$$

and, recursively:

$$\begin{aligned} S_t &= \mathbb{E}_t[M_{t,t+1} \dots M_{T-1,T} \exp(\delta_t + \dots + \delta_{T-1}) S_T] \\ &= \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t \dots - r_{T-1} + \delta_t + \dots + \delta_{T-1}) S_T]. \end{aligned}$$

$\Rightarrow r_t$ replaced by $r_t - \delta_t$.

- Replacing S_t by $\tilde{S} \exp(\delta_t)$ and S_T by $\tilde{S}_T \exp(\delta_T)$, we get

$$\tilde{S}_t = \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{T-1} + \delta_{t+1} + \dots + \delta_T) \tilde{S}_T].$$

- If deterministic δ_t :

$$\Phi_{t,T} = \frac{S_t \exp(-\delta_t - \dots - \delta_{T-1})}{B(t, T - t)}.$$

If r_t deterministic too:

$$\Phi_{t,T} = S_t \exp(r_t + \dots + r_{T-1} - \delta_t - \dots - \delta_{T-1}).$$

5.2 Futures Contracts, Futures Prices

- Futures contract: agreement signed at t to buy/sell an asset (a commodity) at given delivery date $T > t$ at a price $F_{t,T}$ (futures price) decided at t .
- Difference with forward contract: both counterparties are required to deposit into a **margin account** at every trading day $s > t$ the resettlement payment (**margin call**). The latter is equal to:

$$\Delta_{s,T} = F_{s,T} - F_{s-1,T} \quad (\text{for the buyer}).$$

\Rightarrow A Futures contract is actually closed out after every day, starts afresh the next day and therefore is valued zero.

Proposition 5.2 (Pricing futures). *We have:*

$$F_{t,T} = \mathbb{E}_t^{\mathbb{Q}}(S_T),$$

that is $F_{t,T}$ is a \mathbb{Q} -martingale, or $\Delta_{t,T}$ is a \mathbb{Q} -martingale difference.

Proof. At each date $s \geq t$, after the deposit of the resettlement payment, there is a new contract valued zero and paying $F_{s+1,T} - F_{s,T}$ at $s + 1$ (and providing another zero valued contract at $s + 1$). Therefore $0 = \mathbb{E}_s^{\mathbb{Q}}[\exp(-r_s)(F_{s+1,T} - F_{s,T})]$, and $0 = \mathbb{E}_s^{\mathbb{Q}}(F_{s+1,T} - F_{s,T})$ since $\exp(-r_s)$ is known at s . Hence $F_{s,T} = \mathbb{E}_s^{\mathbb{Q}}(F_{s+1,T})$, and the results follows from $F_{T,T} = S_T$. \square

Proposition 5.3 (XXXX).

$$\Phi_{t,T} - F_{t,T} = \frac{cov_t^{\mathbb{Q}} \left[\prod_{s=t}^{T-1} \exp(-r_s), S_T \right]}{B(t, T-t)}.$$

Proof. We have:

$$\begin{aligned} \Phi_{t,T} - F_{t,T} &= \mathbb{E}_t^{\mathbb{Q}} \left[\frac{\prod_{s=t}^{T-1} \exp(-r_s) S_T}{B(t, T-t)} - S_T \right] \\ &= \frac{\mathbb{E}_t^{\mathbb{Q}} \left[\prod_{s=t}^{T-1} \exp(-r_s) S_T \right] - \mathbb{E}_t^{\mathbb{Q}} \left[\prod_{s=t}^{T-1} \exp(-r_s) \right] \mathbb{E}_t^{\mathbb{Q}}(S_T)}{B(t, T-t)} \\ &= \frac{cov_t^{\mathbb{Q}} \left[\prod_{s=t}^{T-1} \exp(-r_s), S_T \right]}{B(t, T-t)}. \end{aligned}$$

□

- $\Phi_{t,T} = F_{t,T}$, if, and only if, $\prod_{s=t}^{T-1} \exp(-r_s)$ and S_T are conditionally uncorrelated under \mathbb{Q} . In particular it is true in the case of deterministic short rates.

5.3 Convenience Yields

:::{.definition #convYield name=“Convenience yield”} A convenience yield is a net benefit associated with holding a physical asset (rather than a forward or futures contract). It is *net* in the sense that it is equal to the positive gain of holding minus the cost of storage. :::

- The notion of convenience yield is valid only for storable commodities (not, e.g., for electricity).
- Can be positive or negative.
- Notion of convenience yield mathematically similar to a dividend yield (but can be < 0 , and latent).

the following, we denote the convenience yield by c_t . * The price is here the *cum* convenience yield price.

Basic pricing approach

- r_t and c_t deterministic. We have:

$$\Phi_{t,T} = F_{t,T} = S_t \exp(r_t + \dots + r_{T-1} - c_t - \dots - c_{T-1}).$$

- Moreover, if r_t and c_t time independent:

$$\Phi_{t,T} = F_{t,T} = S_t \exp[(T-t)(r-c)].$$

- If $r > c$ forward curve function \uparrow of T : situation of **contango**;
- If $r < c$ forward curve function \downarrow of T : **backwardated** forward curve.

Advanced approach: c_t is stochastic.

- We have:

$$\Phi_{t,T} = \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{T-1})S_T]/B(t, T-t) \quad (5.1)$$

$$F_{t,T} = \mathbb{E}_t^{\mathbb{Q}}(S_T) \quad (5.2)$$

$$S_t = \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{T-1} + c_t + \dots + c_{T-1})S_T]. \quad (5.3)$$

- necessitates a joint modelling (at least under \mathbb{Q}) of: S_t (or $s_t = \log S_t$), c_t (convenience yield), r_t (short rate),
- and, in general, other factors (observable or latent).

Direct or backward approach (see XXX).

- Case of a **nonstorable commodity**. No convenience yield; if, moreover r_t deterministic, $\Phi_{t,T} = F_{t,T} = \mathbb{E}_t^{\mathbb{Q}}(S_T)$, but since Eq. (@ref{eq:convenience3}) is not valid we do not have $S_t = B(t, T)\mathbb{E}_t^{\mathbb{Q}}(S_T)$ and therefore we do not have $\Phi_{t,T} = S_t/B(t, T)$.
\end{remark}
- Case of a **storable commodity** The price ex-convenience yield at t , i.e. $S_t \exp(-c_t)$, is the price at t of S_{t+1} :

$$\begin{aligned} S_t \exp(-c_t) &= \exp(-r_t) \mathbb{E}_t^{\mathbb{Q}}(S_{t+1}) \\ \mathbb{E}_t^{\mathbb{Q}}(S_{t+1}) &= S_t \exp(r_t - c_t) \\ \mathbb{E}_t^{\mathbb{Q}} \exp(s_{t+1}) &= \exp(s_t + r_t - c_t). \end{aligned} \quad (5.4)$$

Or if $y_{t+1} = \log \frac{S_{t+1}}{S_t}$ denotes the (geometric) return, then:

$$\mathbb{E}_t^{\mathbb{Q}} \exp(y_{t+1}) = \exp(r_t - c_t).$$

- Eq. (5.3) is automatically satisfied (using Eq. (5.4) recursively).

5.4 Pricing with Affine Models

- If storable commodity:

$$w_t = (s_t, c_t, r_t, x_t')',$$

where $s_t = \log S_t$ and x_t is a vector of additional factors.

- w_t \mathbb{Q} -affine:

$$\mathbb{E}_t^{\mathbb{Q}} \exp(u' w_{t+1}) = \exp[a'(u) w_t + b(u)].$$

- If $S_{t+1} = \exp(e_1' w_{t+1})$, ICCs give:

$$\begin{aligned} \mathbb{E}_t^{\mathbb{Q}} \exp(s_{t+1}) &= \exp(s_t - c_t + r_t) \\ \Leftrightarrow \mathbb{E}_t^{\mathbb{Q}} \exp(e_1' w_{t+1}) &= \exp[(e_1 - e_2 + e_3)' w_t] \\ &\Leftrightarrow \begin{cases} a(e_1) = e_1 - e_2 + e_3 \\ b(e_1) = 0 \end{cases} \end{aligned}$$

(if non storable, no c_t , no ICCs).

- Since (Eqs. (??) and (5.2)):

$$\begin{aligned} \Phi_{t,T} &= \frac{\mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{t-1} + s_T)]}{\mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t - \dots - r_{T-1})]} \\ F_{t,T} &= \mathbb{E}_t^{\mathbb{Q}}[\exp(s_T)], \end{aligned}$$

\Rightarrow quasi explicit formulas using multihorizon Laplace transforms, exponential affine in $w_t = (s_t, c_t, r_t, x_t')'$.

5.5 Historical dynamics

- We have:

$$M_{t,t+1}(\underline{w}_{t+1}) = \exp[-r_t + \alpha_t' w_{t+1} + \psi_t^{\mathbb{Q}}(-\alpha_t)],$$

and

$$\mathbb{E}_t^{\mathbb{Q}}(M_{t,t+1}^{-1}) = \exp(r_t).$$

- Historical dynamics defined through:

$$\psi_t(u) = \psi_t^{\mathbb{Q}}(u - \alpha_t) - \psi_t^{\mathbb{Q}}(-\alpha_t).$$

- Note: no constraint on α_t , possibility to introduce seasonality.

5.6 A Gaussian VAR Model

- State vector: $w_t = (s_t, c_t, r_t)'$.
- **R.N. dynamics:**

$$w_{t+1} = A_0 + A_1 w_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim i.i.d. \mathcal{N}(0, \Sigma) \text{ under } \mathbb{Q}.$$

- We have (see Example 1.6):

$$\begin{aligned} \mathbb{E}_t \exp(u' w_{t+1}) &= \exp \left[u' (A_0 + A_1 w_t) + \frac{1}{2} u' \Sigma u \right] \\ \Rightarrow \quad \begin{cases} a(u) &= A_1' u \\ b(u) &= A_0' u + \frac{1}{2} u' \Sigma u. \end{cases} \end{aligned}$$

- Let us decompose A_0 and A_1 as follows:

$$A_0 = \begin{pmatrix} A_{01} \\ \tilde{A}_0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} A_{11} \\ \tilde{A}_1 \end{pmatrix}.$$

- ICCs (if storable, otherwise no constraint):

$$\begin{cases} a(e_1) &= e_1 - e_2 + e_3 \\ b(e_1) &= 0. \end{cases}$$

\Leftrightarrow The first row of A_1 is $A_{11} = e_1' - e_2' + e_3'$, the first element of A_0 is $A_{01} = -\frac{1}{2}\sigma_1^2$ (σ_1^2 conditional variance of s_{t+1}).

- In other words the \mathbb{Q} -VAR is:

$$\begin{cases} s_{t+1} &= -\frac{1}{2}\sigma_1^2 + s_t - c_t + r_t + \varepsilon_{1,t+1} \\ \begin{pmatrix} c_{t+1} \\ r_{t+1} \end{pmatrix} &= \tilde{A}_0 + \tilde{A}_1 w_t + \begin{pmatrix} \varepsilon_{2,t+1} \\ \varepsilon_{3,t+1} \end{pmatrix}, \end{cases}$$

(where \tilde{A}_0 and \tilde{A}_1 are not constrained).

- Noting $y_{t+1} = \log(S_{t+1}/S_t) = s_{t+1} - s_t$, this implies:

$$\mathbb{E}_t^{\mathbb{Q}} y_{t+1} = r_t - c_t - \frac{1}{2}\sigma_1^2,$$

$\Rightarrow y_{t+1} - r_t + c_t + \frac{1}{2}\sigma_1^2$ is a \mathbb{Q} -martingale difference.

- **Historical dynamics:**

$$\begin{aligned}\psi_t(u) &= \psi_t^{\mathbb{Q}}(u - \alpha_t) - \psi_t^{\mathbb{Q}}(-\alpha_t) \\ &= u'(A_0 + A_1 w_t) + \frac{1}{2}(u - \alpha_t)' \Sigma (u - \alpha_t) - \frac{1}{2} \alpha_t' \Sigma \alpha_t \\ &= u'(A_0 + A_1 w_t) - u' \Sigma \alpha_t + \frac{1}{2} u' \Sigma u.\end{aligned}$$

- If we take $\alpha_t = \alpha_0 + \alpha_1 w_t$, we get:

$$\begin{aligned}\psi_t(u) &= u'[A_0 - \Sigma \alpha_0 + (A_1 - \Sigma \alpha_1) w_t] + \frac{1}{2} u' \Sigma u \\ \Rightarrow w_{t+1} &= A_0 - \Sigma \alpha_0 + (A_1 - \Sigma \alpha_1) w_t + \xi_t, \\ \xi_t &\sim i.i.d. \mathcal{N}(0, \Sigma) \text{ under } \mathbb{P}.\end{aligned}$$

- any VAR(1), with same Σ , can be reached.
- We can also take: $\alpha_t = \alpha_{0t} + \alpha_1 w_t$. We then get the historical dynamics:

$$w_{t+1} = A_0 - \Sigma \alpha_{0t} + (A_1 - \Sigma \alpha_1) w_t + \xi_{t+1}$$

- We can choose α_{0t} such that: $A_0 - \Sigma \alpha_{0t} = \begin{pmatrix} \mu_{1t} \\ \mu_2 \\ \mu_3 \end{pmatrix}$

(μ_{1t}, μ_2, μ_3 given). * In particular:

$$\begin{aligned}s_{t+1} &= \mu_{1t} + (A_{11} - \Sigma_1 \alpha_1) w_t + \xi_{1,t+1} \\ &= \mu_{1t} + \bar{A}_{11} w_t + \xi_{1,t+1} \text{ (say),}\end{aligned}$$

\bar{A}_{11} not constrained, or:

$$s_{t+1} - \nu_{t+1} = \mu_1 + \bar{A}_{11,1}(s_t - \nu_t) + \bar{A}_{11,2}c_t + \bar{A}_{11,3}r_t + \xi_{1,t+1}$$

with

$$\mu_{1t} = \mu_1 + \nu_{t+1} - \bar{A}_{11,1}\nu_t.$$

* In other words any historical seasonal function ν_t can be reached by choosing μ_{1t} , i.e. α_{0t} .

5.7 Applications

- Source: Schwartz (1997).
- Three models whose discrete-time versions are:
- **Model 1:**

$$\begin{aligned}\mathbb{P} &: s_{t+1} = a_0 + a_1 s_t + \sigma \varepsilon_{t+1}, \quad \varepsilon_t \stackrel{\mathbb{P}}{\sim} i.i.d. \mathcal{N}(0, 1) \\ \mathbb{Q} &: s_{t+1} = a_0^* + a_1 s_t + \sigma \xi_{t+1}, \quad \xi_t \stackrel{\mathbb{Q}}{\sim} i.i.d. \mathcal{N}(0, 1),\end{aligned}$$

and $\log F_{t,T}$ affine functions of s_t .

- **Model 2:**

$$\begin{aligned}\mathbb{P} &: s_{t+1} = a_{10} + s_t - c_t + \sigma_1 \varepsilon_{1,t+1} \\ &\quad c_{t+1} = a_{20} + a_{21} c_t + \sigma_2 \varepsilon_{2,t+1} \quad \text{corr}(\varepsilon_{1t}, \varepsilon_{2t}) = \rho \\ \mathbb{Q} &: s_{t+1} = r - \frac{\sigma_1^2}{2} + s_t - c_t + \sigma_1 \xi_{1,t+1} \\ &\quad c_{t+1} = a_{20}^* + a_{21} c_t + \sigma_2 \xi_{2,t+1} \quad \text{corr}(\xi_{1t}, \xi_{2t}) = \rho.\end{aligned}$$

- **Model 3:**

$$\begin{aligned}\mathbb{P} &: s_{t+1} = a_{10} + s_t - c_t + r_t + \sigma_1 \varepsilon_{1,t+1} \\ &\quad c_{t+1} = a_{20} + a_{21} c_t + \sigma_2 \varepsilon_{2,t+1} \\ &\quad r_{t+1} = a_{30} + a_{31} r_t + \sigma_3 \varepsilon_{3,t+1} \\ \mathbb{Q} &: s_{t+1} = -\frac{\sigma_1^2}{2} + s_t - c_t + r_t + \sigma_1 \xi_{1,t+1} \\ &\quad c_{t+1} = a_{20}^* + a_{21} c_t + \sigma_2 \xi_{2,t+1} \\ &\quad r_{t+1} = a_{30} + a_{31} r_t + \sigma_3 \xi_{3,t+1},\end{aligned}$$

$\text{corr}(\varepsilon_{1t}, \varepsilon_{2t}) = \rho_1$, $\text{corr}(\varepsilon_{2t}, \varepsilon_{3t}) = \rho_2$, $\text{corr}(\varepsilon_{1t}, \varepsilon_{3t}) = \rho_3$, idem for the ξ_{it} 's.

$\log F_{t,T}$ affine functions of s_t, c_t, r_t .

- In these models s_t, c_t are latent, the parameters are estimated by the ML method. The likelihood functions are computed by the Kalman filter.
- The measurement equations are of the form: $\log F_{t,T} = \text{affine functions of the state variables} + \text{error for various values of } t \text{ and } T$.

- State variables: s_t or (s_t, c_t) , (the \mathbb{P} and \mathbb{Q} dynamics of r_t is estimated separately).

(Constraints between the \mathbb{P} and \mathbb{Q} parameters \Rightarrow helps identification)

Out of sample means for maturities not used at the estimation stage. F1 contract : the closest to maturity F2 : the second contract to maturity and so one. Source: Schwartz (1997).

Chapter 6

References

Bibliography

- Adrian, T., Crump, R. K., and Moench, E. (2013). Pricing the Term Structure with Linear Regressions. *Journal of Financial Economics*, 110(1):110–138.
- Ang, A., Boivin, J., Dong, S., and Loo-Kung, R. (2011). Monetary Policy Shifts and the Term Structure. *Review of Economic Studies*, 78(2):429–457.
- Ang, A. and Piazzesi, M. (2003). A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics*, 50(4):745–787.
- Black, F. (1995). Interest Rates as Options. *The Journal of Finance*, 50(5):1371–1376.
- Chikhani, P. and Renne, J.-P. (2023). An Analytical Framework to Price Long-Dated Climate-Exposed Assets. University of lausanne working paper, University of Lausanne.
- Chow, G. C. and Lin, A.-l. (1971). Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series. *The Review of Economics and Statistics*, 53(4):372–375.
- Christensen, J. H. E., Diebold, F. X., and Rudebusch, G. D. (2009). An Arbitrage-free Generalized Nelson–Siegel Term Structure Model. *The Econometrics Journal*, 12(3):C33–C64.
- Duffee, G. (2010). Sharpe Ratios in Term Structure Models. Economics Working Paper Archive 575, The Johns Hopkins University, Department of Economics.

- Duffie, D., Pan, J., and Singleton, K. (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica*, 68(6):1343–1376.
- Duffie, D. and Singleton, K. J. (1993). Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica*, 61(4):929–952.
- Duffie, D. and Singleton, K. J. (1997). An Econometric Model of the Term Structure of Interest-Rate Swap Yields. *The Journal of Finance*, 52(4):1287–1321.
- Duffie, D. and Stein, J. C. (2015). Reforming LIBOR and Other Financial Market Benchmarks. *Journal of Economic Perspectives*, 29(2):191–212.
- Gourieroux, C. and Jasiak, J. (2006). Autoregressive Gamma Processes. *Journal of Forecasting*, 25(2):129–152.
- Gouriéroux, C. and Monfort, A. (1997). *Simulation-based Econometric Methods*. Oxford University Press.
- Gourieroux, C., Monfort, A., Mouabbi, S., and Renne, J.-P. (2021). Disastrous Defaults. *Review of Finance*, 25(6):1727–1772.
- Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Theory. *Econometrica*, 52(3):681–700.
- Hansen, L. P. and Jagannathan, R. (1991). Implications of Security Market Data for Models of Dynamic Economies. *Journal of Political Economy*, 99(2):225–262.
- Harvey, A. C. and Pierse, R. G. (1984). Estimating Missing Observations in Economic Time Series. *Journal of the American Statistical Association*, 79:1–36.
- Jardet, C., Monfort, A., and Pegoraro, F. (2013). No-Arbitrage Near-Cointegrated VAR(p) Term Structure Models, Term Premia and GDP Growth. *Journal of Banking & Finance*, 37(2):389–402.
- Joslin, S., Pribsch, M., and Singleton, K. J. (2014). Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks. *The Journal of Finance*, 69(3):1197–1233.

- Kim, D. H. and Orphanides, A. (2005). Term Structure Estimation with Survey Data on Interest Rate Forecasts. Finance and Economics Discussion Series 2005-48, Board of Governors of the Federal Reserve System (U.S.).
- Krippner, L. (2013). A Tractable Framework for Zero Lower Bound Gaussian Term Structure Models. Reserve Bank of New Zealand Discussion Paper Series DP2013/02, Reserve Bank of New Zealand.
- Michael D. Bauer, G. D. R. and Wu, J. C. (2012). Correcting Estimation Bias in Dynamic Term Structure Models. *Journal of Business & Economic Statistics*, 30(3):454–467.
- Monfort, A., Pegoraro, F., Renne, J.-P., and Roussellet, G. (2017). Staying at Zero with Affine Processes: An Application to Term Structure Modelling. *Journal of Econometrics*, 201(2):348–366.
- Nelson, C.-J. and Kim, C. R. (1999). *State-Space Models with Regime-Switching*. M.i.t. press edition.
- Pallara, K. and Renne, J.-P. (2023). Fiscal Limits and the Pricing of Eurobonds. *Management Science*, 0(0):null.
- Pribsch, M. A. (2013). Computing Arbitrage-Free Yields in Multi-Factor Gaussian Shadow-Rate Term Structure Models. Finance and Economics Discussion Series 2013-63, Board of Governors of the Federal Reserve System (U.S.).
- Renne, J.-P. (2017). A Model of the Euro-Area Yield Curve with Discrete Policy Rates. *Studies in Nonlinear Dynamics & Econometrics*, 21(1):99–116.
- Schwartz, E. S. (1997). The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. *The Journal of Finance*, 52(3):923–973.
- Wu, J. C. and Xia, F. D. (2016). Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291.