A Computation of the scaled score in the univariate models

A.1 χ^2 density

We compute the scaled score appearing in the dynamic of the log-variance $\lambda_t^{(i)}$ in Eq. (2.13) in the case of the χ^2 density. To simplify the notation, we suppress the subscript i. The conditional log-likelihood is:

$$\log p_{W_1}(x_t; v_t, \nu) = \frac{1}{2}c(\nu) + \left(\frac{\nu}{2} - 1\right)\log(x_t) - \frac{\nu}{2}\log(v_t) - \frac{\nu}{2}\left(\frac{x_t}{v_t}\right) \tag{A.1}$$

where $c(\nu) = \nu\log(\nu/2) - 2\log\Gamma(\nu/2)$. We now prove the following result (recall that $v_t = e^{\lambda_t}$):

Proposition A.1. *For the density in Eq. (A.1), the score* $\nabla_t^{W_1} = \frac{\partial \log p_{W_1}(x_t;\lambda_t,\nu)}{\partial \lambda_t}$ *is given by:*

$$\nabla_t^{W_1} = \frac{\nu}{2e^{\lambda_t}}\left[x_t - e^{\lambda_t}\right] \tag{A.2}$$

Proof.

$$\begin{aligned}
\nabla_t^{W_1} &= \frac{\partial \log p_{W_1}(x_t;\lambda_t,\nu)}{\partial \lambda_t} = \frac{\partial \log p_{W_1}(x_t;\lambda_t,\nu)}{\partial v_t} \times \frac{\partial v_t}{\partial \lambda_t} \\
&= \frac{\nu}{2v_t^2}\left[x_t - v_t\right] \times v_t \\
&= \frac{\nu}{2e^{\lambda_t}}\left[x_t - e^{\lambda_t}\right]
\end{aligned}$$

□

Then, we compute the information quantity:

Proposition A.2. *For the density in Eq. (A.1), the Fisher information* $\mathcal{I}_{t|t-1}^{W_1} = \mathrm{E}_{t|t-1}[\nabla_t^2]$ *is given by:*

$$\mathcal{I}_{t|t-1}^{W_1} = \frac{\nu}{2} \tag{A.3}$$

Proof.

$$\begin{aligned}
\mathcal{I}_{t|t-1}^{W_1} &= E_{t-1}\left[(\nabla_t^{W_1})^2\right] = E_{t-1}\left[\frac{\nu^2}{4e^{2\lambda_t}}(x_t - e^{\lambda_t})^2\right] \\
&= \frac{\nu^2}{4e^{2\lambda_t}}Var\left[x_t\right] =^* \frac{\nu^2}{4e^{2\lambda_t}}\frac{2e^{2\lambda_t}}{\nu} = \frac{\nu}{2}
\end{aligned}$$

where $(^*)$ follows from $Var\left[x\right] = Var\left[(v_t/\nu)k_\nu\right] = 2\frac{(v_t)^2}{\nu}$ if k_ν is distributed as a χ^2_ν. □

Finally, it is immediate to compute the scaled score:

Proposition A.3. *For the density in Eq. (A.1), the scaled score* $s_t^{W_1} = (\mathcal{I}_{t|t-1}^{W_1})^{-1}\nabla_t^{W_1}$ *is given by:*

$$s_t^{W_1} = \frac{1}{e^{\lambda_t}}\left[x_t - e^{\lambda_t}\right] \tag{A.4}$$

A.2 F density

In the case of the univariate F density, the conditional log-likelihood is:

$$\log p_F\left(x_t; v_t, \nu_1, \nu_2\right) = d(\nu_1, \nu_2) - \frac{\nu_1}{2}\log(v_t) + \left(\frac{\nu_1}{2} - 1\right)\log(x_t) - \frac{\nu_1+\nu_2}{2}\log\left(\tilde{w}_t\right) \tag{A.5}$$

where:

$$\tilde{w}_t = 1 + \frac{\nu_1 x_t}{(\nu_2 - 2)v_t} \tag{A.6}$$

$$d(\nu_1, \nu_2) = \frac{\nu_1}{2}\log\left(\frac{\nu_1}{\nu_2 - 2}\right) + \log\Gamma\left(\frac{\nu_1+\nu_2}{2}\right) - \log\Gamma\left(\frac{\nu_1}{2}\right) - \log\Gamma\left(\frac{\nu_2}{2}\right) \tag{A.7}$$

We compute now the score of the conditional log-likelihood.

Proposition A.4. *For the density in Eq. (A.5), the score* $\nabla_t^F = \frac{\partial \log p_F(x_t; v_t, \nu_1, \nu_2)}{\partial \lambda_t}$ *is given by:*

$$\nabla_t^F = \frac{\nu_1}{2e^{\lambda_t}}\left[\frac{\nu_1+\nu_2}{\nu_2 - 2}\frac{x_t}{\tilde{w}_t} - e^{\lambda_t}\right] \tag{A.8}$$

Proof.

$$\begin{aligned}
\nabla_t^F &= \frac{\partial \log p_F\left(x_t; v_t, \nu_1, \nu_2\right)}{\partial \lambda_t} = \frac{\partial \log p_F\left(x_t; v_t, \nu_1, \nu_2\right)}{\partial v_t} \times \frac{\partial v_t}{\partial \lambda_t} \\
&= -\frac{\nu_1}{2v_t} + \frac{\nu_1+\nu_2}{2}\left[\tilde{w}_t \frac{\nu_1 x_t}{(\nu_2-2)(v_t)^2}\right] \times v_t \\
&= \frac{\nu_1}{2e^{2\lambda_t}}\left[\frac{\nu_1+\nu_2}{\nu_2-2}\frac{x_t}{\tilde{w}_t} - e^{\lambda_t}\right] \times e^{\lambda_t} \\
&= \frac{\nu_1}{2e^{\lambda_t}}\left[\frac{\nu_1+\nu_2}{\nu_2-2}\frac{x_t}{\tilde{w}_t} - e^{\lambda_t}\right]
\end{aligned}$$

□

As in the correlation model (cf. Section 2.2.2), we scale the score by the inverse of the Fisher information of the χ^2 density. Thus, we get:

Proposition A.5. *For the density in Eq. (A.5), the scaled score* $s_t^F = (\mathcal{I}_{t|t-1}^{W_1})^{-1}\nabla_t^F$ *is given by:*

$$s_t^F = \frac{1}{e^{\lambda_t}}\left[\frac{\nu_1+\nu_2}{\nu_2-2}\frac{x_t}{\tilde{w}_t} - e^{\lambda_t}\right] \tag{A.9}$$

B Proposition 2.1

Proof. We need to compute $\frac{\partial l(X_t)}{\partial f_t'} = \frac{\partial l(X_t)}{\partial \mathsf{vech}(Q_t)'}$, where l is given by Eq. (2.17)

$$l\left(X_t\right) = \frac{1}{2}d_X(k,\nu) + \frac{\nu - k - 1}{2}\log|X_t| - \frac{\nu}{2}\log|V_t| - \frac{\nu}{2}tr\left(V_t^{-1}X_t\right).$$

Thanks to the chain rule, we can split our equation as:

$$\frac{\partial l\left(X_t\right)}{\partial \mathsf{vech}\left(Q_t\right)'} = \frac{\partial l\left(X_t\right)}{\partial \mathsf{vec}\left(V_t\right)'} \frac{\partial \mathsf{vec}\left(V_t\right)}{\partial \mathsf{vec}\left(R_t\right)'} \frac{\partial \mathsf{vec}\left(R_t\right)}{\partial \mathsf{vec}\left(Q_t\right)'} \frac{\partial \mathsf{vec}\left(Q_t\right)}{\partial \mathsf{vech}\left(Q_t\right)'}.$$

Then, starting with the first term and considering $d \log |X| = tr\left(X^{-1}\right) dX$ and $d\left(X^{-1}\right) = -X^{-1}\left(dX\right) X^{-1}$, see Magnus and Neudecker (1999),

$$\begin{aligned}\frac{\partial l\left(X_t\right)}{\partial \mathsf{vec}\left(V_t\right)'} &= -\frac{\nu}{2}\left[\mathsf{vec}\left(V_t^{-1}\right)' - \mathsf{vec}\left(X_t\right)'\left(V_t^{-1} \otimes V_t^{-1}\right)\right] \\ &= \frac{\nu}{2}\left[\mathsf{vec}\left(X_t\right) - \mathsf{vec}\left(V_t\right)\right]'\left(V_t^{-1} \otimes V_t^{-1}\right).\end{aligned}$$

The second term is, thanks to the fact that $\mathsf{vec}\left(AXB\right) = \left(B' \otimes A\right) \mathsf{vec}\left(X\right)$

$$\frac{\partial \mathsf{vec}\left(V_t\right)}{\partial \mathsf{vec}\left(R_t\right)'} = \frac{\partial \mathsf{vec}\left(D_t R_t D_t\right)}{\partial \mathsf{vec}\left(R_t\right)'} = \left(D_t \otimes D_t\right).$$

By definition of duplication matrix, see Abadir and Magnus (2005), we have that

$$\frac{\partial \mathsf{vec}\left(Q_t\right)}{\partial \mathsf{vech}\left(Q_t\right)'} = \mathcal{D}_k.$$

The third term is a little bit more complicated, indeed defining $\Delta_t = \left(diag(Q_t)^{1/2}\right)$,

$$\begin{aligned}d\mathsf{vec}\left(R_t\right) =& d\mathsf{vec}\left(\Delta_t^{-1} Q_t \Delta_t^{-1}\right) \\ =& \Delta_t^{-1} \otimes \Delta_t^{-1} d\mathsf{vec}\left(Q_t\right) + \mathsf{vec}\left(d\left(\Delta_t^{-1}\right) Q_t \Delta_t^{-1}\right) + \mathsf{vec}\left(\Delta_t^{-1} Q_t d\left(\Delta_t^{-1}\right)\right) \\ =& \Delta_t^{-1} \otimes \Delta_t^{-1} d\mathsf{vec}\left(Q_t\right) + \left[\left(\Delta_t^{-1} Q_t \otimes I\right) + \left(I \otimes \Delta_t^{-1} Q_t\right)\right] d\mathsf{vec}\left(\Delta_t^{-1}\right) \\ =& \Delta_t^{-1} \otimes \Delta_t^{-1} d\mathsf{vec}\left(Q_t\right) - \left[\left(\Delta_t^{-1} Q_t \otimes I\right) + \left(I \otimes \Delta_t^{-1} Q_t\right)\right] \Delta_t^{-1} \otimes \Delta_t^{-1} d\mathsf{vec}\left(\Delta_t\right) \\ =& \Delta_t^{-1} \otimes \Delta_t^{-1} d\mathsf{vec}\left(Q_t\right) - \left[\left(\Delta_t^{-1} Q_t \otimes I\right) + \left(I \otimes \Delta_t^{-1} Q_t\right)\right] \Delta_t^{-1} \otimes \Delta_t^{-1} W_Q d\mathsf{vec}\left(Q_t\right),\end{aligned}$$

where $q_t = \mathsf{vec}\left(\Delta_t\right)$ and W_Q is a diagonal matrix with its ith diagonal elements equal to $1/2\sqrt{q_t^{(i)}}$ if $q_t^{(i)} \neq 0$ and zero otherwise.

Given these four results combined with the fact that $\left(D_t \otimes D_t\right)$ and $\left(V_t^{-1} \otimes V_t^{-1}\right)$ are symmetric , we get

$$\begin{aligned}\frac{\partial l\left(X_t\right)}{\partial \mathsf{vec}\left(V_t\right)'} =& \frac{\nu}{2}\left[\mathsf{vec}\left(X_t\right) - \mathsf{vec}\left(V_t\right)\right]'\left(V_t^{-1} \otimes V_t^{-1}\right)\left(D_t \otimes D_t\right)\left(\Delta_t^{-1} \otimes \Delta_t^{-1}\right) \\ & \times \left[I - \left(Q_t \Delta_t^{-1} \otimes I\right) + \left(I \otimes Q_t \Delta_t^{-1}\right)\right] W_Q \mathcal{D}_k \\ =& \frac{\nu}{2}\left[\mathsf{vec}\left(X_t\right) - \mathsf{vec}\left(V_t\right)\right]'\left(D_t^{-1} \Delta_t Q_t^{-1} \otimes D_t^{-1} \Delta_t Q_t^{-1}\right) \Psi_t \mathcal{D}_k,\end{aligned}$$

where $\Psi_t = \left[I - \left[\left(\Delta_t^{-1} Q_t \otimes I\right) + \left(I \otimes \Delta_t^{-1} Q_t\right)\right] W_Q\right]$. □

C Proposition 2.2

Proof. Starting with the definition of Fisher information matrix:

$$\begin{aligned}
E\left[\nabla_t^W \nabla_t^{W\prime} | \mathcal{F}_{t-1}\right] =& E\Big[\frac{\nu^2}{4}\mathcal{D}_k^\prime \Psi_t^\prime \left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right)\left[\mathsf{vec}\left(X_t\right) - \mathsf{vec}\left(V_t\right)\right] \\
& \times \left[\mathsf{vec}(X_t) - \mathsf{vec}(V_t)\right]^\prime \left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right)\Psi_t \mathcal{D}_k | \mathcal{F}_{t-1}\Big] \\
=& \frac{\nu^2}{4}\mathcal{D}_k^\prime \Psi_t^\prime \left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right) \times \\
& \times Var\left(\mathsf{vec}(X_t) - \mathsf{vec}(V_t)|\mathcal{F}_{t-1}\right)\left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right)\Psi_t \mathcal{D}_k \\
=^* & \frac{\nu^2}{4}\mathcal{D}_k^\prime \Psi_t^\prime \left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right)\frac{1}{\nu}\left(I_{k^2} + K_k\right) \times \\
& \times \left(V_t \otimes V_t\right)\left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right)\Psi_t \mathcal{D}_k \\
=^{**} & \frac{\nu}{2}\mathcal{D}_k^\prime \Psi_t^\prime \left(D_t^{-1}\Delta_t Q_t^{-1} D_t \Delta_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1} D_t \Delta_t^{-1}\right)\mathcal{D}_k \mathcal{D}_k^+ \Psi_t \mathcal{D}_k
\end{aligned}$$

where * is thank to the vech formulation of the Wishart variance in Abadir and Magnus (2005) and ** is given by $2\mathcal{D}_k\mathcal{D}_k^+ = (I_k + K_k)$ (recall that $V_t = \left(D_t\Delta_t^{-1}Q_t\Delta_t^{-1}D_t\right))$

$$\begin{aligned}
A &= \left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right)\left(V_t \otimes V_t\right)\left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right) \\
&= \left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right)\left(D_t\Delta_t^{-1}Q_t\Delta_t^{-1}D_t D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t\Delta_t^{-1}Q_t\Delta_t^{-1}D_t D_t^{-1}\Delta_t Q_t^{-1}\right) \\
&= \left(D_t^{-1}\Delta_t Q_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1}\right)\left(D_t\Delta_t^{-1} \otimes D_t\Delta_t^{-1}\right) \\
&= \left(D_t^{-1}\Delta_t Q_t^{-1} D_t\Delta_t^{-1} \otimes D_t^{-1}\Delta_t Q_t^{-1} D_t\Delta_t^{-1}\right) \\
&= \left(H_t^{-1}Q_t^{-1}H_t \otimes H_t^{-1}Q_t^{-1}H_t\right)
\end{aligned}$$

□

D Proposition 2.3

Proof. Multiplying the inverse of 2.19 with the 2.18 we get

$$\begin{aligned}
s_t^W &= \left(\mathcal{D}_k{}^\prime \Psi_t^\prime \left(H_t^{-1}Q_t^{-1}H_t \otimes H_t^{-1}Q_t^{-1}H_t\right)\mathcal{D}_k\mathcal{D}_k{}^+\Psi_t\mathcal{D}_k\right)^{-1}\mathcal{D}_k^\prime\Psi_t^\prime \\
&\quad \times \left(H_t^{-1}Q_t^{-1} \otimes H_t^{-1}Q_t^{-1}\right)\left[\mathsf{vec}\left(X_t\right) - \mathsf{vec}\left(V_t\right)\right] \\
&= \left(\mathcal{D}_k{}^\prime \Psi_t^\prime \left(H_t^{-1}Q_t^{-1}H_t \otimes H_t^{-1}Q_t^{-1}H_t\right)\mathcal{D}_k\mathcal{D}_k{}^+\Psi_t\mathcal{D}_k\right)^{-1}\mathcal{D}_k^\prime\Psi_t^\prime \\
&\quad \times \left(H_t^{-1}Q_t^{-1}H_t \otimes H_t^{-1}Q_t^{-1}H_t\right)\left(H_t^{-1} \otimes H_t^{-1}\right)\left[\mathsf{vec}\left(X_t\right) - \mathsf{vec}\left(V_t\right)\right] \\
&= \left(\mathcal{D}_k{}^\prime \Psi_t^\prime \left(H_t^{-1}Q_t^{-1}H_t \otimes H_t^{-1}Q_t^{-1}H_t\right)\mathcal{D}_k\mathcal{D}_k{}^+\Psi_t\mathcal{D}_k\right)^{-1}\mathcal{D}_k^\prime\Psi_t^\prime \\
&\quad \times \left(H_t^{-1}Q_t^{-1}H_t \otimes H_t^{-1}Q_t^{-1}H_t\right)\left(H_t^{-1} \otimes H_t^{-1}\right)\mathcal{D}_k\mathcal{D}_k{}^+\left[\mathsf{vec}\left(X_t\right) - \mathsf{vec}\left(V_t\right)\right] \\
&= \left(\mathcal{D}_k{}^\prime \Psi_t^\prime \left(H_t^{-1}Q_t^{-1}H_t \otimes H_t^{-1}Q_t^{-1}H_t\right)\mathcal{D}_k\mathcal{D}_k{}^+\Psi_t\mathcal{D}_k\right)^{-1}\mathcal{D}_k^\prime\Psi_t^\prime \\
&\quad \times \left(H_t^{-1}Q_t^{-1}H_t \otimes H_t^{-1}Q_t^{-1}H_t\right)\mathcal{D}_k\mathcal{D}_k{}^+\left(H_t^{-1} \otimes H_t^{-1}\right)\left[\mathsf{vec}\left(X_t\right) - \mathsf{vec}\left(V_t\right)\right]
\end{aligned}$$

Now let us simplify $\mathcal{I}_t^W = \frac{\nu}{2}\mathcal{D}_k'\Psi_t'\left(H_t^{-1}Q_t^{-1}H_t \otimes H_t^{-1}Q_t^{-1}H_t\right)\mathcal{D}_k$. This simplification is not restrictive because the matrix Ψ_t is very sparse, moreover using the approximation $\Psi_t = I$ allows us to define the inverse of $\mathcal{I}_t$ which is not full-rank with the original representation.
We get

$$\begin{aligned} s_t =& \mathcal{D}_k{}^+\left(H_t^{-1}\otimes H_t^{-1}\right)\left[\mathsf{vec}\left(X_t\right) - \mathsf{vec}\left(V_t\right)\right] \\ =& \mathcal{D}_k{}^+\left(H_t^{-1}\otimes H_t^{-1}\right)\mathsf{vec}(X_t) - \mathcal{D}_k{}^+\left(H_t^{-1}\otimes H_t^{-1}\right)\left(H_t\otimes H_t\right)\mathsf{vec}(Q_t) \\ =& \mathcal{D}_k{}^+\left(H_t^{-1}\otimes H_t^{-1}\right)\mathsf{vec}(X_t) - \mathsf{vech}\left(Q_t\right) \end{aligned}$$

□

E Proposition 2.4

Proof. As in the Wishart case, we can split

$$\frac{\partial l\left(X_t\right)}{\partial \mathsf{vech}\left(Q_t\right)'} = \frac{\partial l\left(X_t\right)}{\partial \mathsf{vec}\left(V_t\right)'}\frac{\partial \mathsf{vec}\left(V_t\right)}{\partial \mathsf{vec}\left(R_t\right)'}\frac{\partial \mathsf{vec}\left(R_t\right)}{\partial \mathsf{vec}\left(Q_t\right)'}\frac{\partial \mathsf{vec}\left(Q_t\right)}{\partial \mathsf{vech}\left(Q_t\right)'}.$$

The only terms which is different from the previous model is the first one. The log-likelihood function is and get, using again $d\log|X| = tr\left(X^{-1}\right)dX$ and $d\left(X^{-1}\right) = -X^{-1}\left(dX\right)X^{-1}$

$$\begin{aligned} dl_{X_t} =& -\frac{\nu_1}{2}tr\left(V_t^{-1}dV_t\right) - \frac{\nu_1+\nu_2}{2}tr\left(\tilde{W}_t^{-1}d\tilde{W}_t\right) \\ =^* & -\frac{\nu_1}{2}\left(\mathsf{vec}V_t^{-1}\right)'d\mathsf{vec}V_t + \frac{\nu_1+\nu_2}{2}tr\left(\tilde{W}_t^{-1}\frac{\nu_1}{\nu_2-k-1}V_t^{-1}dV_tV_t^{-1}X_t\right) \\ =& -\frac{\nu_1}{2}\left(\mathsf{vec}V_t^{-1}\right)'d\mathsf{vec}V_t + \frac{\nu_1+\nu_2}{2}tr\left(\frac{\nu_1}{\nu_2-k-1}V_t^{-1}X_t\tilde{W}_t^{-1}V_t^{-1}dV_t\right) \\ =& -\frac{\nu_1}{2}\left(\mathsf{vec}V_t^{-1}\right)'d\mathsf{vec}V_t + \frac{\nu_1+\nu_2}{2}\mathsf{vec}\left(\frac{\nu_1}{\nu_2-k-1}V_t^{-1}X_t\tilde{W}_t^{-1}V_t^{-1}\right)'d\mathsf{vec}V_t \end{aligned}$$

hence we obtain

$$\begin{aligned} \frac{\partial l\left(X_t\right)}{\partial \mathsf{vec}\left(V_t\right)'} =& -\frac{\nu_1}{2}\left(\mathsf{vec}V_t^{-1}\right)' + \frac{\nu_1+\nu_2}{2}\mathsf{vec}\left(\frac{\nu_1}{\nu_2-k-1}V_t^{-1}X_t\tilde{W}_t^{-1}V_t^{-1}\right)' \\ =& \frac{\nu_1}{2}\left[\left(V_t^{-1}\otimes V_t^{-1}\right)\left(\frac{\nu_1+\nu_2}{\nu_2-k-1}\mathsf{vec}\left(X_t\tilde{W}_t^{-1}\right) - \mathsf{vec}\left(V_t\right)\right)\right]'. \end{aligned}$$

Combining all the formulas together we get

$$
\begin{aligned}
\frac{\partial l\left(X_t\right)}{\partial \mathsf{vech}\left(Q_t\right)'} =& \frac{\nu_1}{2}\left(\frac{\nu_1+\nu_2}{\nu_2-k-1}\mathsf{vec}\left(X_t\tilde{W}_t^{-1}\right)-\mathsf{vec}\left(V_t\right)\right)'\left(V_t^{-1}\otimes V_t^{-1}\right)\left(D_t\otimes D_t\right) \\
&\times\left(\Delta_t^{-1}\otimes\Delta_t^{-1}\right)\Psi_t\mathcal{D}_k \\
=& \frac{\nu_1}{2}\left(\frac{\nu_1+\nu_2}{\nu_2-k-1}\mathsf{vec}\left(X_t\tilde{W}_t^{-1}\right)-\mathsf{vec}\left(V_t\right)\right)' \\
&\times\left(D_t^{-1}\Delta_tQ_t^{-1}\Delta_tD_t^{-1}\otimes D_t^{-1}\Delta_tQ_t^{-1}\Delta_tD_t^{-1}\right)\left(D_t\Delta_t^{-1}\otimes D_t\Delta_t^{-1}\right)\Psi_t'\mathcal{D}_k' \\
=& \frac{\nu_1}{2}\left(\frac{\nu_1+\nu_2}{\nu_2-k-1}\mathsf{vec}\left(X_t\tilde{W}_t^{-1}\right)-\mathsf{vec}\left(V_t\right)\right)'\left(H_t^{-1}Q_t^{-1}\otimes H_t^{-1}Q_t^{-1}\right)
\end{aligned}
$$

□

F Proposition 2.5

Proof. Consider the multiplication $\mathcal{I}_t^W\nabla_t^F$, considering the same approximation of $\mathcal{I}_t^W$ we used in D,

$$
\begin{aligned}
s_t^F =& \left(\mathcal{D}_k'\left(H_t^{-1}Q_t^{-1}H_t\otimes H_t^{-1}Q_t^{-1}H_t\right)\mathcal{D}_k\right)^{-1}\mathcal{D}_k' \\
&\times\left(H_t^{-1}Q_t^{-1}\otimes H_t^{-1}Q_t^{-1}\right)\left[\frac{\nu_1+\nu_2}{\nu_2-k-1}\mathsf{vec}\left(X_t\tilde{W}_t^{-1}\right)-\mathsf{vec}\left(V_t\right)\right] \\
=& \left(\mathcal{D}_k'\left(H_t^{-1}Q_t^{-1}H_t\otimes H_t^{-1}Q_t^{-1}H_t\right)\mathcal{D}_k\right)^{-1}\mathcal{D}_k'\left(H_t^{-1}Q_t^{-1}H_t\otimes H_t^{-1}Q_t^{-1}H_t\right) \\
&\times\mathcal{D}_k\mathcal{D}_k^{+}\left(H_t^{-1}\otimes H_t^{-1}\right)\left[\frac{\nu_1+\nu_2}{\nu_2-k-1}\mathsf{vec}\left(X_t\tilde{W}_t^{-1}\right)-\mathsf{vec}\left(V_t\right)\right] \\
=& \frac{\nu_1+\nu_2}{\nu_2-k-1}\mathsf{vech}\left(H_t^{-1}X_t\tilde{W}_t^{-1}H_t^{-1}\right)-\mathsf{vech}\left(Q_t\right)
\end{aligned}
$$

□

G Proposition 2.6

Proof. The proof is straightforward:

$$
\begin{aligned}
\lim_{\nu_2\to\infty}s_t^F =& \lim_{\nu_2\to\infty}\frac{\nu_1+\nu_2}{\nu_2-k-1}\mathsf{vech}\left(H_t^{-1}X_t\tilde{W}_t^{-1}H_t^{-1}\right)-\mathsf{vech}\left(Q_t\right) \\
=& \mathsf{vech}\left(H_t^{-1}X_t\tilde{W}_t^{-1}H_t^{-1}\right)-\mathsf{vech}\left(V_t\right)=s_t^W
\end{aligned}
$$

since $\tilde{W}_t = I_k + \frac{\nu_1}{\nu_2-k-1}V_t^{-1}X_t$ □

H Figures of Section (3.2)

In this section, we report the figures of the experiment based on the misspecified DGP's presented in Section (3.2)

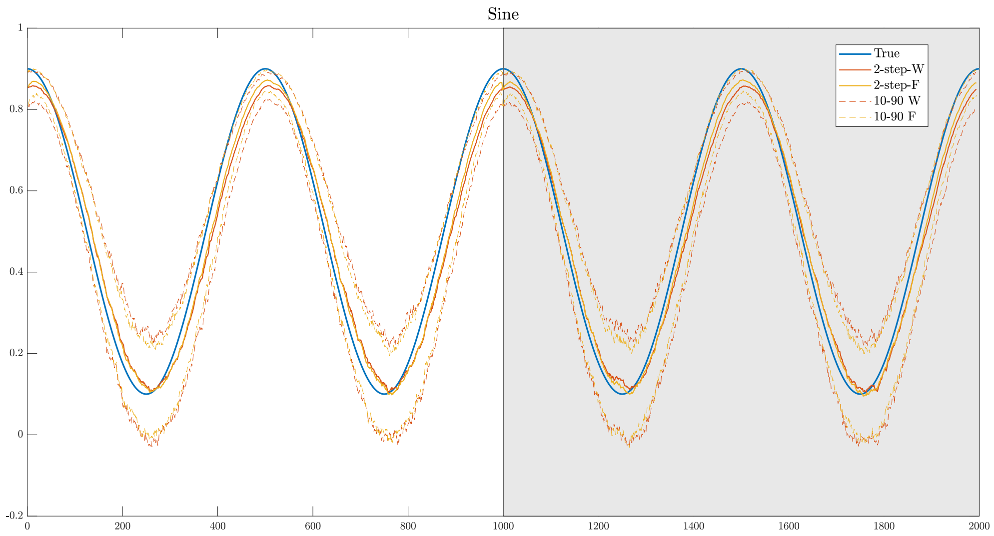

Figure 8: In-sample and out-of-sample filtered estimates of $\rho_t^{(1)}$ from both the 2-step-W and 2-step-F models. Estimates are averaged over the 1000 simulations. Confidence bands are constructed by computing the 10% and 90% empirical quantiles.

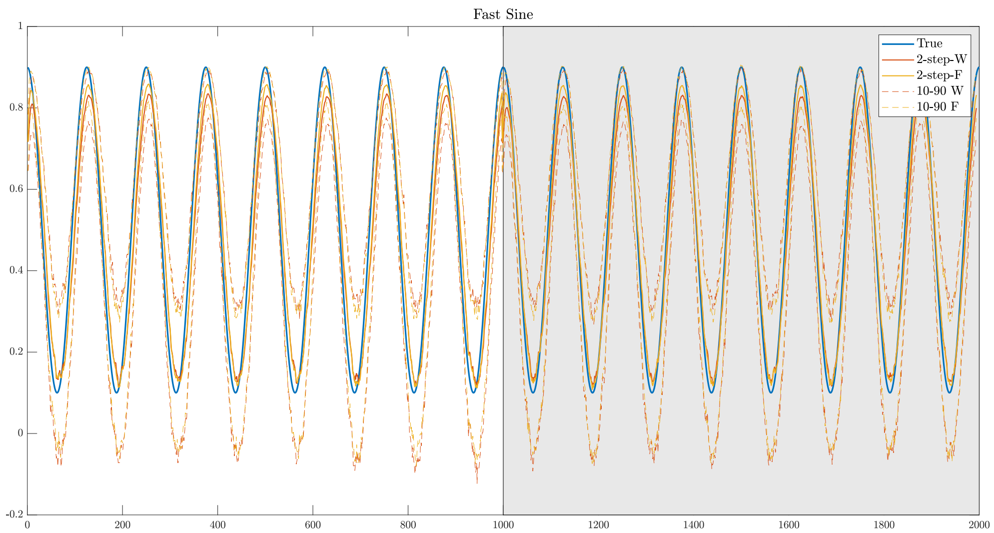

Figure 9: In-sample and out-of-sample filtered estimates of $\rho_t^{(2)}$ from both the 2-step-W and 2-step-F models. Estimates are averaged over the 1000 simulations. Confidence bands are constructed by computing the 10% and 90% empirical quantiles.

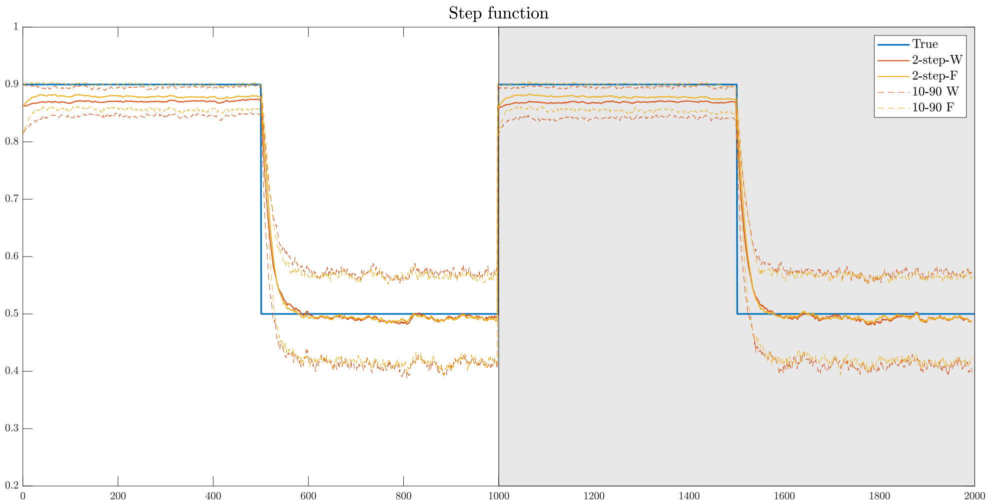

Figure 10: In-sample and out-of-sample filtered estimates of $\rho_t^{(3)}$ from both the 2-step-W and 2-step-F models. Estimates are averaged over the 1000 simulations. Confidence bands are constructed by computing the 10% and 90% empirical quantiles.

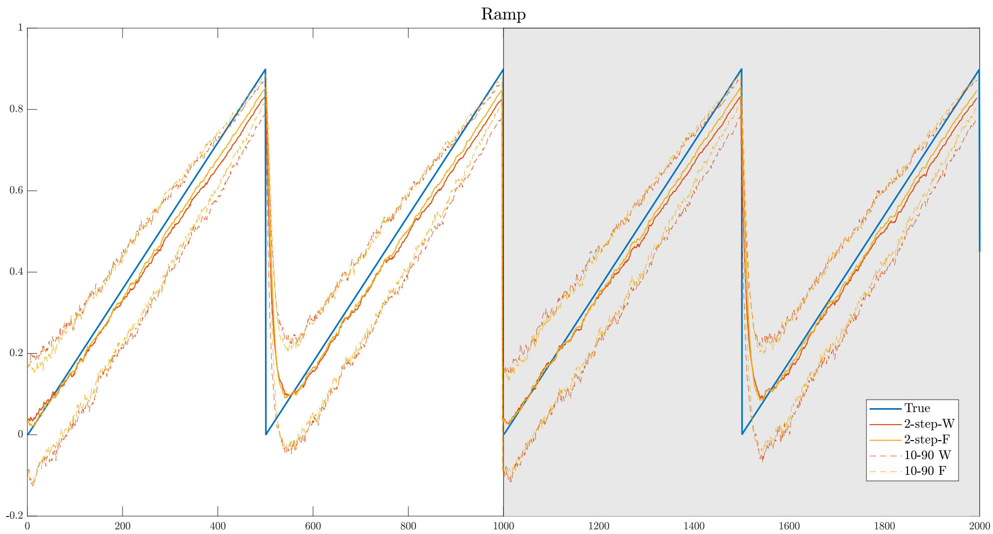

Figure 11: In-sample and out-of-sample filtered estimates of $\rho_t^{(4)}$ from both the 2-step-W and 2-step-F models. Estimates are averaged over the 1000 simulations. Confidence bands are constructed by computing the 10% and 90% empirical quantiles.