Various state space models

Ilya Zarubin

1 Model

Setting: we consider a panel of observations with a fixed cross sectional unit i at some time point t, and the number of income groups denoted as M_{it} (which may vary over both indices though typically changes only with i i.e. $M_{it} \equiv M_i$):

- $\underline{\boldsymbol{n}}_{it} = \left\{ \underline{n}_{it}^{(k)} \right\}_{k=1}^{M_{it}}$: number of observations in the M_{it} income groups
- $\underline{\boldsymbol{z}}_{it} = \left\{ z_{it}^{(k)} \right\}_{k=1}^{M_{it}-1}$: group boundaries with the last boundary $z_{it}^{M_{it}} = \infty$
- $\underline{\bar{\boldsymbol{y}}}_{it} = \left\{ \bar{y}_{it}^{(k)} \right\}_{k=1}^{M_{it}}$: group mean incomes
- individual incomes are generated by some parametric income distribution $F_y(y; \Theta)$ with P-variate parameter vector $\Theta = (\Theta^{(1)}, \ldots, \Theta^{(P)})$. The latter will be modeled componentwise as a latent stochastic processes over time on the logarithmic scale. More precisely, fixing the cross-sectional unit i, for $t = 1, \ldots, T$, and each component $p = 1, \ldots, P$ we have $\log\left(\Theta_t^{(p)}\right) = x_t^p$
- the latent state transitions for each parameter component $p = 1, \ldots, P$ typically follow some autoregressive process

  $$x_t = \phi x_{t-1} + \boldsymbol{z}_t' \boldsymbol{\beta}_{\boldsymbol{Z}} + \varepsilon_x \ , \ \varepsilon_x \sim \mathcal{N}\left(0, \sigma_X^2\right) \ ,$$

  that includes a set of regressors $\boldsymbol{z}_t'$ that describe various effects of globalization such as trade openess, urbanization, level of education etc.

DGP 1: DGP 1 implies fixed n_{it}'s but random group boundaries z_{it}'s and group means $\bar{y}_{it}$'s. Let each $\underline{\bar{y}}_{it}|\underline{z}_{it}^{(k-1)}, \underline{z}_{it}^{(k)} \overset{approx.}{\sim} \mathcal{N}_{it}(\bar{y}_k|z_{it}^{(k)}\Theta)$. Mean incomes are conditionally independent, given regressors $\boldsymbol{Z}_{it}$. Assuming individual components of the likelihood functions are approximated via the central limit theorem, the joint likelihood function can be approximately written as

$$L\left(\underline{\Theta}; \underline{\bar{\boldsymbol{y}}}_{it}|\boldsymbol{Z}_{it}\right) = \prod_{i=1}^{N}\prod_{t=1}^{T}\prod_{k=2}^{M_{it}-1} \underbrace{\mathcal{N}\left(\bar{y}_{it}^{(k)}|z_{it}^{(k-1)}, z_{it}^{(k)}, \Theta_{it}\right)}_{\text{Part I}} \times \underbrace{f_z\left(z_{it}^{(k)}|z_{it}^{(k-1)}, \Theta_{it}\right)}_{\text{Part II}}$$
$$\times \mathcal{N}\left(\bar{y}_{it}^{(1)}|z_{it}^{(1)}, \Theta_{it}\right) \times \mathcal{N}\left(\bar{y}_{it}^{(M_{it})}|z_{it}^{(M_{it})}, \Theta_{it}\right) \times \underbrace{f_z\left(z_{it}^{(1)}|\Theta_{it}\right)}_{\text{Part IV}} .$$

Part I: $\log\left\{\mathcal{N}\left(\bar{y}_{it}^{(k)}|z_{it}^{(k-1)}, z_{it}^{(k)}, \Theta_{it}\right)\right\}$

$$\begin{aligned}
\log\left\{\mathcal{N}\left(\bar{y}_{it}^{(k)}|z_{it}^{(k-1)}, z_{it}^{(k)}, \Theta_{it}\right)\right\} &= \log\left\{\frac{1}{\sqrt{2\pi\sigma_k^2\left(\Theta_{it}\right)}}\exp\left\{-\frac{1}{2}\frac{\left(\bar{y}_{it}^{(k)} - \mu_k\left(\Theta_{it}\right)\right)^2}{\sigma_k^2\left(\Theta_{it}\right)}\right\}\right\} \\
&= -\frac{1}{2}\log\left(2\pi\sigma_k^2\left(\Theta_{it}\right)\right) - \frac{1}{2}\frac{\left(\bar{y}_{it}^{(k)} - \mu_k\left(\Theta_{it}\right)\right)^2}{\sigma_k^2\left(\Theta_{it}\right)} \\
&= -\frac{1}{2}\left[\log\left(2\pi\sigma_k^2\left(\Theta_{it}\right)\right) + \frac{\left(\bar{y}_{it}^{(k)} - \mu_k\left(\Theta_{it}\right)\right)^2}{\sigma_k^2\left(\Theta_{it}\right)}\right] .
\end{aligned}$$

Part II: $\log\left\{f_z\left(z_{it}^{(k)}|z_{it}^{(k-1)}, \Theta_{it}\right)\right\}$

$$\begin{aligned}
&\log\left\{f_z\left(z_{it}^{(k)}|z_{it}^{(k-1)}, \Theta_{it}\right)\right\} = \\
&+\log\left\{\left(n - n_{k-1}^c\right)!\right\} - \log\left\{\left(n_k^c - n_{k-1}^c - 1\right)!\right\} - \log\left\{(n - n_k^c)!\right\} \\
&+ (n - n_k^c)\log\left\{1 - F_y\left(z_{it}^{(k)}; \Theta_{it}\right)\right\} - \left(n - n_{k-1}^c\right)\log\left\{1 - F_y\left(z_{it}^{(k-1)}; \Theta_{it}\right)\right\} \\
&+ \left(n_k^c - n_{k-1}^c - 1\right)\log\left\{F_y\left(z_{it}^{(k)}; \Theta_{it}\right) - F_y\left(z_{it}^{(k-1)}; \Theta_{it}\right)\right\} \\
&+ \log\left\{f_y\left(z_{it}^{(k)}; \Theta_{it}\right)\right\} .
\end{aligned}$$

Part IV: $\log\left\{f_z\left(z_{it}^{(1)}|\Theta_{it}\right)\right\}$

$$\begin{aligned}
&\log\left\{f_z\left(z_{it}^{(1)}|\Theta_{it}\right)\right\} = \\
&+\log\left\{n!\right\} - \log\left\{(n_1^c - 1)!\,(n - n_1^c)!\right\} \\
&+ (n_1^c - 1)\log\left\{F_y\left(z_{it}^{(1)}; \Theta_{it}\right)\right\} + (n - n_1^c)\log\left\{1 - F_y\left(z_{it}^{(1)}; \Theta_{it}\right)\right\} \\
&+ \log\left\{f_y\left(z_{it}^{(1)}; \Theta_{it}\right)\right\} .
\end{aligned}$$

One example of for an individual income distribution is the four parameter GB2:

$$F_{GBP}(y;a,b,p,q) = B(d;p,q) = \frac{\int_0^d t^{p-1}(1-t)^{q-1}dt}{\mathrm{B}(p,q)} \ , \ d = \frac{(y/b)^a}{1+(y/b)^a} \ ,$$

where each parameter is strictly positive $a, b, p, q > 0$.

DGP 2: DGP 2 implies fixed group boundaries z_{it}'s but random n_{it}'s and $\bar{y}_{it}$'s. We consider only the likelihood related to the number of observations in each group i.e. the classical likelihood framework following McDonald (1984).
Let each $\underline{\boldsymbol{n}}_{it} \overset{ind.}{\sim} \text{MNL}_{it}\left(\pi_{it}^{(1)}, \ldots, \pi_{it}^{(M_{it})}\right)$ conditional on regressors $\boldsymbol{Z}_{it}$ s.th. the joint likelihood function is

$$L\left(\Theta; \underline{\boldsymbol{n}}_{it} | \boldsymbol{Z}_{it}\right) = \prod_{i=1}^{N} \prod_{t=1}^{T} \text{MNL}_{it}(\pi_{it}^{(1)}, \ldots, \pi_{it}^{(M_{it})}) \ .$$

The probability for an individual income to occur in income group $k = 1, \ldots, M_{it}$ and is given as $\pi_{it}^{(k)} = \left(F_y(z_{it}^{(k)}; \Theta) - F_y(z_{it}^{(k-1)}; \Theta)\right)$. One example of for an individual income distribution is the four parameter GB2:

$$F_{GBP}(y; a, b, p, q) = B(d; p, q) = \frac{\int_0^d t^{p-1}(1-t)^{q-1} dt}{\text{B}(p, q)} \ , \ d = \frac{(y/b)^a}{1 + (y/b)^a} \ ,$$

where each parameter is strictly positive $a, b, p, q > 0$.

2 SV model

2.1 Mathematical model

The latent state transition is

$$x_t = \phi x_{t-1} + \boldsymbol{z}_t' \boldsymbol{\beta}_{\boldsymbol{Z}} + \varepsilon_x \ , \ \varepsilon_x \sim \mathcal{N}\left(0, \sigma_X^2\right) \ ,$$

and the measurement equation is

$$y_t = \varepsilon_Y \ , \ \beta_Y \sqrt{\exp\left(x_t\right)} \varepsilon_Y \sim \mathcal{N}\left(0, 1\right) \ .$$

This implies

$$\begin{aligned} p\left(x_t | x_{t-1}\right) &= \mathcal{N}\left(x_t \,\middle|\, \phi x_{t-1} + \boldsymbol{z}_t' \boldsymbol{\beta}_Z, \sigma_X^2\right) \ , t = 1, \ldots, T \ , \\ p\left(y_t | x_t\right) &= \mathcal{N}\left(y_t \,\middle|\, 0, \beta_Y^2 \exp\left(x_t\right)\right) \ , t = 1, \ldots, T \ . \end{aligned}$$

Finally, we need prior assumptions on ε_x and ε_Y:

$$\begin{aligned} \sigma_X^2 &\sim \mathcal{IG}\left(a_X, b_X\right) \ , a_X = b_X = 0.001 \ , \\ \beta_Y^2 &\sim \mathcal{IG}\left(a_Y, b_Y\right) \ , a_Y = b_Y = 0.001 \ . \end{aligned}$$

Note that the model can be written in matrix form as:

$$\begin{aligned} x_{2:T} &= x_{1:T-1} \phi + \boldsymbol{z}_{2:T} \boldsymbol{\beta}_{\boldsymbol{Z}} + \boldsymbol{\varepsilon}_{x,2:T} \ , \\ x_{2:T} &= \boldsymbol{Z}_{2:T} \times \left(\phi, \boldsymbol{\beta}_{\boldsymbol{Z}}'\right)' + \boldsymbol{\varepsilon}_{x,2:T} \ , \end{aligned}$$

where $\boldsymbol{Z}_{2:T}$ is a matrix containing as first column $x_{1:T-1}$ and the remaining K regressors in $\boldsymbol{z}_{2:T}$. This makes it easier to calculate the Gibbs block for ϕ below.

2.2 Gibbs-Part

To derive $p\left(\boldsymbol{\theta}|x_{0:T}, y_{1:T}\right)$, consider the full probabilistic model with $\boldsymbol{\theta} = (\sigma_X^2, \sigma_Y^2, \phi, \boldsymbol{\beta}_{\boldsymbol{Z}})$ as

$$
\begin{aligned}
p\left(\boldsymbol{\theta}, x_{0:T}, y_{1:T}\right) = p\left(y_{1:T}|\boldsymbol{\theta}, x_{0:T}\right) p\left(x_{0:T}, \boldsymbol{\theta}\right) &= \prod_{t=1}^{T} p\left(y_t|x_t, \boldsymbol{\theta}\right) \prod_{t=1}^{T} p\left(x_t|x_{t-1}, \boldsymbol{\theta}\right) p\left(x_0|\boldsymbol{\theta}\right) p\left(\boldsymbol{\theta}\right) \ , \\
&= \frac{1}{\left(2\pi\beta_Y^2 \exp\left(x_t\right)\right)^{T/2}} \prod_{t=1}^{T} \exp\left(-\frac{y_t^2}{2\beta_Y^2 \exp\left(x_t\right)}\right) \\
&\times \frac{1}{\left(2\pi\sigma_X^2\right)^{T/2}} \prod_{t=1}^{T} \exp\left(-\frac{\left(x_t - \phi x_{t-1} - \boldsymbol{z}_t'\boldsymbol{\beta}_{\boldsymbol{Z}}\right)^2}{2\sigma_X^2}\right) \\
&\times \frac{b_X^{a_X}}{\Gamma\left(a_X\right)} \left(\sigma_X^2\right)^{-a_X-1} \exp\left(-\frac{b_X}{\sigma_X^2}\right) \\
&\times \frac{b_Y^{a_Y}}{\Gamma\left(a_Y\right)} \left(\beta_Y^2\right)^{-a_Y-1} \exp\left(-\frac{b_Y}{\beta_Y^2}\right) \ .
\end{aligned}
$$

Then, the conditional parameter distributions are conjugate and given as

$$
\begin{aligned}
p\left(\sigma_X^2|x_{0:T}, y_{1:T}, \sigma_Y^2\right) &= \frac{1}{\left(2\pi\sigma_X^2\right)^{T/2}} \prod_{t=1}^{T} \exp\left(-\frac{\left(x_t - \phi x_{t-1} - \boldsymbol{z}_t'\boldsymbol{\beta}_{\boldsymbol{Z}}\right)^2}{2\sigma_X^2}\right) \\
&\times \frac{b_X^{a_X}}{\Gamma\left(a_X\right)} \left(\sigma_X^2\right)^{-a_X-1} \exp\left(-\frac{b_X}{\sigma_X^2}\right) \\
&\propto \left(\sigma_X^2\right)^{-(a_X+T/2)-1} \times \exp\left(-\frac{1}{\sigma_X^2}\left(b_X + \frac{\sum_{t=1}^{T}\left(x_t - \phi x_{t-1} - \boldsymbol{z}_t'\boldsymbol{\beta}_{\boldsymbol{Z}}\right)^2}{2}\right)\right) \\
p\left(\beta_Y^2|x_{0:T}, y_{1:T}, \sigma_X^2\right) &\propto \left(\beta_Y^2\right)^{-(a_Y+T/2)-1} \times \exp\left(-\frac{1}{\beta_Y^2}\left(b_Y + \frac{\sum_{t=1}^{T}\exp\left(-x_t\right) y_t^2}{2}\right)\right) \ .
\end{aligned}
$$

With e.g. $a_Y = a_X = b_Y = b_X = 0.001$, we have

$$
\begin{aligned}
\sigma_X^2 &\sim \mathcal{IG}\left(a_X^*, b_X^*\right) \ , \ a_X^* = a_X + T/2 \ , \ b_X^* = b_X + \frac{\sum_{t=1}^{T}\left(x_t - \phi x_{t-1} - \boldsymbol{z}_t'\boldsymbol{\beta}_{\boldsymbol{Z}}\right)^2}{2} \\
\sigma_Y^2 &\sim \mathcal{IG}\left(a_Y^*, b_Y^*\right) \ , \ a_Y^* = a_Y + T/2 \ , \ b_Y^* = b_Y + \frac{\sum_{t=1}^{T}\exp\left(-x_t\right) y_t^2}{2} \ .
\end{aligned}
$$

For $\boldsymbol{\beta}_{\boldsymbol{Z}}^{*} = (\phi, \boldsymbol{\beta}_{\boldsymbol{Z}}')'$ with a normal prior $\boldsymbol{\beta}_{\boldsymbol{Z}}^{*} \sim \mathcal{N}_{K+1}\left(\underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}}, \underline{\boldsymbol{\Omega}_{Z}}\right)$ and the previous $x_{2:T} = \boldsymbol{Z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}}^{*} + \boldsymbol{\varepsilon}_{x,2:T}$, we have

$$
\begin{aligned}
p\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^{*} | x_{2:T}, \boldsymbol{Z}_{2:T}, \sigma_X^2, \sigma_Y^2\right) \propto \exp & \left\{-\frac{1}{2}\left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}\right)' \boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x,2:T}}^{-1}\left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}\right)\right\} \\
\times \exp & \left\{-\frac{1}{2}\left(\boldsymbol{\beta}_{\boldsymbol{Z}} - \underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}}\right)' \underline{\boldsymbol{\Omega}_{Z}}^{-1}\left(\boldsymbol{\beta}_{\boldsymbol{Z}} - \underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}}\right)\right\}
\end{aligned}
$$

Because we have

$$
\begin{aligned}
\left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}\right)' \boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x,2:T}}^{-1}\left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}\right) &= \boldsymbol{\beta}_{\boldsymbol{Z}}^{*\prime}\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x,2:T}}^{-1}\boldsymbol{Z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}}^{*} - 2\boldsymbol{\beta}_{\boldsymbol{Z}}^{*\prime}\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x,2:T}}^{-1}x_{2:T} \\
&\quad + x_{2:T}'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x,2:T}}^{-1}x_{2:T} \\
\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^{*} - \underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}}\right)' \underline{\boldsymbol{\Omega}_{Z}}^{-1}\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^{*} - \underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}}\right) &= \boldsymbol{\beta}_{\boldsymbol{Z}}^{*\prime}\underline{\boldsymbol{\Omega}_{Z}}^{-1}\boldsymbol{\beta}_{\boldsymbol{Z}}^{*} - 2\boldsymbol{\beta}_{\boldsymbol{Z}}^{*\prime}\underline{\boldsymbol{\Omega}_{Z}}^{-1}\underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}} \\
&\quad + \underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^{*\prime}} * \underline{\boldsymbol{\Omega}_{Z}}^{-1}\underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^{*}}
\end{aligned}
$$

we obtain

$$
\begin{aligned}
\Rightarrow p\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^{*} | x_{2:T}, \boldsymbol{Z}_{2:T}, \sigma_X^2, \sigma_Y^2\right) &= \mathcal{N}_{K+1}\left(\overline{\boldsymbol{\beta}_{\boldsymbol{Z}}}, \overline{\boldsymbol{\Omega}_{Z}}\right) \\
\overline{\boldsymbol{\Omega}_{Z}} &= \left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x,2:T}}^{-1}\boldsymbol{Z}_{2:T} + \underline{\boldsymbol{\Omega}_{Z}}^{-1}\right]^{-1} \\
\overline{\boldsymbol{\beta}_{\boldsymbol{Z}}} &= \overline{\boldsymbol{\Omega}_{Z}} \times \left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x,2:T}}^{-1}x_{2:T} + \underline{\boldsymbol{\Omega}_{Z}}^{-1}\boldsymbol{\beta}_{\boldsymbol{Z}}\right]
\end{aligned}
$$

In Detail:

In Detail:

3 GZ model

3.1 Mathematical model

The latent state transitions are

$$\begin{aligned}
x_{a,t,i} &= \phi_a x_{a,t-1,i} + \boldsymbol{z}'_{a,t,i}\boldsymbol{\beta}_{\boldsymbol{Z_a}} + \varepsilon_{x_a} \ , \varepsilon_{x_a} \sim \mathcal{N}\left(0, \sigma^2_{x_a}\right) \ , \\
x_{b,t,i} &= \phi_b x_{b,t-1,i} + \boldsymbol{z}'_{b,t,i}\boldsymbol{\beta}_{\boldsymbol{Z_b}} + \varepsilon_{x_b} \ , \varepsilon_{x_b} \sim \mathcal{N}\left(0, \sigma^2_{x_b}\right) \ , \\
x_{p,t,i} &= \phi_p x_{p,t-1,i} + \boldsymbol{z}'_{p,t,i}\boldsymbol{\beta}_{\boldsymbol{Z_p}} + \varepsilon_{x_p} \ , \varepsilon_{x_p} \sim \mathcal{N}\left(0, \sigma^2_{x_p}\right) \ , \\
x_{q,t,i} &= \phi_q x_{q,t-1,i} + \boldsymbol{z}'_{q,t,i}\boldsymbol{\beta}_{\boldsymbol{Z_q}} + \varepsilon_{x_q} \ , \varepsilon_{x_q} \sim \mathcal{N}\left(0, \sigma^2_{x_q}\right) \ .
\end{aligned}$$

The parameters of the state processes can be grouped as

$$\begin{aligned}
\boldsymbol{\theta}_a &= \left(\phi_a, \boldsymbol{\beta}_{\boldsymbol{Z_a}}, \sigma^2_{x_a}\right) \ , \\
\boldsymbol{\theta}_b &= \left(\phi_b, \boldsymbol{\beta}_{\boldsymbol{Z_b}}, \sigma^2_{x_b}\right) \ , \\
\boldsymbol{\theta}_p &= \left(\phi_p, \boldsymbol{\beta}_{\boldsymbol{Z_p}}, \sigma^2_{x_p}\right) \ , \\
\boldsymbol{\theta}_q &= \left(\phi_q, \boldsymbol{\beta}_{\boldsymbol{Z_q}}, \sigma^2_{x_q}\right) \ .
\end{aligned}$$

The measurement equation takes the general, highly nonlinear form, of

$$\boldsymbol{y}_{t,i} = g\left(\exp\left(x_{a,t,i}\right), \exp\left(x_{b,t,i}\right), \exp\left(x_{p,t,i}\right), \exp\left(x_{q,t,i}\right)\right) \ .$$

This implies $\forall\ t = 1, \ldots, T$ and $\forall\ i = 1, \ldots, N$

$$\begin{aligned}
p_{\boldsymbol{\theta}_a}\left(x_{a,t,i}|x_{a,t-1,i}, \boldsymbol{z}'_{a,t,i}\right) &= \mathcal{N}\left(x_{a,t,i} \mid \phi_a x_{a,t-1,i} + \boldsymbol{z}'_{a,t,i}\boldsymbol{\beta}_{\boldsymbol{Z_a}}, \sigma^2_{x_a}\right) \ , \\
p_{\boldsymbol{\theta}_b}\left(x_{b,t,i}|x_{b,t-1,i}, \boldsymbol{z}'_{b,t,i}\right) &= \mathcal{N}\left(x_{b,t,i} \mid \phi_b x_{b,t-1,i} + \boldsymbol{z}'_{b,t,i}\boldsymbol{\beta}_{\boldsymbol{Z_b}}, \sigma^2_{x_b}\right) \ , \\
p_{\boldsymbol{\theta}_p}\left(x_{p,t,i}|x_{p,t-1,i}, \boldsymbol{z}'_{p,t,i}\right) &= \mathcal{N}\left(x_{p,t,i} \mid \phi_p x_{p,t-1,i} + \boldsymbol{z}'_{p,t,i}\boldsymbol{\beta}_{\boldsymbol{Z_p}}, \sigma^2_{x_p}\right) \ , \\
p_{\boldsymbol{\theta}_q}\left(x_{q,t,i}|x_{q,t-1,i}, \boldsymbol{z}'_{q,t,i}\right) &= \mathcal{N}\left(x_{q,t,i} \mid \phi_q x_{q,t-1,i} + \boldsymbol{z}'_{q,t,i}\boldsymbol{\beta}_{\boldsymbol{Z_q}}, \sigma^2_{x_q}\right) \ , \\
p\left(\boldsymbol{y}_{t,i}|x_{a,t,i}, x_{b,t,i}, x_{p,t,i}, x_{q,t,i}\right) &= \mathcal{MNL}\left(\boldsymbol{y}_{t,i}|\pi^{(1)}_{it}, \ldots, \pi^{(M_{it})}_{it}\right) \ ,
\end{aligned}$$

where $\pi^{(k)}_{it} = \left(F_{\text{GB2}}\left(c^{(k)}_{it}; x_{a,t,i}, x_{b,t,i}, x_{p,t,i}, x_{q,t,i}\right) - F_{\text{GB2}}\left(c^{(k-1)}_{it}; x_{a,t,i}, x_{b,t,i}, x_{p,t,i}, x_{q,t,i}\right)\right)$. The income distribution function is a four-parameter GB2:

$$\begin{aligned}
F(c^{(k)}_{it}; x_{a,t,i}, x_{b,t,i}, x_{p,t,i}, x_{q,t,i}) = B(d^{(k)}_{t,i}; x_{p,t,i}, x_{q,t,i}) &= \frac{\int_0^{d^{(k)}_{t,i}} t^{\exp(x_{p,t,i})-1}(1-t)^{\exp(x_{q,t,i})-1}dt}{\mathrm{B}\left(\exp\left(x_{p,t,i}\right), \exp\left(x_{p,t,i}\right)\right)} \ , \\
d^{(k)}_{t,i} &= \frac{(c^{(k)}_{it}/\exp\left(x_{b,t,i}\right))^{\exp(x_{a,t,i})}}{1 + (c^{(k)}_{it}/\exp\left(x_{b,t,i}\right))^{\exp(x_{a,t,i})}} \ .
\end{aligned}$$

Finally, we need prior assumptions on $(\boldsymbol{\theta}_a, \boldsymbol{\theta}_b, \boldsymbol{\theta}_p, \boldsymbol{\theta}_q)$:

$$
\begin{aligned}
\sigma^2_{x_a} &\sim \mathcal{IG}\left(a_{x_a}, b_{x_a}\right) \ , a_{x_a} = b_{x_a} = 0.001 \ , \\
\sigma^2_{x_b} &\sim \mathcal{IG}\left(a_{x_b}, b_{x_b}\right) \ , a_{x_b} = b_{x_b} = 0.001 \ , \\
\sigma^2_{x_p} &\sim \mathcal{IG}\left(a_{x_p}, b_{x_p}\right) \ , a_{x_p} = b_{x_p} = 0.001 \ , \\
\sigma^2_{x_q} &\sim \mathcal{IG}\left(a_{x_q}, b_{x_q}\right) \ , a_{x_q} = b_{x_q} = 0.001 \ , \\
\left(\phi_a, \boldsymbol{\beta}_{\boldsymbol{Z_a}}\right) &\sim \mathcal{N}\left(0, \boldsymbol{I}_a\right) \ , \\
\left(\phi_b, \boldsymbol{\beta}_{\boldsymbol{Z_b}}\right) &\sim \mathcal{N}\left(0, \boldsymbol{I}_b\right) \ , \\
\left(\phi_p, \boldsymbol{\beta}_{\boldsymbol{Z_p}}\right) &\sim \mathcal{N}\left(0, \boldsymbol{I}_p\right) \ , \\
\left(\phi_q, \boldsymbol{\beta}_{\boldsymbol{Z_q}}\right) &\sim \mathcal{N}\left(0, \boldsymbol{I}_q\right) \ ,
\end{aligned}
$$

where $\boldsymbol{I}_a, \boldsymbol{I}_b, \boldsymbol{I}_p, \boldsymbol{I}_q$ are identity matrices of appropriate dimension. Alternatively, instead of $\boldsymbol{I}_a$, one can model the variances of the parameters explicitly i.e. using $\sigma^2_{\boldsymbol{\beta}_{\boldsymbol{Z}_a}, \phi_a} \times \boldsymbol{I}$ with an additional (hierarchical) $\mathcal{IG}\left(,\right)$-prior on $\sigma^2_{\boldsymbol{\beta}_{\boldsymbol{Z}_a}, \phi_a}$ (and similarly for $\boldsymbol{I}_b, \boldsymbol{I}_p, \boldsymbol{I}_q$).
Let $x_{2:T}$ be a generic state proces for some fixed $i = 1, \ldots, N$ i.e. either $\boldsymbol{x}_{a,2:T}, \boldsymbol{x}_{b,2:T}, \boldsymbol{x}_{p,2:T}$ or $\boldsymbol{x}_{q,2:T}$. Note that the model can be written in matrix form as:

$$
\begin{aligned}
x_{2:T} &= x_{1:T-1}\phi + \boldsymbol{z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}} + \boldsymbol{\varepsilon}_{x,2:T} \ , \\
x_{2:T} &= \boldsymbol{Z}_{2:T} \times \left(\phi, \boldsymbol{\beta}'_{\boldsymbol{Z}}\right)' + \boldsymbol{\varepsilon}_{x,2:T} \ ,
\end{aligned}
$$

where $\boldsymbol{Z}_{2:T}$ is a matrix containing as first column $x_{1:T-1}$ and the remaining K'regressors in $\boldsymbol{z}_{2:T}$. This makes it easier to calculate the Gibbs block for ϕ in the next sections.

3.2 Gibbs-Part: univariate (for one $i = 1, \dots, N$)

We derive $p(\boldsymbol{\theta}|x_{0:T}, \boldsymbol{y}_{1:T})$ for a particular $x_t \in \{x_{a,t,i}, x_{b,t,i}, x_{p,t,i}, x_{q,t,i}\}$ i.e. fixing the cross sectional unit i and picking one of the four GB2 parameters. The full probabilistic model with $\boldsymbol{\theta} = (\sigma_X^2, \phi, \boldsymbol{\beta_Z})$ can then be factorized according to

$$
\begin{aligned}
p(\boldsymbol{\theta}, x_{0:T}, \boldsymbol{y}_{1:T}) = p(\boldsymbol{y}_{1:T}|\boldsymbol{\theta}, x_{0:T})\, p(x_{0:T}, \boldsymbol{\theta}) &= \prod_{t=1}^{T} p(\boldsymbol{y}_t|x_t, \boldsymbol{\theta}) \prod_{t=1}^{T} p(x_t|x_{t-1}, \boldsymbol{\theta})\, p(x_0|\boldsymbol{\theta})\, p(\boldsymbol{\theta}) \ , \\
&= \prod_{t=1}^{T} (\boldsymbol{y}_t|\boldsymbol{\theta}, x_t) \\
&\times \frac{1}{(2\pi\sigma_X^2)^{T/2}} \prod_{t=1}^{T} \exp\left(-\frac{(x_t - \phi x_{t-1} - \boldsymbol{z}_t'\boldsymbol{\beta_Z})^2}{2\sigma_X^2}\right) \\
&\times p(\boldsymbol{\theta}) \ .
\end{aligned}
$$

Then, the conditional parameter distributions are conjugate and given as

$$
\begin{aligned}
p\left(\sigma_X^2|x_{0:T}, \boldsymbol{y}_{1:T}\right) &= \frac{1}{(2\pi\sigma_X^2)^{T/2}} \prod_{t=1}^{T} \exp\left(-\frac{(x_t - \phi x_{t-1} - \boldsymbol{z}_t'\boldsymbol{\beta_Z})^2}{2\sigma_X^2}\right) \\
&\times \frac{b_X^{a_X}}{\Gamma(a_X)} \left(\sigma_X^2\right)^{-a_X-1} \exp\left(-\frac{b_X}{\sigma_X^2}\right) \\
&\propto \left(\sigma_X^2\right)^{-(a_X+T/2)-1} \times \exp\left(-\frac{1}{\sigma_X^2}\left(b_X + \frac{\sum_{t=1}^{T}(x_t - \phi x_{t-1} - \boldsymbol{z}_t'\boldsymbol{\beta_Z})^2}{2}\right)\right) \ .
\end{aligned}
$$

With e.g. $a_X = b_X = 0.001$, we have

$$
\sigma_X^2 \sim \mathcal{IG}(a_X^*, b_X^*) \ , \ a_X^* = a_X + T/2 \ , \ b_X^* = b_X + \frac{\sum_{t=1}^{T}(x_t - \phi x_{t-1} - \boldsymbol{z}_t'\boldsymbol{\beta_Z})^2}{2} \ .
$$

For $\boldsymbol{\beta}_{\boldsymbol{Z}}^* = (\phi, \boldsymbol{\beta}_{\boldsymbol{Z}}')'$ with a normal prior $\boldsymbol{\beta}_{\boldsymbol{Z}}^* \sim \mathcal{N}_{K+1}\left(\underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^*}, \underline{\boldsymbol{\Omega}_Z}\right)$ and the previous $x_{2:T} = \boldsymbol{Z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}}^* + \boldsymbol{\varepsilon}_{x,2:T}$, we have

$$
\begin{aligned}
p\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^*|x_{2:T}, \boldsymbol{Z}_{2:T}, \sigma_X^2\right) &\propto \exp\left\{-\frac{1}{2}(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}}^*)' \boldsymbol{\Omega}_{\varepsilon_{x,2:T}}^{-1} (x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}_{\boldsymbol{Z}}^*)\right\} \\
&\times \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta_Z} - \underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^*}\right)' \underline{\boldsymbol{\Omega}_Z}^{-1} \left(\boldsymbol{\beta_Z} - \underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^*}\right)\right\}
\end{aligned}
$$

Because we have

$$
\begin{aligned}
\left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}^*_{\boldsymbol{Z}}\right)' \boldsymbol{\Omega}^{-1}_{\varepsilon_{x,2:T}} \left(x_{2:T} - \boldsymbol{Z}_{2:T}\boldsymbol{\beta}^*_{\boldsymbol{Z}}\right) &= \boldsymbol{\beta}^{*\prime}_{\boldsymbol{Z}}\boldsymbol{Z}'_{2:T}\boldsymbol{\Omega}^{-1}_{\varepsilon_{x,2:T}}\boldsymbol{Z}_{2:T}\boldsymbol{\beta}^*_{\boldsymbol{Z}} - 2\boldsymbol{\beta}^{*\prime}_{\boldsymbol{Z}}\boldsymbol{Z}'_{2:T}\boldsymbol{\Omega}^{-1}_{\varepsilon_{x,2:T}}x_{2:T} \\
&\quad + x'_{2:T}\boldsymbol{\Omega}^{-1}_{\varepsilon_{x,2:T}}x_{2:T} \\
\left(\boldsymbol{\beta}^*_{\boldsymbol{Z}} - \underline{\boldsymbol{\beta}^*_{\boldsymbol{Z}}}\right)' \underline{\boldsymbol{\Omega}_Z}^{-1} \left(\boldsymbol{\beta}^*_{\boldsymbol{Z}} - \underline{\boldsymbol{\beta}^*_{\boldsymbol{Z}}}\right) &= \boldsymbol{\beta}^{*\prime}_{\boldsymbol{Z}}\underline{\boldsymbol{\Omega}_Z}^{-1}\boldsymbol{\beta}^*_{\boldsymbol{Z}} - 2\boldsymbol{\beta}^{*\prime}_{\boldsymbol{Z}}\underline{\boldsymbol{\Omega}_Z}^{-1}\underline{\boldsymbol{\beta}^*_{\boldsymbol{Z}}} \\
&\quad + \underline{\boldsymbol{\beta}^{*\prime}_{\boldsymbol{Z}}} * \underline{\boldsymbol{\Omega}_Z}^{-1}\underline{\boldsymbol{\beta}^*_{\boldsymbol{Z}}}
\end{aligned}
$$

we obtain

$$
\begin{aligned}
p\left(\boldsymbol{\beta}^*_{\boldsymbol{Z}}|x_{2:T}, \boldsymbol{Z}_{2:T}, \sigma^2_X\right) &= \mathcal{N}_{K+1}\left(\overline{\boldsymbol{\beta}_{\boldsymbol{Z}}}, \overline{\boldsymbol{\Omega}_Z}\right) \\
\overline{\boldsymbol{\Omega}_Z} &= \left[\boldsymbol{Z}'_{2:T}\boldsymbol{\Omega}^{-1}_{\varepsilon_{x,2:T}}\boldsymbol{Z}_{2:T} + \underline{\boldsymbol{\Omega}_Z}^{-1}\right]^{-1} \\
\overline{\boldsymbol{\beta}_{\boldsymbol{Z}}} &= \overline{\boldsymbol{\Omega}_Z} \times \left[\boldsymbol{Z}'_{2:T}\boldsymbol{\Omega}^{-1}_{\varepsilon_{x,2:T}}x_{2:T} + \underline{\boldsymbol{\Omega}_Z}^{-1}\boldsymbol{\beta}_{\boldsymbol{Z}}\right]
\end{aligned}
$$

In Detail:

3.3 Gibbs-Part: multivariate (full cross section $\forall i = 1, \ldots, N$)

We now consider vector valued processes stacked along the cross sectional dimension as e.g. $x_{t,1:N}, \boldsymbol{y}_{t,1:N}$. All the corresponding state transition and measurement equations factorize along the time dimension and given as

$$
\begin{aligned}
p_{\boldsymbol{\theta}_a}\left(x_{a,t,1:N} | x_{a,t-1,1:N}, \boldsymbol{z}_{a,t,1:N}\right) &= \mathcal{N}_{1:N}\left(x_{a,t,1:N} \mid \phi_a x_{a,t-1,1:N} + \boldsymbol{z}_{a,t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z_a}}, \sigma^2_{x_a}\boldsymbol{I}_N\right) \ , \\
p_{\boldsymbol{\theta}_b}\left(x_{b,t,1:N} | x_{b,t-1,1:N}, \boldsymbol{z}_{b,t,1:N}\right) &= \mathcal{N}_{1:N}\left(x_{b,t,1:N} \mid \phi_b x_{b,t-1,1:N} + \boldsymbol{z}_{b,t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z_b}}, \sigma^2_{x_b}\boldsymbol{I}_N\right) \ , \\
p_{\boldsymbol{\theta}_p}\left(x_{p,t,1:N} | x_{p,t-1,1:N}, \boldsymbol{z}_{p,t,1:N}\right) &= \mathcal{N}_{1:N}\left(x_{p,t,1:N} \mid \phi_p x_{p,t-1,1:N} + \boldsymbol{z}_{p,t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z_p}}, \sigma^2_{x_p}\boldsymbol{I}_N\right) \ , \\
p_{\boldsymbol{\theta}_q}\left(x_{q,t,1:N} | x_{q,t-1,1:N}, \boldsymbol{z}_{q,t,1:N}\right) &= \mathcal{N}_{1:N}\left(x_{q,t,1:N} \mid \phi_q x_{q,t-1,1:N} + \boldsymbol{z}_{q,t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z_q}}, \sigma^2_{x_q}\boldsymbol{I}_N\right) \ , \\
p\left(\boldsymbol{y}_{t,i} | x_{a,t,i}, x_{b,t,i}, x_{p,t,i}, x_{q,t,i}\right) &= \mathcal{MNL}\left(\boldsymbol{y}_{t,i} | \pi^{(1)}_{it}, \ldots, \pi^{(M_{it})}_{it}\right) \ ,
\end{aligned}
$$

with $\pi^{(k)}_{it} = \left(F_{\text{GB2}}\left(c^{(k)}_{it}; x_{a,t,i}, x_{b,t,i}, x_{p,t,i}, x_{q,t,i}\right) - F_{\text{GB2}}\left(c^{(k-1)}_{it}; x_{a,t,i}, x_{b,t,i}, x_{p,t,i}, x_{q,t,i}\right)\right)$ and income distribution function as a four-parameter GB2

$$
F(c^{(k)}_{it}; x_{a,t,i}, x_{b,t,i}, x_{p,t,i}, x_{q,t,i}) = B(d^{(k)}_{t,i}; x_{p,t,i}, x_{q,t,i}) = \frac{\int_0^{d^{(k)}_{t,i}} t^{\exp(x_{p,t,i})-1}(1-t)^{\exp(x_{q,t,i})-1} dt}{\mathrm{B}\left(\exp\left(x_{p,t,i}\right), \exp\left(x_{p,t,i}\right)\right)} \ ,
$$

$$
d^{(k)}_{t,i} = \frac{(c^{(k)}_{it} / \exp\left(x_{b,t,i}\right))^{\exp(x_{a,t,i})}}{1 + (c^{(k)}_{it} / \exp\left(x_{b,t,i}\right))^{\exp(x_{a,t,i})}} \ .
$$

However, as all state transitions share the same structure, we derive them for a particular $x_{t,1:N} \in \{x_{a,t,1:N}, x_{b,t,1:N}, x_{p,t,1:N}, x_{q,t,1:N}\}$. Now, to obtain $p\left(\boldsymbol{\theta} | x_{0:T,1:N}, \boldsymbol{y}_{0:T,1:N}\right)$, consider the full probabilistic model with $\boldsymbol{\theta} = (\sigma^2_X, \phi, \boldsymbol{\beta}_{\boldsymbol{Z}})$ as

$$
\begin{aligned}
p\left(\boldsymbol{\theta}, x_{0:T,1:N}, \boldsymbol{y}_{0:T,1:N}\right) &= p\left(\boldsymbol{y}_{0:T,1:N} | \boldsymbol{\theta}, x_{0:T,1:N}\right) p\left(x_{0:T,1:N}, \boldsymbol{\theta}\right) \\
&= \prod_{t=1}^{T} p\left(\boldsymbol{y}_{t,1:N} | x_{t,1:N}, \boldsymbol{\theta}\right) \prod_{t=1}^{T} p\left(x_{t,1:N} | x_{t-1,1:N}, \boldsymbol{\theta}\right) p\left(x_{0,1:N} | \boldsymbol{\theta}\right) p\left(\boldsymbol{\theta}\right) \\
&= \prod_{t=1}^{T} p\left(\boldsymbol{y}_{t,1:N} | x_{t,1:N}\right) \times \prod_{t=1}^{T} \frac{1}{\left(2\pi\right)^{N/2}\left(\det\left(\sigma^2_X \boldsymbol{I}_N\right)\right)^{1/2}} \\
&\times \prod_{t=1}^{T} \exp\left(-\frac{1}{2\sigma^2_X}\left(x_{t,1:N} - \phi x_{t-1,1:N} - \boldsymbol{z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}\right)'\left(x_{t,1:N} - \phi x_{t-1,1:N} - \boldsymbol{z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}\right)\right) \\
&\times p\left(\boldsymbol{\theta}\right) \\
&= \prod_{t=1}^{T} p\left(\boldsymbol{y}_{t,1:N} | x_{t,1:N}\right) \times \left(2\pi\sigma^2_X\right)^{-NT/2} \\
&\times \exp\left(-\frac{1}{2\sigma^2_X}\sum_{t=1}^{T}\left(x_{t,1:N} - \boldsymbol{\mu}_{x,t}\right)'\left(x_{t,1:N} - \boldsymbol{\mu}_{x,t}\right)\right) \\
&\times p\left(\boldsymbol{\theta}\right) \ , \ \boldsymbol{\mu}_{x,t} = \phi x_{t-1,1:N} + \boldsymbol{z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}} \ .
\end{aligned}
$$

Then, the conditional parameter distributions are conjugate and given as

$$
\begin{aligned}
p\left(\sigma_X^2|x_{0:T,1:N}\right) = \left(2\pi\sigma_X^2\right)^{-NT/2} \times \exp\left(-\frac{1}{2\sigma_X^2}\sum_{t=1}^{T}\left(x_{t,1:N}-\boldsymbol{\mu}_{x,t}\right)'\left(x_{t,1:N}-\boldsymbol{\mu}_{x,t}\right)\right) \\
\times \frac{\underline{b_X}^{\underline{a_X}}}{\Gamma\left(\underline{a_X}\right)}\left(\sigma_X^2\right)^{-\underline{a_X}-1}\exp\left(-\frac{\underline{b_X}}{\sigma_X^2}\right)
\end{aligned}
$$

$$
\propto \left(\sigma_X^2\right)^{-(\underline{a_X}+NT/2)-1} \times \exp\left(-\frac{1}{\sigma_X^2}\left(\underline{b_X}+\frac{\sum_{t=1}^{T}\left(x_{t,1:N}-\boldsymbol{\mu}_{x,t}\right)'\left(x_{t,1:N}-\boldsymbol{\mu}_{x,t}\right)}{2}\right)\right) .
$$

With e.g. $\underline{a_X} = \underline{b_X} = 0.001$, we have

$$
\sigma_X^2|x_{0:T,1:N} \sim \mathcal{IG}\left(\overline{a_X}, b_X^*\right) \ , \ \overline{a_X} = \underline{a_X} + NT/2 \ , \ \overline{b_X} = \underline{b_X} + \frac{\sum_{t=1}^{T}\left(x_{t,1:N}-\boldsymbol{\mu}_{x,t}\right)'\left(x_{t,1:N}-\boldsymbol{\mu}_{x,t}\right)}{2} .
$$

For $\boldsymbol{\beta}_{\boldsymbol{Z}}^* = (\phi, \boldsymbol{\beta}_{\boldsymbol{Z}}')'$ with a normal prior $\boldsymbol{\beta}_{\boldsymbol{Z}}^* \sim \mathcal{N}_{K+1}\left(\underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^*}, \underline{\boldsymbol{\Omega}_Z}\right)$ and $x_{t,1:N} = \phi x_{t-1,1:N} + \boldsymbol{z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}} + \boldsymbol{\varepsilon}_{x_{t,1:N}} = \boldsymbol{Z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}^* + \boldsymbol{\varepsilon}_{x_{t,1:N}}$, we have

$$
\begin{aligned}
p\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^*|x_{0:T,1:N}, \boldsymbol{Z}_{0:T,1:N}, \sigma_X^2\right) \propto \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\left(x_{t,1:N}-\boldsymbol{Z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}^*\right)'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x_{t,1:N}}}^{-1}\left(x_{t,1:N}-\boldsymbol{Z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}^*\right)\right\} \\
\times \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^*-\underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^*}\right)'\underline{\boldsymbol{\Omega}_Z}^{-1}\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^*-\underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^*}\right)\right\} ,
\end{aligned}
$$

which can as a whole expression be written as

$$
\exp\left\{-\frac{1}{2}\left[\sum_{t=1}^{T}\left(x_{t,1:N}-\boldsymbol{Z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}^*\right)'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x_{t,1:N}}}^{-1}\left(x_{t,1:N}-\boldsymbol{Z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}^*\right)+\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^*-\underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^*}\right)'\underline{\boldsymbol{\Omega}_Z}^{-1}\left(\boldsymbol{\beta}_{\boldsymbol{Z}}^*-\underline{\boldsymbol{\beta}_{\boldsymbol{Z}}^*}\right)\right]\right\}
$$

Because we have for every $t = 1, \dots, T$

$$
\begin{aligned}
\left(x_{t,1:N}-\boldsymbol{Z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}^*\right)'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x_{t,1:N}}}^{-1}\left(x_{t,1:N}-\boldsymbol{Z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}^*\right) = \boldsymbol{\beta}_{\boldsymbol{Z}}^{*\prime}\boldsymbol{Z}_{t,1:N}'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x_{t,1:N}}}^{-1}\boldsymbol{Z}_{t,1:N}\boldsymbol{\beta}_{\boldsymbol{Z}}^* - 2\boldsymbol{\beta}_{\boldsymbol{Z}}^{*\prime}\boldsymbol{Z}_{t,1:N}'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x_{t,1:N}}}^{-1}x_{t,1:N} \\
+ x_{t,1:N}'\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_{x_{t,1:N}}}^{-1}x_{t,1:N}
\end{aligned}
$$

and prior $\left(\boldsymbol{\beta}^*_{\boldsymbol{Z}} - \underline{\boldsymbol{\beta}^*_{\boldsymbol{Z}}}\right)' \underline{\boldsymbol{\Omega}_Z}^{-1} \left(\boldsymbol{\beta}^*_{\boldsymbol{Z}} - \underline{\boldsymbol{\beta}^*_{\boldsymbol{Z}}}\right) = \boldsymbol{\beta}^{*\prime}_{\boldsymbol{Z}} \underline{\boldsymbol{\Omega}_Z}^{-1} \boldsymbol{\beta}^*_{\boldsymbol{Z}} - 2\boldsymbol{\beta}^{*\prime}_{\boldsymbol{Z}} \underline{\boldsymbol{\Omega}_Z}^{-1} \underline{\boldsymbol{\beta}^*_{\boldsymbol{Z}}} + \underline{\boldsymbol{\beta}^{*\prime}_{\boldsymbol{Z}}} * \underline{\boldsymbol{\Omega}_Z}^{-1} \underline{\boldsymbol{\beta}^*_{\boldsymbol{Z}}}$ we obtain

$$
\begin{aligned}
p\left(\boldsymbol{\beta}^*_{\boldsymbol{Z}} | x_{0:T,1:N}, \boldsymbol{Z}_{0:T,1:N}, \sigma^2_X\right) = \mathcal{N}_{K+1}\left(\overline{\boldsymbol{\beta}_{\boldsymbol{Z}}}, \overline{\boldsymbol{\Omega}_Z}\right) \\
\overline{\boldsymbol{\Omega}_Z} &= \left[\sum_{t=1}^{T} \boldsymbol{Z}'_{t,1:N} \boldsymbol{\Omega}^{-1}_{\varepsilon_{x_{t,1:N}}} \boldsymbol{Z}_{t,1:N} + \underline{\boldsymbol{\Omega}_Z}^{-1}\right]^{-1} \\
\overline{\boldsymbol{\beta}_{\boldsymbol{Z}}} &= \overline{\boldsymbol{\Omega}_Z} \times \left[\sum_{t=1}^{T} \boldsymbol{Z}'_{t,1:N} \boldsymbol{\Omega}^{-1}_{\varepsilon_{x_{t,1:N}}} x_{t,1:N} + \underline{\boldsymbol{\Omega}_Z}^{-1} \boldsymbol{\beta}_{\boldsymbol{Z}}\right]
\end{aligned}
$$

In Detail:

In Detail:

ALGORITHM: *Conditional BPF*

START **I. Initiliaze** $(t=0)$:

For $i=1,\ldots,N$:

1. Sample $x_0^i \sim p(x_0)$
3. Set $w_0^i = \frac{1}{N}$
2. Set $x_0^N = x_0^{\mathcal{R}}$ (conditioning)

II. For $t=1$ to T:

For $i=1,\ldots,N$:

1. Draw $a_t^i \sim \mathcal{C}\left(\{w_t^i\}_{j=1}^N\right)$
2. Sample $x_t^i \sim q(x_t|x_{t-1}^{a_t^i},\boldsymbol{\theta}) \underbrace{=}_{\text{BPF}} p(x_t|x_{t-1}^{a_t^i},\boldsymbol{\theta})$

For $i=N$:

3. Set $x_t^N = x_t^{\mathcal{R}}$ (conditioning)
4. Sample $a_t^N \in \{1,\ldots,N\}$ with probability

$$\mathbb{P}\left(a_t^N = i \propto w_{t-1}^i p\left(x_t^{\mathcal{R}}|x_{t-1}^i,\boldsymbol{\theta}\right)\right) , \quad \text{(AS-step)}$$

For $i=1,\ldots,N$:

5. Set $\tilde{w}_t^i = p\left(y_t|x_t^i\right)$
6. Normalize weights $w_t^i = \frac{\tilde{w}_t^i}{\sum_{t=1}^T \tilde{w}_t^i}$

III. For t=T:

1. Draw $b \sim \mathcal{C}\left(\{w_T^i\}_{i=1}^N\right)$ and compute $x_{0:T}^b$

END *Output* $x_{0:T}^{\mathcal{R}} = x_{0:T}^b$

ALGORITHM: *PGAS with conditional BPF*

START **Initiliaze (m=1):**

1. Set $x_{0:T}\left[1\right]$ and $\boldsymbol{\theta}\left[1\right]$ arbitrarily

For $m=1,\ldots,N$:

2. Draw $\boldsymbol{\theta}\left[m\right] \sim p\left(\boldsymbol{\theta}|x_{0:T}\left[m-1\right], x_{0:T}^{\mathcal{R}}\right)$
3. Draw $x_{0:T}\left[m\right] \sim \kappa_{N,\boldsymbol{\theta}[m]}\left(x_{0:T}\left[m-1\right], x_{0:T}^{\mathcal{R}}\right)$

END *Output* $\boldsymbol{\theta}\left[1:m\right]$ and $x_{0:T}\left[1:m\right]$

PEIS: standard SMC constructed from the output of an EIS algorithm:

⇒: replace resampling and propagation scheme by a scheme that favours particles that are more likely to survive the next resampling step

⇒: this is achieved by augmenting standard targets $\pi_T(x_{1:T})$ to include future y_t-measurements → forward looking

⇒: these extended auxiliary densities are then targeted by a SMC based on an EIS proposals and appropriately adjusted weights (that take into account the discrepancy between the new auxiliary targets and the particular EIS proposals)

- The auxiliary targets are:

$$\pi_t(x_{1:t}) \propto \gamma_t(x_{1:t}) \equiv p_{\boldsymbol{\theta}}(x_{1:t}, y_{1:t})\, \chi_{t+1}(x_t; \hat{c}_{t+1}) = p_{\boldsymbol{\theta}}(x_{1:t}, y_{1:t})\, p_{\boldsymbol{\theta}}\left(y_{(t+1):T}|x_t\right) ,$$

  where the last equality holds for the conceptual globally fully adapted PEIS where $\chi_{t+1}(x_t; \hat{c}_{t+1}) = p_{\boldsymbol{\theta}}\left(y_{(t+1):T}|x_t\right)$ together with the corresponding optimal proposals sets the global IS ratio equal to zero. However, these are typically not feasible and an approximation to these optimal proposals and optimal normalizing constant is obtained via auxiliary EIS regressions.

- The proposal is:

$$q_t\left(x_t|x_{1:(t-1)}\right) = q_t\left(x_t|x_{1:(t-1)}; \hat{c}_t\right) = \frac{k_t\left(x_t, x_{t-1}; \hat{c}_t\right)}{\chi_t(x_{t-1}; \hat{c}_t)} .$$

- The resulting weights are:

$$\begin{aligned}
\tilde{w}_t^i &= w_{t-1}^i \frac{\gamma_t(x_{1:t}^i)}{\gamma_{t-1}\left(x_{1:(t-1)}^i\right) q_t\left(x_t^i|x_{1:(t-1)}^i\right)} \\
&= w_{t-1}^i \frac{p_{\boldsymbol{\theta}}(x_{1:t}^i, y_{1:t}^i)\, \chi_{t+1}^i(x_t^i; \hat{c}_{t+1})\, \chi_t^i\left(x_{t-1}^i; \hat{c}_t\right)}{p_{\boldsymbol{\theta}}\left(x_{1:(t-1)}^i, y_{1:(t-1)}^i\right) \chi_t^i(x_{t-1}^i; \hat{c}_t)\, k_t(x_t^i, x_{t-1}^i; \hat{c}_t)} \\
&= w_{t-1}^i \frac{p_{\boldsymbol{\theta}}\left(x_t^i, y_t^i|x_{1:(t-1)}^i, y_{1:(t-1)}\right) \chi_{t+1}^i(x_t^i; \hat{c}_{t+1})}{k_t(x_t^i, x_{t-1}^i; \hat{c}_t)} \\
&= w_{t-1}^i \frac{g_{\boldsymbol{\theta}}(y_t|x_t^i)\, f_{\boldsymbol{\theta}}\left(x_t^i|x_{t-1}^i\right) \chi_{t+1}^i(x_t^i; \hat{c}_{t+1})}{k_t(x_t^i, x_{t-1}^i; \hat{c}_t)} .
\end{aligned}$$

GZ model

Derivation of SMC-EIS weights when resampling is performed in every period and for one exemplary state trajectory x_t e.g. $x_{a,t,n} \equiv x_t$ for some $n = 1, \ldots, N$, $t = 1, \ldots, T$.

$$w_t^i = \frac{g_{\boldsymbol{\theta}}\left(y_t|x_t^i\right) f_{\boldsymbol{\theta}}\left(x_t^i|x_{t-1}^i\right) \chi_{t+1}^i\left(x_t^i; \hat{c}_{t+1}\right)}{k_t\left(x_t^i, x_{t-1}^i; \hat{c}_t\right)} \ .$$

$$\log\left(w_t^i\right) = \log\left(g_{\boldsymbol{\theta}}\left(y_t|x_t^i\right)\right) + \log\left(f_{\boldsymbol{\theta}}\left(x_t^i|x_{t-1}^i\right)\right) + \log\left(\chi_{t+1}^i\left(x_t^i; \hat{c}_{t+1}\right)\right) - \log\left(k_t\left(x_t^i, x_{t-1}^i; \hat{c}_t\right)\right) \ .$$

Deriving each part of the log-weights:

1.

$$g_{\boldsymbol{\theta}}\left(y_t|x_t^i\right) = \frac{K!}{y_{t,1}! \cdot y_{t,K}!} \prod_{k=1}^{K} \pi_k^{y_{t,k}}$$

$$\log\left(g_{\boldsymbol{\theta}}\left(y_t|x_t^i\right)\right) = \log\left(K!\right) - \sum_{k=1}^{K} \log\left(y_{t,k}!\right) + \sum_{k=1}^{K} \log\left(y_{t,k}\right)\log\left(\pi_k\right)$$

$$\log\left(g_{\boldsymbol{\theta}}\left(y_t|x_t^i\right)\right) \propto \sum_{k=1}^{K} \log\left(y_{t,k}\right)\log\left(\pi_k\right) \ .$$

2.

$$f_{\boldsymbol{\theta}}\left(x_t^i|x_{t-1}^i\right) = \frac{1}{\sqrt{2\pi\sigma_{x_a}^2}} \exp\left\{-\frac{1}{2}\frac{\left(x_t^i - \mu_{x_a}\right)^2}{\sigma_{x_a}^2}\right\}$$

$$\log\left(f_{\boldsymbol{\theta}}\left(x_t^i|x_{t-1}^i\right)\right) = -\frac{1}{2}\log\left(2\pi\sigma_{x_a}^2\right) - \frac{1}{2}\frac{\left(x_t^i - \mu_x\right)^2}{\sigma_{x_a}^2} \ ,$$

where $\mu_x = \boldsymbol{z}_a \boldsymbol{\beta}_{z_a}^*$ if e.g. $x_t \equiv x_{t,a} = \phi x_{t-1,a} + z_a \boldsymbol{\beta}_{z_a} + \varepsilon_{x_a}$ and $\boldsymbol{\beta}_{z_a}^* = \left(\phi, \boldsymbol{\beta}_{z_a}\right)$.

3.

$$\chi_{t+1}^i\left(x_t^i; \hat{c}_{t+1}\right) =$$

$$\log\left(\chi_{t+1}^i\left(x_t^i; \hat{c}_{t+1}\right)\right) =$$

4.

$$k_t\left(x_t^i, x_{t-1}^i; \hat{c}_t\right) =$$

$$\log\left(k_t\left(x_t^i, x_{t-1}^i; \hat{c}_t\right)\right) =$$

Derivation of AS weights for PEIS when resampling is performed in every period and for one exemplary state trajectory x_t e.g. $x_{a,t,n} \equiv x_t$ for some $n = 1, \ldots, N$, $t = 1, \ldots, T$.

- In general, the AS-weights are computed as

$$\widehat{w}^i_{t-1|T} \propto w^i_{t-1} \times \frac{\gamma_T\left(x^i_{1:(t-1)}, x^{\mathcal{R}}_{t:T}\right)}{\gamma_{t-1}\left(x^i_{1:(t-1)}\right)} \ .$$

- The globally fully adapted SMC targeted by PEIS uses targets $\pi_t\left(x_{1:t}\right) \propto p_{\boldsymbol{\theta}}\left(x_{1:t}, y_{1:t}\right) p_{\boldsymbol{\theta}}\left(y_{(t+1):T}|x_t\right)$ and $q_t\left(x_t|x_{1:(t-1)}\right) \equiv p_{\boldsymbol{\theta}}\left(x_t|x_{t-1}, y_{t:T}\right)$. In this setting the AS weights become

$$\begin{aligned}\widehat{w}^i_{t-1|T} &\propto 1 \times \frac{\gamma_T\left(x^i_{1:(t-1)}, x^{\mathcal{R}}_{t:T}\right)}{\gamma_{t-1}\left(x^i_{1:(t-1)}\right)} = 1 \times \frac{p_{\boldsymbol{\theta}}\left(y_{t:T}, x^{\mathcal{R}}_{t:T}|x^i_{t-1}\right)}{p_{\boldsymbol{\theta}}\left(y_{t:T}|x^i_{t-1}\right)} \\ &= 1 \times p_{\boldsymbol{\theta}}\left(x^{\mathcal{R}}_t|x^i_{t-1}, y_{t:T}\right) \ .\end{aligned}$$

  Under PEIS, which is designed to provide the best possible approximation to the globally fully adapted SMC, the constant prior weights, $\propto 1$, are replaced by w^i_{t-1} and the optimal EIS density by the PEIS density (which is obtained by solving the least squares optimization problem before propagating particles from $t = 1, \ldots, T$) and given as

$$q_t\left(x_t|x_{1:(t-1)}\right) = q_t\left(x_t|x_{1:(t-1)}; \hat{c}_t\right) = \frac{k_t\left(x_t, x_{t-1}; \hat{c}_t\right)}{\chi_t\left(x_{t-1}; \hat{c}_t\right)} \ .$$