Supplemental appendix for

Indirect estimation of large conditionally heteroskedastic factor models, with an application to the Dow 30 stocks

Enrique Sentana
CEMFI, Casado del Alisal 5, E-28014 Madrid, Spain
<sentana@cemfi.es>

Giorgio Calzolari
Università di Firenze, Viale Morgagni 59, I-50134 Firenze, Italy
<calzolar@ds.unifi.it>

Gabriele Fiorentini
Università di Firenze, Viale Morgagni 59, I-50134 Firenze, Italy
<fiorentini@ds.unifi.it>

Revised: April 2008

B The score of the HRS approximate likelihood function

The log-likelihood function of the HRS model that we consider is given by $\bar{l}_T(\boldsymbol{\theta}) = T^{-1}\sum_{t=1}^{T} l_t(\boldsymbol{\theta})$, where

$$l_t(\boldsymbol{\theta}) = -\frac{N}{2}\log 2\pi - \frac{1}{2}\log\left|\mathbf{C}\boldsymbol{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}' + \boldsymbol{\Gamma}_t(\boldsymbol{\theta})\right| - \frac{1}{2}\mathbf{x}_t'[\mathbf{C}\boldsymbol{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}' + \boldsymbol{\Gamma}_t(\boldsymbol{\theta})]^{-1}\mathbf{x}_t, \tag{B3}$$

$\boldsymbol{\Lambda}_t(\boldsymbol{\theta})$ is a $k \times k$ diagonal matrix with typical element

$$\lambda_{jt}(\boldsymbol{\theta}) = \varpi_j + \alpha_j[g^2_{jt-1|t-1}(\boldsymbol{\theta}) + \omega_{jjt-1|t-1}(\boldsymbol{\theta})] + \beta_j\lambda_{jt-1}(\boldsymbol{\theta}),$$

with $\varpi_j = (1-\alpha_j-\beta_j)\lambda_j$, $\boldsymbol{\Gamma}_t(\boldsymbol{\theta})$ is a $N \times N$ diagonal matrix with typical element

$$\gamma_{it}(\boldsymbol{\theta}) = \varpi_i^* + \alpha_i^*[v^2_{it-1|t-1}(\boldsymbol{\theta}) + \xi_{iit-1|t-1}(\boldsymbol{\theta})] + \beta_i^*\gamma_{it-1}(\boldsymbol{\theta}),$$

with $\varpi_i^* = (1-\alpha_i^*-\beta_i^*)\gamma_i$ and where $g_{jt|t}(\boldsymbol{\theta})$, $v_{it|t}(\boldsymbol{\theta})$, $\omega_{jlt|t}(\boldsymbol{\theta})$ and $\xi_{ilt|t}(\boldsymbol{\theta})$ are typical elements of the outputs of the Kalman filter updating equations:

$$\begin{aligned}
\mathbf{g}_{t|t}(\boldsymbol{\theta}) &= E(\mathbf{g}_t|X_t;\boldsymbol{\theta}) = \boldsymbol{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}'[\mathbf{C}\boldsymbol{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}' + \boldsymbol{\Gamma}_{\mathbf{t}}(\boldsymbol{\theta})]^{-1}\mathbf{x}_t,\\
\mathbf{v}_{t|t}(\boldsymbol{\theta}) &= E(\mathbf{v}_t|X_t;\boldsymbol{\theta}) = \mathbf{x}_t - \mathbf{C}\mathbf{g}_{t|t}(\boldsymbol{\theta}),\\
\boldsymbol{\Omega}_{t|t}(\boldsymbol{\theta}) &= V(\mathbf{g}_t|X_t;\boldsymbol{\theta}) = \boldsymbol{\Lambda}_t(\boldsymbol{\theta}) - \boldsymbol{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}'[\mathbf{C}\boldsymbol{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}' + \boldsymbol{\Gamma}_{\mathbf{t}}(\boldsymbol{\theta})]^{-1}\mathbf{C}\boldsymbol{\Lambda}_t(\boldsymbol{\theta}),\\
\boldsymbol{\Xi}_{t|t}(\boldsymbol{\theta}) &= V(\mathbf{v}_t|X_t;\boldsymbol{\theta}) = \mathbf{C}\boldsymbol{\Omega}_{t|t}(\boldsymbol{\theta})\mathbf{C}'.
\end{aligned}$$

Bollerslev and Wooldridge (1992) show that the score function $\mathbf{s}_t(\boldsymbol{\theta}) = \partial l_t(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ of any multivariate conditionally heteroskedastic dynamic regression model with conditional mean vector $\boldsymbol{\mu}_t(\boldsymbol{\theta})$ and conditional covariance matrix $\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$ is given by the following expression:

$$\begin{aligned}
\mathbf{s}_t(\boldsymbol{\theta}) &= \frac{\partial\boldsymbol{\mu}_t'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})[\mathbf{x}_t - \boldsymbol{\mu}_t(\boldsymbol{\theta})]\\
&\quad + \frac{1}{2}\frac{\partial vec'\left[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right]}{\partial\boldsymbol{\theta}}\left[\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\otimes\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\right] vec\left\{[\mathbf{x}_t - \boldsymbol{\mu}_t(\boldsymbol{\theta})][\mathbf{x}_t - \boldsymbol{\mu}_t(\boldsymbol{\theta})]' - \boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right\}.
\end{aligned}$$

In our case the first term disappears because $\boldsymbol{\mu}_t(\boldsymbol{\theta}) = \mathbf{0}$. As for the second, given that the differential of $\boldsymbol{\Sigma}_t$ is

$$d(\mathbf{C}\boldsymbol{\Lambda}_t\mathbf{C}' + \boldsymbol{\Gamma}_t) = (d\mathbf{C})\boldsymbol{\Lambda}_t\mathbf{C}' + \mathbf{C}(d\boldsymbol{\Lambda}_t)\mathbf{C}' + \mathbf{C}\boldsymbol{\Lambda}_t(d\mathbf{C}') + d\boldsymbol{\Gamma}_t, \tag{B4}$$

(cf. Magnus and Neudecker (1999)), we have that Jacobian of $\boldsymbol{\Sigma}_t(\boldsymbol{\theta})$ will be:

$$\frac{\partial vec\left[\boldsymbol{\Sigma}_t(\boldsymbol{\theta})\right]}{\partial\boldsymbol{\theta}'} = (\mathbf{I}_{N^2} + \mathbf{K}_{NN})[\mathbf{I}_N \otimes \mathbf{C}\boldsymbol{\Lambda}_t(\boldsymbol{\theta})]\frac{\partial\mathbf{c}}{\partial\boldsymbol{\theta}'} + \mathbf{E}_N\frac{\partial\boldsymbol{\gamma}_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'} + (\mathbf{C}\otimes\mathbf{C})\mathbf{E}_k\frac{\partial\boldsymbol{\lambda}_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'},$$

where $\mathbf{E}_n$ is the unique $n^2 \times n$ "diagonalisation" matrix which transforms $vec(\mathbf{A})$ into $vecd(\mathbf{A})$ as $vecd(\mathbf{A}) = \mathbf{E}_n'vec(\mathbf{A})$, and $\mathbf{K}_{mn}$ is the commutation matrix of orders m and n (see Magnus (1988)).

After some straightforward algebraic manipulations, we get:

$$\begin{aligned}
\mathbf{s}_t(\boldsymbol{\theta}) &= \frac{\partial \mathbf{c}'}{\partial \boldsymbol{\theta}} vec\left[\boldsymbol{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}'\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\mathbf{x}_t\mathbf{x}_t'\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}'\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\right] \\
&+\frac{1}{2}\frac{\partial \boldsymbol{\gamma}_t'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} vecd\left[\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\mathbf{x}_t\mathbf{x}_t'\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta}) - \boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\right] \\
&+\frac{1}{2}\frac{\partial \boldsymbol{\lambda}_t'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} vecd\left[\mathbf{C}'\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\mathbf{x}_t\mathbf{x}_t'\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\mathbf{C} - \mathbf{C}'\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\theta})\mathbf{C}\right].
\end{aligned}$$

In view of (9), we will have that:

$$\begin{aligned}
\frac{\partial \lambda_{jt}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \alpha_j\left[2g_{jt-1|t-1}(\boldsymbol{\theta})\frac{\partial g_{jt-1|t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \frac{\partial \omega_{jjt-1|t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right] + \beta_j \frac{\partial \lambda_{jt-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\
&+\frac{\partial \varpi_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \frac{\partial \alpha_j}{\partial \boldsymbol{\theta}}\left[g^2_{jt-1|t-1}(\boldsymbol{\theta}) + \omega_{jjt-1|t-1}(\boldsymbol{\theta})\right] + \frac{\partial \beta_j}{\partial \boldsymbol{\theta}'}\lambda_{jt-1}(\boldsymbol{\theta}),
\end{aligned}$$

where $\varpi_j(\boldsymbol{\theta}) = (1-\alpha_j-\beta_j)\lambda_j$. Similarly, (10) implies that

$$\begin{aligned}
\frac{\partial \gamma_{it}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \alpha_i^*\left[2v_{it-1|t-1}(\boldsymbol{\theta})\frac{\partial v_{it-1|t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \frac{\partial \xi_{iit-1|t-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right] + \beta_i^* \frac{\partial \gamma_{it-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\
&+\frac{\partial \varpi_i^*(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \frac{\partial \alpha_i^*}{\partial \boldsymbol{\theta}}\left[v^2_{it-1|t-1}(\boldsymbol{\theta}) + \xi_{iit-1|t-1}(\boldsymbol{\theta})\right] + \frac{\partial \beta_i^*}{\partial \boldsymbol{\theta}'}\gamma_{it-1}(\boldsymbol{\theta}),
\end{aligned}$$

where $\varpi_i^*(\boldsymbol{\theta}) = (1-\alpha_i^*-\beta_i^*)\gamma_i$. If we impose the restriction $\alpha_i^* = \alpha^*$ and $\beta_i^* = \beta^*\ \forall i$, then the usual chain rule implies that $s_{\alpha^* t}(\boldsymbol{\theta}) = \sum_{i=1}^N s_{\alpha_i^* t}(\boldsymbol{\theta})$ and $s_{\beta^* t}(\boldsymbol{\theta}) = \sum_{i=1}^N s_{\beta_i^* t}(\boldsymbol{\theta})$.

Finally, it is worth mentioning that if we fix the factor scales by setting $c_{jj} = 1$ instead of $\lambda_j = 1$ for $j = 1, \ldots, k$, then we must exclude the elements of the score corresponding to those factor loadings, and replace them with the derivatives with respect to λ_j, which can be trivially found from the previous expressions because the unconditional variance parameters only appear directly in the expression for the pseudo log-likelihood function $l_t(\boldsymbol{\theta})$ in (B3) through the constant term in the conditional variance expressions, $\varpi_j(\boldsymbol{\theta})$. Either way, since we initialise the conditional variances with $\lambda_{j1}(\boldsymbol{\theta}) = \lambda_j$ and $\gamma_{i1}(\boldsymbol{\theta}) = \gamma_i$, then we must always start up the derivative recursions with $\partial\lambda_{j1}(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \partial\lambda_j/\partial\boldsymbol{\theta}$ and $\partial\gamma_{i1}(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \partial\gamma_i/\partial\boldsymbol{\theta}$.

If $\boldsymbol{\Gamma}_t > \mathbf{0}$, then we can use the Woodbury formula to prove that

$$\begin{aligned}
\mathbf{g}_{t|t} &= \boldsymbol{\Omega}_{t|t}\mathbf{C}'\boldsymbol{\Gamma}_t^{-1}\mathbf{x}_t, \\
\boldsymbol{\Omega}_{t|t} &= \left(\mathbf{C}'\boldsymbol{\Gamma}_t^{-1}\mathbf{C}+\boldsymbol{\Lambda}_t^{-1}\right)^{-1}, \\
\boldsymbol{\Lambda}_t\mathbf{C}'\boldsymbol{\Sigma}_t^{-1}\mathbf{x}_t\mathbf{x}_t'\boldsymbol{\Sigma}_t^{-1} - \mathbf{C}\boldsymbol{\Sigma}_t^{-1}\mathbf{C}\boldsymbol{\Lambda}_t &= [\mathbf{g}_{t|t}\mathbf{x}_t - (\mathbf{g}_{t|t}\mathbf{g}_{t|t}' + \boldsymbol{\Omega}_{t|t})\mathbf{C}']\boldsymbol{\Gamma}_t^{-1}, \\
\boldsymbol{\Sigma}_t^{-1}\mathbf{x}_t\mathbf{x}_t'\boldsymbol{\Sigma}_t^{-1} - \boldsymbol{\Sigma}_t^{-1} &= \boldsymbol{\Gamma}_t^{-1}[(\mathbf{x}_t - \mathbf{C}\mathbf{g}_{t|t})(\mathbf{x}_t - \mathbf{C}\mathbf{g}_{t|t})' + \mathbf{C}\boldsymbol{\Omega}_{t|t}\mathbf{C}' - \boldsymbol{\Gamma}_t]\boldsymbol{\Gamma}_t^{-1},
\end{aligned}$$

and

$$\mathbf{C}'\boldsymbol{\Sigma}_t^{-1}\mathbf{x}_t\mathbf{x}_t'\boldsymbol{\Sigma}_t^{-1}\mathbf{C} - \mathbf{C}'\boldsymbol{\Sigma}_t^{-1}\mathbf{C} = \boldsymbol{\Lambda}_t^{-1}[(\mathbf{g}_{t|t}\mathbf{g}_{t|t}' + \boldsymbol{\Omega}_{t|t}) - \boldsymbol{\Lambda}_t]\boldsymbol{\Lambda}_t^{-1},$$

which greatly simplifies the computations (see Sentana (2000)).

Under the same assumption, the differential of $\mathbf{\Omega}_{t|t}$ will be $-\mathbf{\Omega}_{t|t}d\left(\mathbf{C}'\mathbf{\Gamma}_t^{-1}\mathbf{C}+\mathbf{\Lambda}_t^{-1}\right)\mathbf{\Omega}_{t|t}$, where

$$d\left(\mathbf{C}'\mathbf{\Gamma}_t^{-1}\mathbf{C}+\mathbf{\Lambda}_t^{-1}\right)=(d\mathbf{C}')\mathbf{\Gamma}_t^{-1}\mathbf{C}+\mathbf{C}'\mathbf{\Gamma}_t^{-1}(d\mathbf{C})-\mathbf{C}'\mathbf{\Gamma}_t^{-1}(d\mathbf{\Gamma}_t)\mathbf{\Gamma}_t^{-1}\mathbf{C}-\mathbf{\Lambda}_t^{-1}(d\mathbf{\Lambda}_t)\mathbf{\Lambda}_t^{-1}.$$

If we call $\boldsymbol{\omega}_{t|t}=vech(\mathbf{\Omega}_{t|t})=\mathbf{D}_k^+vec(\mathbf{\Omega}_{t|t})$, where $\mathbf{D}_k$ is the duplication matrix of order k and $\mathbf{D}_k^+$ its Moore-Penrose inverse, then we will have that

$$\begin{aligned}\frac{\partial\boldsymbol{\omega}'_{t|t}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} &= \left[-2\frac{\partial\mathbf{c}'}{\partial\boldsymbol{\theta}}(\mathbf{\Gamma}_t^{-1}\mathbf{C}\mathbf{\Omega}_{t|t}\otimes\mathbf{\Omega}_{t|t})+\frac{\partial\boldsymbol{\gamma}'_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{E}'_N(\mathbf{\Gamma}_t^{-1}\mathbf{C}\mathbf{\Omega}_{t|t}\otimes\mathbf{\Gamma}_t^{-1}\mathbf{C}\mathbf{\Omega}_{t|t})\right.\\ &\qquad\left.+\frac{\partial\boldsymbol{\lambda}'_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{E}'_k(\mathbf{\Lambda}_t^{-1}\mathbf{\Omega}_{t|t}\otimes\mathbf{\Lambda}_t^{-1}\mathbf{\Omega}_{t|t})\right]\mathbf{D}_k^{+\prime}.\end{aligned}$$

In addition, the differential of $\mathbf{g}_{t|t}$ when $\mathbf{\Gamma}_t$ has full rank will be given by

$$d\mathbf{g}_{t|t}=(d\mathbf{\Omega}_{t|t})\mathbf{C}'\mathbf{\Gamma}_t^{-1}\mathbf{x}_t+\mathbf{\Omega}_{t|t}d(\mathbf{C}')\mathbf{\Gamma}_t^{-1}\mathbf{x}_t-\mathbf{\Omega}_{t|t}\mathbf{C}'\mathbf{\Gamma}_t^{-1}d(\mathbf{\Gamma}_t)\mathbf{\Gamma}_t^{-1}\mathbf{x}_t.$$

As a result, we will have that

$$\frac{\partial\mathbf{g}'_{t|t}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}=\frac{\partial\mathbf{c}'}{\partial\boldsymbol{\theta}}(\mathbf{\Gamma}_t^{-1}\mathbf{x}_t\otimes\mathbf{\Omega}_{t|t})-\frac{\partial\boldsymbol{\gamma}'_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{E}'_N(\mathbf{\Gamma}_t^{-1}\mathbf{x}_t\otimes\mathbf{\Gamma}_t^{-1}\mathbf{C}\mathbf{\Omega}_{t|t})+\frac{\partial\boldsymbol{\omega}'_{t|t}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{D}'_k(\mathbf{C}'\mathbf{\Gamma}_t^{-1}\mathbf{x}_t\otimes\mathbf{I}_k).$$

Similarly, given that $\mathbf{v}_{t|t}=\mathbf{x}_t-\mathbf{C}\mathbf{g}_{t|t}$, we will have that

$$d\mathbf{v}_{t|t}=-(d\mathbf{C})\mathbf{g}_{t|t}-\mathbf{C}(d\mathbf{g}_{t|t}),$$

whence

$$\frac{\partial\mathbf{v}'_{t|t}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}=-\frac{\partial\mathbf{c}'}{\partial\boldsymbol{\theta}}(\mathbf{I}_N\otimes\mathbf{g}_{t|t})-\frac{\partial\mathbf{g}'_{t|t}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{C}'.$$

In addition, since $\mathbf{\Xi}_{t|t}=\mathbf{C}\mathbf{\Omega}_{t|t}\mathbf{C}'$, then

$$d\mathbf{\Xi}_{t|t}=(d\mathbf{C})\mathbf{\Omega}_{t|t}\mathbf{C}'+\mathbf{C}(d\mathbf{\Omega}_{t|t})\mathbf{C}'+\mathbf{C}\mathbf{\Omega}_{t|t}(d\mathbf{C}').$$

Hence, if we call $\boldsymbol{\xi}_{t|t}=vech(\mathbf{\Xi}_{t|t})=\mathbf{D}_N^+vec(\mathbf{\Xi}_{t|t})$, then after some algebraic manipulations we will have that

$$\frac{\partial\boldsymbol{\xi}'_{t|t}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}=\left[2\frac{\partial\mathbf{c}'}{\partial\boldsymbol{\theta}}(\mathbf{I}_N\otimes\mathbf{\Omega}_{t|t}\mathbf{C}')+\frac{\partial\boldsymbol{\omega}'_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{D}'_k(\mathbf{C}'\otimes\mathbf{C}')\right]\mathbf{D}_N^{+\prime}.$$

If some $\gamma_{it}=0$, though, the above expressions become invalid. Nevertheless, appropriately modified expressions can be developed along the lines of Sentana (2000). For the sake of brevity, though, we only obtain the score when $rank(\mathbf{\Gamma}_t)=N-k$, so that there are as many Heywood cases as factors. To do so, let us partition the original set of variables in two subsets, say $\mathbf{x}_{at}$ and $\mathbf{x}_{bt}$, of dimensions k and $N-k$ respectively. With this notation, we can re-write the auxiliary model as

$$\begin{pmatrix}\mathbf{x}_{at}\\ \mathbf{x}_{bt}\end{pmatrix}=\begin{pmatrix}\mathbf{C}_a\\ \mathbf{C}_b\end{pmatrix}\mathbf{g}_t+\begin{pmatrix}\mathbf{v}_{at}\\ \mathbf{v}_{bt}\end{pmatrix},$$

where

$$\begin{pmatrix} \mathbf{g}_t \\ \mathbf{v}_{at} \\ \mathbf{v}_{bt} \end{pmatrix} |X_{t-1} \sim N \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{\Lambda}_t & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Gamma}_{at} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Gamma}_{bt} \end{pmatrix} \right].$$

In this context, it is convenient to factorise the joint log-likelihood function of $\mathbf{x}_{at}$ and $\mathbf{x}_{bt}$ (given X_{t-1}) as the marginal log-likelihood function of $\mathbf{x}_{at}$ (given X_{t-1}) plus the conditional log-likelihood function of $\mathbf{x}_{bt}$ given $\mathbf{x}_{at}$ (and X_{t-1}). More formally, we can write

$$l_t(\boldsymbol{\theta}) = l_{at}(\boldsymbol{\theta}) + l_{bt|at}(\boldsymbol{\theta}),$$

so that

$$\mathbf{s}_t(\boldsymbol{\theta}) = \mathbf{s}_{at}(\boldsymbol{\theta}) + \mathbf{s}_{bt|at}(\boldsymbol{\theta}).$$

The two log-likelihood components will be given by

$$l_{at}(\boldsymbol{\theta}) = -\frac{k}{2}\log 2\pi - \frac{1}{2}\log|\mathbf{\Sigma}_{at}| - \frac{1}{2}\mathbf{x}'_{at}\mathbf{\Sigma}_{at}^{-1}\mathbf{x}_{at}$$

and

$$l_{bt|at}(\boldsymbol{\theta}) = -\frac{N-k}{2}\log 2\pi - \frac{1}{2}\log\left|\mathbf{\Sigma}_{bt|at}(\boldsymbol{\theta})\right| - \frac{1}{2}\boldsymbol{\varepsilon}_{bt|at}(\boldsymbol{\theta})'\mathbf{\Sigma}_{bt|at}^{-1}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_{bt|at}(\boldsymbol{\theta}),$$

where

$$\begin{aligned}
\mathbf{\Sigma}_{at}(\boldsymbol{\theta}) &= V(\mathbf{x}_{at}|X_{t-1};\boldsymbol{\theta}) = \mathbf{C}_a\mathbf{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}'_a + \mathbf{\Gamma}_{at}(\boldsymbol{\theta}),\\
\boldsymbol{\varepsilon}_{bt|at}(\boldsymbol{\theta}) &= \mathbf{x}_{bt} - \boldsymbol{\mu}_{bt|at}(\boldsymbol{\theta}),\\
\boldsymbol{\mu}_{bt|at}(\boldsymbol{\theta}) &= E(\mathbf{x}_{bt}|\mathbf{x}_{at}, X_{t-1};\boldsymbol{\theta}) = \mathbf{C}_{bt}\mathbf{g}_{t|at}(\boldsymbol{\theta}),\\
\mathbf{g}_{t|at}(\boldsymbol{\theta}) &= E(\mathbf{g}_t|\mathbf{x}_{at}, X_{t-1};\boldsymbol{\theta}) = \mathbf{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}'_a\mathbf{\Sigma}_{at}^{-1}(\boldsymbol{\theta})\mathbf{x}_{at},\\
\mathbf{\Sigma}_{bt|at}(\boldsymbol{\theta}) &= E(\mathbf{g}_t|\mathbf{x}_{at}, X_{t-1};\boldsymbol{\theta}) = \mathbf{C}_b\mathbf{\Omega}_{t|at}(\boldsymbol{\theta})\mathbf{C}'_b + \mathbf{\Gamma}_{bt}(\boldsymbol{\theta}),
\end{aligned}$$

and

$$\mathbf{\Omega}_{t|at}(\boldsymbol{\theta}) = V(\mathbf{g}_t|\mathbf{x}_{at}, X_{t-1};\boldsymbol{\theta}) = \mathbf{\Lambda}_t(\boldsymbol{\theta}) - \mathbf{\Lambda}_t(\boldsymbol{\theta})\mathbf{C}'_a\mathbf{\Sigma}_{at}^{-1}(\boldsymbol{\theta})\mathbf{C}_a\mathbf{\Lambda}_t(\boldsymbol{\theta}).$$

Therefore, if we partition $\mathbf{c}$ and $\boldsymbol{\gamma}$ as $(\mathbf{c}'_a, \mathbf{c}'_b)'$ and $(\boldsymbol{\gamma}'_a, \boldsymbol{\gamma}'_b)'$, respectively, where $\mathbf{c}_a = vec(\mathbf{C}'_a)$, $\mathbf{c}_b = vec(\mathbf{C}'_b)$, $\boldsymbol{\gamma}_{at} = vecd(\mathbf{\Gamma}_{at})$, and $\boldsymbol{\gamma}_{bt} = vecd(\mathbf{\Gamma}_{bt})$, then we can use the expressions derived before to find

$$\begin{aligned}
\mathbf{s}_{at}(\boldsymbol{\theta}) = {} & \frac{\partial\mathbf{c}'_a}{\partial\boldsymbol{\theta}}vec(\mathbf{\Lambda}_t\mathbf{C}'_a\mathbf{\Sigma}_{at}^{-1}\mathbf{x}_{at}\mathbf{x}'_{at}\mathbf{\Sigma}_{at}^{-1} - \mathbf{\Lambda}_t\mathbf{C}'_a\mathbf{\Sigma}_{at}) + \frac{1}{2}\frac{\partial\boldsymbol{\gamma}'_{at}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}vecd(\mathbf{\Sigma}_{at}^{-1}\mathbf{x}_{at}\mathbf{x}'_{at}\mathbf{\Sigma}_{at}^{-1} - \mathbf{\Sigma}_{at}^{-1})\\
& + \frac{1}{2}\frac{\partial\boldsymbol{\lambda}'_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}vecd(\mathbf{C}'_a\mathbf{\Sigma}_{at}^{-1}\mathbf{x}_{at}\mathbf{x}'_{at}\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a - \mathbf{C}'_a\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a).
\end{aligned}$$

In order to obtain $\mathbf{s}_{bt|at}(\boldsymbol{\theta})$, though, we first need to find the Jacobian matrices $\partial\boldsymbol{\mu}_{bt|at}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}'$ and $\partial vec[\Sigma_{bt|at}(\boldsymbol{\theta})]/\partial\boldsymbol{\theta}'$. Straightforward algebra shows that

$$\frac{\partial\boldsymbol{\mu}'_{bt|at}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \frac{\partial\mathbf{c}'_b}{\partial\boldsymbol{\theta}}(\mathbf{I}_{N-k}\otimes\mathbf{g}_{t|at}) + \frac{\partial\mathbf{g}'_{t|at}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{C}'_b$$

and

$$\frac{\partial vec'\left[\mathbf{\Sigma}_{bt|at}(\boldsymbol{\theta})\right]}{\partial\boldsymbol{\theta}} = \frac{\partial\mathbf{c}_b'}{\partial\boldsymbol{\theta}}(\mathbf{I}_{N-k}\otimes\mathbf{\Omega}_{bt|at}\mathbf{C}_b')(\mathbf{I}_{(N-k)^2k^2}+\mathbf{K}_{(N-k)k,(N-k)k}) + \frac{\partial\boldsymbol{\gamma}_{bt}'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{E}_{N-k}' + \frac{\partial\boldsymbol{\omega}_{bt|at}'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{D}_k'(\mathbf{C}_b'\otimes\mathbf{C}_b').$$

Hence,

$$\mathbf{s}_{bt|at}(\boldsymbol{\theta}) = \frac{\partial\mathbf{g}_{t|at}'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{C}_b'\mathbf{\Sigma}_{bt|at}^{-1}\boldsymbol{\varepsilon}_{bt|at}$$
$$+\frac{\partial\mathbf{c}_b'}{\partial\boldsymbol{\theta}}vec\left(\mathbf{\Sigma}_{bt|at}^{-1}\boldsymbol{\varepsilon}_{bt|at}\mathbf{g}_{t|at}+\mathbf{\Sigma}_{bt|at}^{-1}\boldsymbol{\varepsilon}_{bt|at}'\boldsymbol{\varepsilon}_{bt|at}\mathbf{\Sigma}_{bt|at}^{-1}\mathbf{C}_b\mathbf{\Omega}_{t|at}-\mathbf{\Sigma}_{bt|at}^{-1}\mathbf{C}_b\mathbf{\Omega}_{t|at}\right)$$
$$+\frac{1}{2}\frac{\partial\boldsymbol{\gamma}_{bt}'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}vecd\left(\mathbf{\Sigma}_{bt|at}^{-1}\boldsymbol{\varepsilon}_{bt|at}\boldsymbol{\varepsilon}_{bt|at}'\mathbf{\Sigma}_{bt|at}^{-1}-\mathbf{\Sigma}_{bt|at}^{-1}\right)$$
$$+\frac{1}{2}\frac{\partial\boldsymbol{\omega}_{bt|at}'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{D}_k'vec\left(\mathbf{C}_b'\mathbf{\Sigma}_{bt|at}^{-1}\boldsymbol{\varepsilon}_{bt|at}\boldsymbol{\varepsilon}_{bt|at}'\mathbf{\Sigma}_{bt|at}^{-1}\mathbf{C}_b-\mathbf{C}_b'\mathbf{\Sigma}_{bt|at}^{-1}\mathbf{C}_b\right).$$

In this case, the differential of $\mathbf{g}_{t|at}(\boldsymbol{\theta})$ will be

$$d\mathbf{g}_{t|at} = (d\mathbf{\Lambda}_t)\mathbf{C}_a'\mathbf{\Sigma}_{at}^{-1}\mathbf{x}_{at}+\mathbf{\Lambda}_t(d\mathbf{C}_a')\mathbf{\Sigma}_{at}^{-1}\mathbf{x}_{at}-\mathbf{\Lambda}_t\mathbf{C}_a'\mathbf{\Sigma}_{at}^{-1}(d\mathbf{\Sigma}_{at})\mathbf{\Sigma}_{at}^{-1}\mathbf{x}_{at},$$

where $d\mathbf{\Sigma}_{at}$ is analogous to (B4). As a result,

$$\frac{\partial\mathbf{g}_{t|at}'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \frac{\partial\mathbf{c}_a'}{\partial\boldsymbol{\theta}}\left[(\mathbf{\Sigma}_{at}^{-1}\mathbf{x}_{at}\otimes\mathbf{\Omega}_{t|at})-(\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a\mathbf{\Lambda}_t\otimes\mathbf{g}_{t|at})\right] - \frac{\partial\boldsymbol{\gamma}_{at}'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{E}_k'(\mathbf{\Sigma}_{at}^{-1}\mathbf{x}_{at}\otimes\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a\mathbf{\Lambda}_t)+\frac{\partial\boldsymbol{\lambda}_t'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{E}_k'(\mathbf{\Lambda}_t^{-1}\mathbf{g}_{t|at}\otimes\mathbf{\Lambda}_t^{-1}\mathbf{\Omega}_{t|at}).$$

Similarly, the differential of $\mathbf{\Omega}_{t|at}$ will be given by

$$d\mathbf{\Omega}_{t|at} = (d\mathbf{\Lambda}_t)-(d\mathbf{\Lambda}_t)\mathbf{C}_a'\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a\mathbf{\Lambda}_t-\mathbf{\Lambda}_t(d\mathbf{C}_a')\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a\mathbf{\Lambda}_t+\mathbf{\Lambda}_t\mathbf{C}_a'\mathbf{\Sigma}_{at}^{-1}(d\mathbf{\Sigma}_{at})\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a\mathbf{\Lambda}_t - \mathbf{\Lambda}_t\mathbf{C}_a'\mathbf{\Sigma}_{at}^{-1}(d\mathbf{C}_a)\mathbf{\Lambda}_t-\mathbf{\Lambda}_t\mathbf{C}_a'\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a(d\mathbf{\Lambda}_t).$$

Hence,

$$\frac{\partial\boldsymbol{\omega}_{t|at}'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \left\{-2\frac{\partial\mathbf{c}_a'}{\partial\boldsymbol{\theta}}(\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a\mathbf{\Lambda}_t\otimes\mathbf{\Omega}_{t|at})-\frac{\partial\boldsymbol{\gamma}_{at}'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{E}_k'(\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a\mathbf{\Lambda}_t\otimes\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a\mathbf{\Lambda}_t) + \frac{\partial\boldsymbol{\lambda}_t'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{E}_k'[(\mathbf{\Lambda}_t^{-1}\mathbf{\Omega}_{t|at}\otimes\mathbf{I}_k)-(\mathbf{C}_a'\mathbf{\Sigma}_{at}^{-1}\mathbf{C}_a\mathbf{\Lambda}_t\otimes\mathbf{\Lambda}_t^{-1}\mathbf{\Omega}_{t|at})]\right\}\mathbf{D}_k'.$$

Finally, we need to obtain $\partial\mathbf{g}_{t|t}'(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ and $\partial\boldsymbol{\omega}_{t|t}'(\boldsymbol{\theta})/\partial\boldsymbol{\theta}$. But since

$$\mathbf{g}_{t|t}(\boldsymbol{\theta}) = \mathbf{g}_{t|at}(\boldsymbol{\theta})+\mathbf{\Omega}_{t|at}(\boldsymbol{\theta})\mathbf{C}_b'\mathbf{\Sigma}_{bt|at}^{-1}(\boldsymbol{\theta})\boldsymbol{\varepsilon}_{bt|at}$$

and

$$\mathbf{\Omega}_{t|t} = \mathbf{\Omega}_{t|at}(\boldsymbol{\theta})-\mathbf{\Omega}_{t|at}(\boldsymbol{\theta})\mathbf{C}_b'\mathbf{\Sigma}_{bt|at}^{-1}(\boldsymbol{\theta})\mathbf{C}_b\mathbf{\Omega}_{t|at}(\boldsymbol{\theta}),$$

we can obtain the required derivatives by combining the previous expressions.

Fortunately, all the above formulae simplify considerably when $\mathbf{\Gamma}_{at} = \mathbf{0}$. Specifically, let $\mathring{\boldsymbol{\theta}}$ denote the value of $\boldsymbol{\theta}$ when $\mathbf{\Gamma}_{at} = \mathbf{0}$. Then, it is immediate to see that

$$\begin{aligned}\mathbf{\Sigma}_{at}(\mathring{\boldsymbol{\theta}}) &= \mathbf{C}_a\mathbf{\Lambda}_t\mathbf{C}_a',\\ \mathbf{g}_{t|at}(\mathring{\boldsymbol{\theta}}) &= \mathbf{C}_a^{-1}\mathbf{x}_{at},\end{aligned}$$

and $\mathbf{\Omega}_{t|at}(\mathring{\boldsymbol{\theta}}) = \mathbf{0}$, so that $\boldsymbol{\varepsilon}_{bt|at}(\mathring{\boldsymbol{\theta}}) = \mathbf{x}_{bt} - \mathbf{C}_b^*\mathbf{x}_{at}$, with $\mathbf{C}_b^* = \mathbf{C}_b\mathbf{C}_a^{-1}$, and $\mathbf{\Sigma}_{bt|at}(\mathring{\boldsymbol{\theta}}) = \mathbf{\Gamma}_{bt}$. Moreover,

$$\mathbf{\Sigma}_{at}(\mathring{\boldsymbol{\theta}})\mathbf{x}_{at}\mathbf{x}_{at}'\mathbf{\Sigma}_{at}(\mathring{\boldsymbol{\theta}}) - \mathbf{\Sigma}_{at}(\mathring{\boldsymbol{\theta}}) = \mathbf{C}_a'^{-1}\mathbf{\Lambda}_t^{-1}(\mathring{\boldsymbol{\theta}})\left[\mathbf{g}_{t|at}(\mathring{\boldsymbol{\theta}})\mathbf{g}_{t|at}'(\mathring{\boldsymbol{\theta}}) - \mathbf{\Lambda}_t(\mathring{\boldsymbol{\theta}})\right]\mathbf{\Lambda}_t^{-1}(\mathring{\boldsymbol{\theta}})\mathbf{C}_a^{-1}.$$

As a result, we can write

$$\begin{aligned}\mathbf{s}_{at}(\mathring{\boldsymbol{\theta}}) &= \frac{\partial \mathbf{c}_a'}{\partial \boldsymbol{\theta}} vec\left[\left(\mathbf{g}_{t|at}\mathbf{g}_{t|at}' - \mathbf{\Lambda}_t\right)\mathbf{\Lambda}_t^{-1}\mathbf{C}_a^{-1}\right]\\ &+\frac{1}{2}\frac{\partial \boldsymbol{\gamma}_{at}'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} vecd[\mathbf{C}_a'^{-1}\mathbf{\Lambda}_t^{-1}(\mathbf{g}_{t|at}\mathbf{g}_{t|at}' - \mathbf{\Lambda}_t)\mathbf{\Lambda}_t^{-1}\mathbf{C}_a^{-1}] + \frac{1}{2}\frac{\partial \boldsymbol{\lambda}_t'(\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} vecd[\mathbf{\Lambda}_t^{-1}(\mathbf{g}_{t|at}\mathbf{g}_{t|at}' - \mathbf{\Lambda}_t)\mathbf{\Lambda}_t^{-1}]\end{aligned}$$

and

$$\begin{aligned}\mathbf{s}_{bt|at}(\mathring{\boldsymbol{\theta}}) &= -\frac{\partial \mathbf{c}_a'}{\partial \boldsymbol{\theta}} vec[\mathbf{C}_b^{*\prime}\mathbf{\Gamma}_{bt}^{-1}\boldsymbol{\varepsilon}_{bt|at}(\mathring{\boldsymbol{\theta}})\mathbf{g}_{t|at}'(\mathring{\boldsymbol{\theta}})] + \frac{\partial \mathbf{c}_b'}{\partial \boldsymbol{\theta}} vec[\mathbf{\Gamma}_{bt}^{-1}\boldsymbol{\varepsilon}_{bt|at}(\mathring{\boldsymbol{\theta}})\mathbf{g}_{t|at}'(\mathring{\boldsymbol{\theta}})]\\ &+\frac{\partial \boldsymbol{\gamma}_{at}'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\left[\frac{1}{2} vecd\{\mathbf{C}_b^{*\prime}\mathbf{\Gamma}_{bt}^{-1}[\boldsymbol{\varepsilon}_{bt|at}(\mathring{\boldsymbol{\theta}})\boldsymbol{\varepsilon}_{bt|at}'(\mathring{\boldsymbol{\theta}}) - \mathbf{\Gamma}_{bt}]\mathbf{\Gamma}_{bt}^{-1}\mathbf{C}_b^{*\prime}\} - \mathbf{E}_k'\mathbf{C}_b^{*\prime}\mathbf{\Gamma}_{bt}^{-1}\boldsymbol{\varepsilon}_{bt|at}(\mathring{\boldsymbol{\theta}})\mathbf{g}_{t|at}'(\mathring{\boldsymbol{\theta}})\mathbf{\Lambda}_t^{-1}\mathbf{C}_a^{-1}\right]\\ &+\frac{1}{2}\frac{\partial \boldsymbol{\gamma}_{bt}'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} vecd\{\mathbf{\Gamma}_{bt}^{-1}[\boldsymbol{\varepsilon}_{bt|at}(\mathring{\boldsymbol{\theta}})\boldsymbol{\varepsilon}_{bt|at}'(\mathring{\boldsymbol{\theta}}) - \mathbf{\Gamma}_{bt}]\mathbf{\Gamma}_{bt}^{-1}\}.\end{aligned}$$

Finally, we obtain

$$\frac{\partial \mathbf{g}_{t|t}'(\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = \frac{\partial \mathbf{g}_{t|at}'(\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{\omega}_{t|at}'(\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\mathbf{D}_k[\mathbf{C}_b^{*\prime}\mathbf{\Gamma}_{bt}^{-1}\boldsymbol{\varepsilon}_{bt|at}(\mathring{\boldsymbol{\theta}}) \otimes \mathbf{I}_k]$$

and

$$\frac{\partial \boldsymbol{\omega}_{t|t}'(\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = \frac{\partial \boldsymbol{\omega}_{t|at}'(\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}},$$

where

$$\frac{\partial \mathbf{g}_{t|at}'(\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = -\frac{\partial \mathbf{c}_a'}{\partial \boldsymbol{\theta}}[\mathbf{C}_a^{-1\prime} \otimes \mathbf{g}_{t|at}'(\mathring{\boldsymbol{\theta}})] + \frac{\partial \boldsymbol{\gamma}_{at}'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\mathbf{E}_k'[\mathbf{C}_a^{-1\prime}\mathbf{\Lambda}_t^{-1}\mathbf{g}_{t|at}(\mathring{\boldsymbol{\theta}}) \otimes \mathbf{C}_a^{-1\prime}],$$

and

$$\frac{\partial \boldsymbol{\omega}_{t|at}'(\mathring{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = \frac{\partial \boldsymbol{\gamma}_{at}'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\mathbf{E}_k'(\mathbf{C}_a^{-1\prime} \otimes \mathbf{C}_a^{-1\prime})\mathbf{D}_k.$$

Although these expressions are strictly speaking only valid when an idiosyncratic variance is identically 0, we recommend their use whenever some γ_{it} is less than .0001 because the expressions for $\boldsymbol{\gamma}_t > \mathbf{0}$ become numerically unreliable for smaller values.