1 Model

Setting: we consider a panel of multinomial observations with cross sectional units $i = 1, \ldots, N$ at time points $t = 1, \ldots, T$:

- $\boldsymbol{y}_{it} = (y_{it1}, \ldots, y_{itD})$: vector of $d = 1, \ldots, D$ time dependent counts to be modelled for each individual unit i
- for each i and t, denote the total number of counts of components of $\boldsymbol{y}_{it}$ as $n_{it} = \sum_{d=1}^{D} y_{itd}$
- The likelihood of all observations given the parameters is

$$L\left(\underline{\boldsymbol{\alpha}}; \boldsymbol{y}_{11}, \ldots, \boldsymbol{y}_{NT}\right) = \prod_{t=1}^{T} \prod_{i=1}^{N} \left\{ \frac{n_{it}!}{y_{it1}! \times y_{it2}! \times \ldots \times y_{itD}!} \prod_{d=1}^{D} \left(y_{itd}\right)^{\exp(x_{itd})} \right\} .$$

- Here, $\underline{\boldsymbol{\alpha}} = (\alpha_{11}, \ldots, \alpha_{1T}, \alpha_{21}, \ldots, \alpha_{2T}, \ldots, \alpha_{N1}, \ldots, \alpha_{NT})$. For each individual unit i, shares $\boldsymbol{\alpha_{it}} = (\alpha_{it1}, \ldots, \alpha_{itD})$ 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 $d = 1, \ldots, D$ we have $\log\left(\alpha_{itd}\right) = x_{itd}$. The loglink ensures the positive scale of the Dirichelet parameters i.e. that $\alpha_{itd} > 0$.
- For each cross sectional unit i, latent state transitions of each parameter component d follows some autoregressive process over time:

$$x_{itd} = \phi_x x_{i(t-1)d} + \left(\boldsymbol{z}_{itd}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z}} + \left(\boldsymbol{z}_{itd}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z}} + \varepsilon_{itd} \ , \ \varepsilon_{itd} \sim \mathcal{N}\left(0, \sigma_{\varepsilon}^{2}\right) \ ,$$

  In general, the parameters $\phi, \boldsymbol{\beta}_{\boldsymbol{Z}}, \boldsymbol{\gamma}_{\boldsymbol{Z}}, \sigma_{\varepsilon}^{2}$ will vary with different α_d's but for notational convenience we have omitted dependence on the index d for the moment. Regressors include $\boldsymbol{z}_{itd}^{lin}$ to model linear effects on the response as well as nonlinear covariate effects $\boldsymbol{z}_{itd}^{spl}$ via polynomial splines.

2 KZ model

2.1 Mathematical model

The latent state transitions are

$$\begin{aligned}
x_{it1} &= \phi_1 x_{i(t-1)1} + \left(\boldsymbol{z}_{it1}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z_1}} + \left(\boldsymbol{z}_{it1}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z_1}} + \varepsilon_1 \ , \varepsilon_1 \sim \mathcal{N}\left(0, \sigma_{\varepsilon_1}^2\right) \ , \\
x_{it2} &= \phi_2 x_{i(t-1)2} + \left(\boldsymbol{z}_{it2}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z_2}} + \left(\boldsymbol{z}_{it2}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z_2}} + \varepsilon_2 \ , \varepsilon_2 \sim \mathcal{N}\left(0, \sigma_{\varepsilon_2}^2\right) \ , \\
&\vdots \\
x_{itD} &= \phi_D x_{i(t-1)D} + \left(\boldsymbol{z}_{itD}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z_D}} + \left(\boldsymbol{z}_{itD}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z_D}} + \varepsilon_D \ , \varepsilon_D \sim \mathcal{N}\left(0, \sigma_{\varepsilon_D}^2\right) \ ,
\end{aligned}$$

where there are D different parameters $\alpha_{it1}, \ldots, \alpha_{itD}$ of the multinomial distribution distribution $\forall i, t$ with link $x_{itd} = \log(\alpha_{itd})$. The parameters of the state processes are grouped as

$$\begin{aligned}
\boldsymbol{\theta}_1 &= \left(\phi_1, \boldsymbol{\beta}_{\boldsymbol{Z_1}}, \boldsymbol{\gamma}_{\boldsymbol{Z_1}}, \sigma_{\varepsilon_1}^2\right) \ , \\
\boldsymbol{\theta}_2 &= \left(\phi_2, \boldsymbol{\beta}_{\boldsymbol{Z_2}}, \boldsymbol{\gamma}_{\boldsymbol{Z_2}}, \sigma_{\varepsilon_2}^2\right) \ , \\
&\vdots \\
\boldsymbol{\theta}_D &= \left(\phi_D, \boldsymbol{\beta}_{\boldsymbol{Z_D}}, \boldsymbol{\gamma}_{\boldsymbol{Z_D}}, \sigma_{\varepsilon_D}^2\right) \ , \\
\underline{\boldsymbol{\theta}} &= \left(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_D\right) \ .
\end{aligned}$$

The measurement equation is given as the Dirichlet density

$$p_{\underline{\boldsymbol{\theta}}}\left(\boldsymbol{y}_{it} | x_{it1}, \ldots, x_{itD}\right) = \frac{n_{it}!}{y_{it1}! \times y_{it2}! \times \ldots \times y_{itD}!} \prod_{d=1}^{D} \left(y_{itd}\right)^{\exp(x_{itd})} \ .$$

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

$$\begin{aligned}
p_{\boldsymbol{\theta}_1}\left(x_{it1} | x_{i(t-1)1}, \boldsymbol{z}_{it1}^{lin}, \boldsymbol{z}_{it1}^{spl}\right) &= \mathcal{N}\left(x_{it1} \mid \mu_{it1}; \sigma_{\varepsilon_1}^2\right) \ , \\
p_{\boldsymbol{\theta}_2}\left(x_{it2} | x_{i(t-1)2}, \boldsymbol{z}_{it2}^{lin}, \boldsymbol{z}_{it2}^{spl}\right) &= \mathcal{N}\left(x_{it2} \mid \mu_{it2}; \sigma_{\varepsilon_2}^2\right) \ , \\
&\vdots \\
p_{\boldsymbol{\theta}_D}\left(x_{itD} | x_{i(t-1)D}, \boldsymbol{z}_{itD}^{lin}, \boldsymbol{z}_{itD}^{spl}\right) &= \mathcal{N}\left(x_{itD} \mid \mu_{itD}; \sigma_{\varepsilon_D}^2\right) \ , \\
p_{\underline{\boldsymbol{\theta}}}\left(\boldsymbol{y}_{it} | x_{it1}, x_{it2}, \ldots, x_{itD}\right) &= \mathbf{Dir}\left(\boldsymbol{y}_{it} | \exp\left(x_{it1}\right), \ldots, \exp\left(x_{itD}\right)\right) \ ,
\end{aligned}$$

where $\mu_{itd} = \phi_d x_{i(t-1)d} + \left(\boldsymbol{z}_{itd}^{lin}\right)' \boldsymbol{\beta}_{\boldsymbol{Z_d}} + \left(\boldsymbol{z}_{itd}^{spl}\right)' \boldsymbol{\gamma}_{\boldsymbol{Z_d}}$ for $d = 1, \ldots, D$.

2.2 Gibbs-Part: univariate (for a single $i = 1, \ldots, N$)

W.l.o.g. we can discard the dependence on α_d and i since the derivations are similar $\forall i, d$. Thus, let $p\left(\boldsymbol{\theta}|x_{0:T}, \boldsymbol{y}_{1:T}\right)$ for some $x_{0:T} \equiv x_{i(0:T)d,\cdot}$. The full probabilistic model with $\boldsymbol{\theta} = (\phi, \boldsymbol{\beta_Z}, \boldsymbol{\gamma_Z}, \sigma_\varepsilon^2)$ and $\mu_t = \phi x_{t-1} + \left(\boldsymbol{z}_t^{lin}\right)' \boldsymbol{\beta_Z} + \left(\boldsymbol{z}_t^{spl}\right)' \boldsymbol{\gamma_Z}$ is factorized as

$$
\begin{aligned}
p\left(\boldsymbol{\theta}, x_{0:T}, \boldsymbol{y}_{1:T}\right) &= p\left(\boldsymbol{y}_{1:T}|\boldsymbol{\theta}, x_{0:T}\right) p\left(x_{0:T}, \boldsymbol{\theta}\right) = \prod_{t=1}^{T} p\left(\boldsymbol{y}_t|x_t\right) \prod_{t=1}^{T} p\left(x_t|x_{t-1}, \boldsymbol{\theta}\right) p\left(x_0|\boldsymbol{\theta}\right) p\left(\boldsymbol{\theta}\right) \\
&= \prod_{t=1}^{T} p\left(\boldsymbol{y}_t|x_t\right) \times \frac{1}{\left(2\pi\sigma_\varepsilon^2\right)^{T/2}} \times \prod_{t=1}^{T} \exp\left(-\frac{\left(x_t - \mu_t\right)^2}{2\sigma_\varepsilon^2}\right) \times p\left(x_0|\boldsymbol{\theta}\right) \times p\left(\boldsymbol{\theta}\right) \ .
\end{aligned}
$$

For $\boldsymbol{\beta} = (\phi, \boldsymbol{\beta}'_{\boldsymbol{Z}}, \boldsymbol{\gamma}'_{\boldsymbol{Z}})'$ we assume a normal prior $\boldsymbol{\beta} \sim \mathcal{N}_{K_\beta+K_\gamma+1}\left(\underline{\boldsymbol{\beta}}, \underline{\boldsymbol{\Omega}_\beta}\right)$ where the dimensions of $\boldsymbol{\beta_Z}$ and $\boldsymbol{\gamma_Z}$ are given as K_β and K_γ, respectively. In our settings, we use a priori independent distributions between components $(\phi, \boldsymbol{\beta}'_{\boldsymbol{Z}}, \boldsymbol{\gamma}'_{\boldsymbol{Z}})'$ in $\boldsymbol{\beta}$ with

$$
\underline{\boldsymbol{\beta}} = (0, 0, \ldots, 0,)_{K_\beta+K_\gamma+1} \ , \quad \underline{\boldsymbol{\Omega}_\beta} = \begin{bmatrix} \underline{\boldsymbol{\Omega}_{lin}} & \boldsymbol{0} \\ \boldsymbol{0} & \underline{\boldsymbol{\Omega}_{spl}} \end{bmatrix} , \underline{\boldsymbol{\Omega}_{lin}} = \begin{bmatrix} \tau_\phi^2 & 0 & \ldots 0 \\ 0 & \tau_{\boldsymbol{\beta_Z}}^2 & \ldots 0 \\ \ldots & \ldots & \ldots \\ 0 & 0 & \tau_{\boldsymbol{\beta_Z}}^2 \end{bmatrix}_{\left(K_\beta+1\right)\times\left(K_\beta+1\right)} .
$$

An uninformative prior at $(\phi, \boldsymbol{\beta_Z})$ would let $\tau_\phi^2, \tau_{\boldsymbol{\beta_Z}}^2 \to \infty$ s.th. $\underline{\boldsymbol{\Omega}_{lin}}^{-1} \to \boldsymbol{0}$, where $\boldsymbol{0}$ is the zero matrix.
For the spline components we either set

$$
\underline{\boldsymbol{\Omega}_{spl}} = \left[\tau_{\boldsymbol{\gamma_Z}}^2 \boldsymbol{K}_1^{-1}\right] \ , \quad \boldsymbol{K}_1 = \boldsymbol{D}_1'\boldsymbol{D}_1 = \begin{bmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{bmatrix} \ , \quad \boldsymbol{D}_1 = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix} ,
$$

or, alternatively, we may use a second order random prior that leads to

$$
\underline{\boldsymbol{\Omega}_{spl}} = \left[\tau_{\boldsymbol{\gamma_Z}}^2 \boldsymbol{K}_2^{-1}\right] \ , \quad \boldsymbol{K}_2 = \boldsymbol{D}_2'\boldsymbol{D}_2 = \begin{bmatrix} 1 & -2 & 1 & & & & \\ -2 & 5 & -4 & 1 & & & \\ 1 & -4 & 6 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 6 & -4 & 1 \\ & & & 1 & -4 & 5 & -2 \\ & & & & 1 & -2 & 1 \end{bmatrix} ,
$$

where $\boldsymbol{D}_2$ is given as

$$\boldsymbol{D}_2 = \begin{bmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{bmatrix} .$$

For $\tau^2_{\boldsymbol{\beta_Z}}$ we assume an $\mathcal{IG}\left(a_\tau, b_\tau\right)$-prior with $a_\tau = b_\tau = 0.001$

$$p\left(\tau^2_{\boldsymbol{\beta_Z}}\right) = \frac{b_\tau^{a_\tau}}{\Gamma\left(a_\tau\right)} \left(\tau^2_{\boldsymbol{\beta_Z}}\right)^{-a_\tau - 1} \exp\left(-\frac{b_\tau}{\tau^2_{\boldsymbol{\beta_Z}}}\right) .$$

Analogously, for σ^2_ε we assume an $\mathcal{IG}\left(a_\tau, b_\sigma\right)$-prior with $a_\sigma = b_\sigma = 0.001$ as well:

$$p\left(\sigma^2_\varepsilon\right) = \frac{b_\varepsilon^{a_\varepsilon}}{\Gamma\left(a_\varepsilon\right)} \left(\sigma^2_\varepsilon\right)^{-a_\varepsilon - 1} \exp\left(-\frac{b_\varepsilon}{\sigma^2_\varepsilon}\right) .$$

The overall prior distribution can be written as

$$\begin{aligned} p(\boldsymbol{\theta}) &= p(\boldsymbol{\beta}) p(\sigma^2_\varepsilon) p(\tau^2_{\boldsymbol{\beta_Z}}) \\ &= (2\pi)^{-\frac{K_{\boldsymbol{\beta}}+1}{2}} \det\left(\underline{\boldsymbol{\Omega}_{lin}}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\phi, \boldsymbol{\beta}'_{\boldsymbol{Z}})' \underline{\boldsymbol{\Omega}_{lin}}^{-1} (\phi, \boldsymbol{\beta}'_{\boldsymbol{Z}})\right) \\ &\times (2\pi\tau^2_{\boldsymbol{\beta_Z}})^{-\frac{K_{\gamma_{\boldsymbol{Z}}}-1}{2}} \exp\left(-\frac{1}{2\tau^2_{\boldsymbol{\beta_Z}}} \sum_{j=2}^{K_{\gamma_{\boldsymbol{Z}}}} \left(\gamma_j - \gamma_{j-1}\right)^2\right) \\ &\times \frac{b_\tau^{a_\tau}}{\Gamma\left(a_\tau\right)} \left(\tau^2_{\boldsymbol{\beta_Z}}\right)^{-a_\tau - 1} \exp\left(-\frac{b_\tau}{\tau^2_{\boldsymbol{\beta_Z}}}\right) \\ &\times \frac{b_\varepsilon^{a_\varepsilon}}{\Gamma\left(a_\varepsilon\right)} \left(\sigma^2_\varepsilon\right)^{-a_\varepsilon - 1} \exp\left(-\frac{b_\varepsilon}{\sigma^2_\varepsilon}\right) \end{aligned}$$

Then, all conditional parameter distributions are conjugate, as will be shown now.

The full conditional for $\tau^2_{\beta_{\boldsymbol{Z}}}$ is obtained as

$$
\begin{aligned}
p\left(\tau^2_{\beta_{\boldsymbol{Z}}}|x_{0:T},\boldsymbol{y}_{1:T},\boldsymbol{\beta},\sigma_\varepsilon\right) &= (2\pi\tau^2_{\beta_{\boldsymbol{Z}}})^{-\frac{K_{\gamma_{\boldsymbol{Z}}}-1}{2}}\exp\left(-\frac{1}{2\tau^2_{\beta_{\boldsymbol{Z}}}}\sum_{j=2}^{K_{\gamma_{\boldsymbol{Z}}}}\left(\gamma_j-\gamma_{j-1}\right)^2\right)\\
&\times\frac{b_\tau^{a_\tau}}{\Gamma\left(a_\tau\right)}\left(\tau^2_{\beta_{\boldsymbol{Z}}}\right)^{-a_\tau-1}\exp\left(-\frac{b_\tau}{\tau^2_{\beta_{\boldsymbol{Z}}}}\right)\\
&= (2\pi\tau^2_{\beta_{\boldsymbol{Z}}})^{-\frac{K_{\gamma_{\boldsymbol{Z}}}-1}{2}}\prod_{j=2}^{K_{\gamma_{\boldsymbol{Z}}}}\exp\left(-\frac{\left(\gamma_j-\gamma_{j-1}\right)^2}{2\tau^2_{\beta_{\boldsymbol{Z}}}}\right)\\
&\times\frac{b_\tau^{a_\tau}}{\Gamma\left(a_\tau\right)}\left(\tau^2_{\beta_{\boldsymbol{Z}}}\right)^{-a_\tau-1}\exp\left(-\frac{b_\tau}{\tau^2_{\beta_{\boldsymbol{Z}}}}\right)\\
&\propto\left(\tau^2_{\beta_{\boldsymbol{Z}}}\right)^{-(a_\tau+(K_{\gamma_{\boldsymbol{Z}}}-1)/2)-1}\times\exp\left(-\frac{1}{\tau^2_{\beta_{\boldsymbol{Z}}}}\left(b_\tau+\frac{\sum_{j=2}^{K_{\gamma_{\boldsymbol{Z}}}}\left(\gamma_j-\gamma_{j-1}\right)^2}{2}\right)\right)\ .
\end{aligned}
$$

With e.g. $a_\sigma=b_\sigma=0.001$, we have

$$
\tau^2_{\beta_{\boldsymbol{Z}}}\sim\mathcal{IG}\left(a_\tau^*,b_\tau^*\right)\ ,\ a_\tau^*=a_\tau+(K_{\gamma_{\boldsymbol{Z}}}-1)/2\ ,\ b_\tau^*=b_\tau+\frac{\sum_{j=2}^{K_{\gamma_{\boldsymbol{Z}}}}\left(\gamma_j-\gamma_{j-1}\right)^2}{2}\ .
$$

The full conditional for σ^2_ε is obtained as

$$
\begin{aligned}
p\left(\sigma^2_\varepsilon|x_{0:T},\boldsymbol{y}_{1:T},\boldsymbol{\beta},\tau_{\beta_{\boldsymbol{Z}}}\right) &= \frac{1}{\left(2\pi\sigma^2_\varepsilon\right)^{T/2}}\prod_{t=1}^{T}\exp\left(-\frac{\left(x_t-\boldsymbol{\mu}_x\right)^2}{2\sigma^2_\varepsilon}\right)\times\frac{b_\varepsilon^{a_\varepsilon}}{\Gamma\left(a_\varepsilon\right)}\left(\sigma^2_\varepsilon\right)^{-a_\varepsilon-1}\exp\left(-\frac{b_\varepsilon}{\sigma^2_\varepsilon}\right)\\
&\propto\left(\sigma^2_\varepsilon\right)^{-(a_\varepsilon+T/2)-1}\times\exp\left(-\frac{1}{\sigma^2_\varepsilon}\left(b_\varepsilon+\frac{\sum_{t=1}^{T}\left(x_t-\boldsymbol{\mu}_x\right)^2}{2}\right)\right)\ ,
\end{aligned}
$$

where $\mu_t=\phi x_{t-1}+\left(\boldsymbol{z}_t^{lin}\right)'\boldsymbol{\beta}_{\boldsymbol{Z}}+\left(\boldsymbol{z}_t^{spl}\right)'\boldsymbol{\gamma}_{\boldsymbol{Z}}$. With e.g. $a_\sigma=b_\sigma=0.001$, we have

$$
\sigma^2_\varepsilon\sim\mathcal{IG}\left(a_\sigma^*,b_\sigma^*\right)\ ,\ a_\sigma^*=a_\sigma+T/2\ ,\ b_\sigma^*=b_\sigma+\frac{\sum_{t=1}^{T}\left(x_t-\boldsymbol{\mu}_x\right)^2}{2}\ .
$$

The full conditional for $\boldsymbol{\beta}$ is obtained easier if one changes to matrix notation. Let

$$x_{2:T} = \boldsymbol{Z}_{2:T}\boldsymbol{\beta} + \varepsilon_{2:T} \ ,$$

$$x_{2:T} = \begin{bmatrix} x_2 \\ \vdots \\ x_T \end{bmatrix} \ , \quad x_{1:T-1} = \begin{bmatrix} x_1 \\ \vdots \\ x_{T-1} \end{bmatrix} \ , \quad \varepsilon_{2:T} = \begin{bmatrix} \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{bmatrix} \ , \quad \mathbb{V}\left(\varepsilon_{2:T}\right) = \boldsymbol{\Omega}_{\varepsilon} = \begin{bmatrix} \sigma_{\varepsilon}^2 & 0 & \dots & 0 \\ 0 & \sigma_{\varepsilon}^2 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ \dots & \dots & \dots & \sigma_{\varepsilon}^2 \end{bmatrix} \ ,$$

$$\boldsymbol{Z}_{2:T} = \left[x_{1:T-1}, \boldsymbol{Z}_{2:T}^{lin}, \boldsymbol{Z}_{2:T}^{spl}\right], \boldsymbol{Z}_{2:T}^{lin} = \begin{bmatrix} z_{21} & z_{22} & \dots & z_{2K_{\boldsymbol{\beta}}} \\ z_{31} & z_{32} & \dots & z_{3K_{\boldsymbol{\beta}}} \\ \vdots & \vdots & \ddots & \vdots \\ z_{T1} & z_{12} & \dots & z_{TK_{\boldsymbol{\beta}}} \end{bmatrix}, \boldsymbol{Z}_{2:T}^{spl} = \begin{bmatrix} B(z_{21})^l & B(z_{22}) & \dots & B(z_{2K_{\boldsymbol{\gamma_Z}}}) \\ B(z_{21})^l & B(z_{22}) & \dots & B(z_{3K_{\boldsymbol{\gamma_Z}}}) \\ \vdots & \vdots & \ddots & \vdots \\ B(z_{T1})^l & B(z_{T2}) & \dots & B(z_{TK_{\boldsymbol{\gamma_Z}}}) \end{bmatrix}$$

To derive the full conditional Gibbs block for $\boldsymbol{\beta}$ we start as follows:

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

Now we have

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

But then $\boldsymbol{\beta}\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\varepsilon}^{-1}\boldsymbol{Z}_{2:T}\boldsymbol{\beta} + \boldsymbol{\beta}\underline{\boldsymbol{\Omega}_{\beta}}^{-1}\boldsymbol{\beta} = \boldsymbol{\beta}\left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\varepsilon}^{-1}\boldsymbol{Z}_{2:T} + \underline{\boldsymbol{\Omega}_{\beta}}^{-1}\right]\boldsymbol{\beta}$ implies

$$\overline{\boldsymbol{\Omega}_{\beta}} = \left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\varepsilon}^{-1}\boldsymbol{Z}_{2:T} + \underline{\boldsymbol{\Omega}_{\beta}}^{-1}\right]^{-1} \ ,$$

and $2\boldsymbol{\beta}\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\varepsilon}^{-1}x_{2:T} - 2\boldsymbol{\beta}\underline{\boldsymbol{\Omega}_{\beta}}^{-1}\underline{\boldsymbol{\beta}} = -2\boldsymbol{\beta}\left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\varepsilon}^{-1}x_{2:T} + \underline{\boldsymbol{\Omega}_{\beta}}^{-1}\underline{\boldsymbol{\beta}}\right]$ implies

$$\overline{\boldsymbol{\beta}} = \overline{\boldsymbol{\Omega}_{\beta}} \times \left[\boldsymbol{Z}_{2:T}'\boldsymbol{\Omega}_{\varepsilon}^{-1}x_{2:T} + \underline{\boldsymbol{\Omega}_{\beta}}^{-1}\boldsymbol{\beta}\right] \ .$$

Finally, we have

$$p\left(\boldsymbol{\beta}|x_{2:T}, \boldsymbol{Z}_{2:T}, \sigma_{\varepsilon}^2, \tau_{\boldsymbol{\beta_Z}}^2\right) = \mathcal{N}_{K+1}\left(\overline{\boldsymbol{\beta}}, \overline{\boldsymbol{\Omega}_{\beta}}\right) \ .$$

2.3 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

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

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

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

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

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

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

Then, the conditional parameter distributions are conjugate and given as

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

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

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

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

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

which can as a whole expression be written as

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

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

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

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

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

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 $\tilde{w}_t^i = p\left(\boldsymbol{y}_{it}|x_t^i\right)$
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 $x_{0:T}[1]$ and $\boldsymbol{\theta}[1]$ arbitrarily

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

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

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