Web Appendix to Modeling financial sector joint tail risk in the euro area

André Lucas,[(a)] *Bernd Schwaab,*[(b)] *Xin Zhang*[(c)]

[(a)] VU University Amsterdam and Tinbergen Institute

[(b)] European Central Bank, Financial Research

[(c)] Sveriges Riksbank, Research Division

October 9, 2015

Appendix A: Univariate volatility models

This appendix summarizes our univariate marginal modeling strategy. We estimate univariate dynamic GHST models using the equity returns for each firm i based on the maximum likelihood approach. Using the estimated univariate models with parameters are $\tilde{\mu}_{it}$, $\tilde{\sigma}_{it}$, γ_i, and ν_i, we transform the observations into their probability integral transforms $\hat{u}_{it} \in [0,1]$. The univariate GHST density is given by

$$
\begin{aligned}
p(y_t; \tilde{\sigma}_t^2, \gamma, \nu) &= \frac{\nu^{\frac{\nu}{2}} 2^{1-\frac{\nu+1}{2}}}{\Gamma(\frac{\nu}{2}) \pi^{\frac{1}{2}} \tilde{\sigma}_t} \cdot \frac{K_{\frac{\nu+1}{2}}\left(\sqrt{d(y_t)(\gamma^2)}\right) e^{\gamma(y_t - \tilde{\mu}_t)/\tilde{\sigma}_t}}{d(y_t)^{\frac{\nu+1}{4}} \cdot (\gamma^2)^{-\frac{\nu+1}{4}}}, \\
d(y_t) &= \nu + (y_t - \tilde{\mu}_t)^2 / \tilde{\sigma}_t^2, \\
\tilde{\mu}_t &= -\frac{\nu}{\nu - 2} \tilde{\sigma}_t \gamma, \quad \tilde{\sigma}_t = \sigma_t T, \\
T &= \left(\frac{\nu}{\nu - 2} + \frac{2\nu^2 \gamma^2}{(\nu - 2)^2 (\nu - 4)} \right)^{-1/2},
\end{aligned}
$$

where $\tilde{\sigma}_t$ and σ_t are the univariate analogues to the Choleski square roots of the scale and covariance matrix, respectively. We here chose $\sigma_t = \sigma(f_t) = \exp(f_t)$. For this specification, the score driven dynamics are given by

$$
\begin{aligned}
f_{t+1} &= \tilde{\omega} + \sum_{i=0}^{p-1} A_i s_{t-i} + \sum_{j=0}^{q-1} B_j f_{t-j} + C(s_t - s_t^{\tilde{\mu}}) 1\{y_t < \tilde{\mu}_t\}, \\
s_t &= \mathcal{S}_t \nabla_t, \quad \nabla_t = \partial \ln p(y_t | \mathcal{F}_{t-1}; f_t, \theta) / \partial f_t, \\
s_t^{\tilde{\mu}} &= \mathcal{S}_t \nabla_t^{\tilde{\mu}}, \quad \nabla_t^{\tilde{\mu}} = \partial \ln p(\tilde{\mu}_t | \mathcal{F}_{t-1}; f_t, \theta) / \partial f_t, \\
\nabla_t &= \Psi_t H_t \left(w_t \cdot y_t^2 - \tilde{\sigma}_t^2 - \left(1 - \frac{\nu}{\nu - 2} w_t \right) \tilde{\sigma}_t \gamma y_t \right), \\
w_t &= \frac{\nu + 1}{2 d(y_t)} - \frac{k'_{0.5(\nu+1)}\left(\sqrt{d(y_t) \cdot (\gamma^2)}\right)}{\sqrt{d(y_t)/(\gamma^2)}}, \\
\Psi_t &= \frac{\partial \sigma_t}{\partial f_t}, \quad H_t = T \tilde{\sigma}^{-3}.
\end{aligned}
$$

The above score dynamics allow for a leverage effect, see, for example, Rodriguez and Ruiz (2012) for a survey and Lucas, Schwaab, and Zhang (2015). Univariate filtered volatilities for N=10 and for N=73 firms can be obtained in this way as well.

Appendix B: The dynamic GHST score

This appendix derives expression (9) for the GHST dynamic score, and discusses the choice of scaling matrix in (8).

We use the following two matrix calculus results for an invertible matrix X, and column vectors α and β.

$$\frac{\partial(\alpha' X^{-1}\beta)}{\partial \mathrm{vec}(X)} = -((X')^{-1}\otimes X^{-1})\mathrm{vec}(\alpha\beta'), \quad \frac{\partial(\log|X|)}{\partial \mathrm{vec}(X)} = ((X')^{-1}\otimes X^{-1})\mathrm{vec}(X).$$

We denote $k_\lambda(\cdot) = \ln K_\lambda(\cdot)$, with first derivative $k'_\lambda(\cdot)$. Based on these results one obtains

$$\begin{aligned}
\nabla_t &= \frac{\partial \mathrm{vec}(\tilde{\Sigma}_t)'}{\partial f_t}\frac{\partial \ln p(y_t;\tilde{\Sigma}_t,\gamma,\nu)}{\partial \mathrm{vec}(\tilde{\Sigma}_t)} \\
&= \Psi_t'\left(-w_t\frac{\partial d(y_t)}{\partial \mathrm{vec}(\tilde{\Sigma}_t)} + \breve{w}_t\frac{\partial d(\gamma)}{\partial \mathrm{vec}(\tilde{\Sigma}_t)} + \frac{\partial(\gamma'\tilde{\Sigma}_t^{-1}(y_t-\tilde{\mu}))}{\partial \mathrm{vec}(\tilde{\Sigma}_t)} - \frac{1}{2}\frac{\partial(\log(|\tilde{\Sigma}_t|))}{\partial \mathrm{vec}(\tilde{\Sigma}_t)}\right) \\
&= \Psi_t' H_t \mathrm{vec}\left(w_t(y_t-\tilde{\mu})(y_t-\tilde{\mu})' - \breve{w}_t\gamma\gamma' - \gamma(y_t-\tilde{\mu})' - 0.5\tilde{\Sigma}_t\right),
\end{aligned}$$

with $H_t = \tilde{\Sigma}_t^{-1}\otimes\tilde{\Sigma}_t^{-1}$, and

$$w_t = \frac{\nu+N}{4d(y_t)} - \frac{k'_{0.5(\nu+N)}\left(\sqrt{d(y_t)d(\gamma)}\right)}{2\sqrt{d(y_t)/d(\gamma)}}, \quad \breve{w}_t = \frac{\nu+N}{4d(\gamma)} + \frac{k'_{0.5(\nu+N)}\left(\sqrt{d(y_t)d(\gamma)}\right)}{2\sqrt{d(\gamma)/d(y_t)}}.$$

To scale the score updates, we set the scaling matrix $\mathcal{S}_t$ equal to the inverse conditional Fisher information matrix of the symmetric Student's t distribution,

$$\mathcal{S}_t = \left\{\Psi_t'(\tilde{\Sigma}_t^{-1}\otimes\tilde{\Sigma}_t^{-1})'[gG - \mathrm{vec}(\mathrm{I})\mathrm{vec}(\mathrm{I})'](\tilde{\Sigma}_t^{-1}\otimes\tilde{\Sigma}_t^{-1})\Psi_t\right\}^{-1}, \qquad \text{(B.1)}$$

where $g = (\nu+N)/(\nu+2+N)$, and $G = \mathrm{E}[x_t x_t' \otimes x_t x_t']$ for $x_t \sim \mathrm{N}(0, \mathrm{I}_N)$. We do not use the inverse conditional Fisher information matrix of the multivariate GHST distribution because it is not available in closed form.

As a practical hint, researchers may also consider working with the simpler expression

$$\mathcal{S}_t = \left\{ \Psi_t' (\tilde{\Sigma}_t^{-1} \otimes \tilde{\Sigma}_t^{-1}) \Psi_t \right\}^{-1}, \tag{B.2}$$

which takes similar values as (B.1) but does not require the repeated multiplication of large matrices of dimension $N^2 \times N^2$. The information theoretic optimality results for score-driven models derived in Blasques, Koopman, and Lucas (2015) depend on the properties of the score, and apply in either case.

Appendix C: The GHST multi-block equicorrelation model

This appendix derives the score updating equation for the m-block equicorrelation model, BEq[m].

We assume that N firms can be divided up into m different groups. There are n_i firms within each group $i = 1, \ldots, m$. Firms have equicorrelation ρ_i^2 within each group, and $\rho_i \cdot \rho_j$ between groups i and j, $j = i+1, \ldots, m$. The BEq[m] model proposed here differs from Engle and Kelly (2012) in that there is a direct relation between the equicorrelation in the off-diagonal blocks and the diagonal blocks. We impose this restriction to maintain the factor copula structure that permits the quick computations of the joint and conditional risk measures as discussed in Section 3 of the main paper.

In the dynamic GHST m-block equicorrelation case the scale matrix at time t is given by

$$\tilde{\Sigma}_t = \begin{pmatrix} (1-\rho_{1,t}^2)\mathrm{I}_1 & \ldots & \ldots & 0 \\ 0 & (1-\rho_{2,t}^2)\mathrm{I}_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & (1-\rho_{m,t}^2)\mathrm{I}_m \end{pmatrix} + \begin{pmatrix} \rho_{1,t}\ell_1 \\ \rho_{2,t}\ell_2 \\ \vdots \\ \rho_{m,t}\ell_m \end{pmatrix} \cdot \begin{pmatrix} \rho_{1,t}\ell_1 \\ \rho_{2,t}\ell_2 \\ \vdots \\ \rho_{m,t}\ell_m \end{pmatrix}', \qquad \text{(C.1)}$$

where I_i is an $n_i \times n_i$ identity matrix, $\ell_i \in \mathbb{R}^{n_i \times 1}$ is a column vector of ones, and $|\rho_{it}| < 1$. The block structure of the matrix allows us to obtain analytical solutions for the determinant,

$$\det(\tilde{\Sigma}_t) = \left[1 + \frac{n_1\rho_{1,t}^2}{1-\rho_{1,t}^2} + \cdots + \frac{n_m\rho_{m,t}^2}{1-\rho_{m,t}^2}\right](1-\rho_{1,t}^2)^{n_1}\cdots(1-\rho_{m,t}^2)^{n_m}, \qquad \text{(C.2)}$$

see Harville (2008). This matrix determinant result further facilitates the computation of the likelihood and the score steps in high dimensions.

We assume that y_t follows a GHST distribution with time-varying scale matrix $\tilde{\Sigma}_t$ (C.1). Factor f_t is an $m \times 1$ vector such that $\rho_i = (1 + \exp(-f_{it}))^{-1}$ for $i = 1, 2, \cdots, m$. It follows that

$$
\begin{aligned}
\tilde{\Psi}_t &= \frac{\partial \text{vec}(\tilde{\Sigma}_t)'}{\partial f_t} = \frac{\partial \text{vec}(\tilde{\Sigma}_t)'}{\partial \rho_t}\frac{\text{d}\rho_t'}{\text{d}f_t}, \\
\frac{\text{d}\rho_t'}{\text{d}f_t} &= \begin{bmatrix} \frac{\exp(-f_{1,t})}{(1+\exp(-f_{1,t}))^2} & 0 & \dots & 0 \\ 0 & \frac{\exp(-f_{2,t})}{(1+\exp(-f_{2,t}))^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\exp(-f_{m,t})}{(1+\exp(-f_{m,t}))^2} \end{bmatrix}, \\
\frac{\partial \text{vec}(\tilde{\Sigma}_t)}{\partial \rho_t{}'} &= \left[\text{vec}\begin{pmatrix} \text{I}_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \cdots, \text{vec}\begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \text{I}_m \end{pmatrix} \right] \cdot \begin{pmatrix} -2\rho_{1,t} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -2\rho_{2,t} \end{pmatrix} \\
&\quad + \left[\begin{pmatrix} \rho_{1,t}\ell_1 \\ \rho_{2,t}\ell_2 \\ \vdots \\ \rho_{m,t}\ell_m \end{pmatrix} \otimes \text{I}_N + \text{I}_N \otimes \begin{pmatrix} \rho_{1,t}\ell_1 \\ \rho_{2,t}\ell_2 \\ \vdots \\ \rho_{m,t}\ell_m \end{pmatrix} \right] \cdot \left[\begin{pmatrix} \ell_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \ell_{n_m} \end{pmatrix} \right].
\end{aligned}
$$

To derive this result we used (B), the fact that

$$
\frac{\partial \text{vec}(\tilde{\Sigma}_t)'}{\partial \rho_{it}} = -2\rho_{it} \cdot \text{vec}\begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \text{I}_i & 0 \\ 0 & 0 & \dots & 0 \end{bmatrix} + \left[\begin{pmatrix} \rho_{1,t}\ell_1 \\ \rho_{2,t}\ell_2 \\ \vdots \\ \rho_{m,t}\ell_m \end{pmatrix} \otimes \text{I}_N + \text{I}_N \otimes \begin{pmatrix} \rho_{1,t}\ell_1 \\ \rho_{2,t}\ell_2 \\ \vdots \\ \rho_{m,t}\ell_m \end{pmatrix} \right] \cdot \begin{pmatrix} 0 \\ \vdots \\ \ell_{n_i} \\ 0 \end{pmatrix}, \tag{C.3}
$$

and combined expression C.3 across the m groups.

For the multi-block model, also the two systemic risk indicators take a slightly different form. The numerical integrations and the computation $\kappa^*_{N,t}(\bar{c}, \varsigma^{(g)})$ can still be done in the same efficient way as described in the main text. We have

$$
D_{N,t} \approx \frac{1}{N}\sum_{i=1}^{N} \text{E}[1\{y_{it} < y^*_{it}\} \mid \kappa_t, \varsigma_t] = \frac{1}{N}\sum_{i=1}^{N} \text{P}[y_{it} < y^*_{it} \mid \kappa_t, \varsigma_t] \equiv C_{N,t}, \tag{C.4}
$$

where

$$\mathrm{P}[y_{it} < y_{it}^* | \kappa_t, \varsigma_t] = \Phi\left(\frac{y_{it}^* - (\varsigma_t - \mu_\varsigma)\,\gamma_i - \sqrt{\varsigma_t}\,\rho_{i,t}\kappa_t}{\sqrt{\varsigma_t\;(1-\rho_{i,t}^2)}}\right). \tag{C.5}$$

Note that the correlations $\rho_{i,t}$ are now different across the different firms in the sample, and in particular on the group number j_i. Similarly, we have

$$\begin{aligned}
\mathrm{P}\Big(C_{N-1,t}^{(-i)} > \bar{c}^{(-i)} \;\Big|\; y_{it} < y_{it}^*\Big) &= p_{it}^{-1} \cdot \mathrm{P}(C_{N-1,t}^{(-i)} > \bar{c}^{(-i)}\;,\; y_{it} < y_{it}^*) \\
&= p_{it}^{-1} \cdot \int \mathrm{P}\left(\kappa_t < \kappa_{i,N-1,t}^*(\bar{c}^{(-i)}, \varsigma_t)\;,\; y_{it} < y_{it}^* \;\middle|\; \varsigma_t\right)\; p(\varsigma_t)\mathrm{d}\varsigma_t \\
&= p_{it}^{-1} \cdot \int \Phi_2(\kappa_{i,N-1,t}^*(\bar{c}^{(-i)}, \varsigma_t)\;,\; y_{it}^{**}(\varsigma_t)\;;\; \rho_{i,t})\; p(\varsigma_t)\mathrm{d}\varsigma_t,
\end{aligned} \tag{C.6}$$

where $\kappa_{i,N-1,t}^*(\bar{c}^{(-i)}, \varsigma_t)$ is the efficiently computed threshold value for κ_t if firm i is left out; see the main paper for more details.

Appendix D: Data summary statistics

Tables D.1 and D.2 provide sample summary statistics for our small-scale and large-scale studies, respectively.

Table D.1: Sample descriptive statistics, 10 firms

The table reports descriptive statistics for $N = 10$ weekly equity returns between January 1999 and August 2013. All excess kurtosis and skewness coefficients are significantly different from 0 at the 5% significance level.

	Median	Std.Dev.	Skewness	Kurtosis	Minimum	Maximum
Bank of Ireland	0.004	0.099	-0.614	14.736	-0.658	0.581
Banco Comercial Portugues	0.004	0.049	-0.375	6.753	-0.284	0.230
Santander	0.002	0.049	-0.558	7.289	-0.261	0.212
National Bank of Greece	0.003	0.061	-1.253	12.909	-0.475	0.292
UniCredito	0.001	0.084	-0.235	15.968	-0.705	0.613
BNP Paribas	0.002	0.056	-0.364	10.344	-0.367	0.339
Deutsche Bank	0.000	0.059	-0.489	16.145	-0.529	0.398
Dexia	0.008	0.087	-0.569	13.735	-0.529	0.494
Eerste Bank Group	0.002	0.061	-1.143	16.100	-0.552	0.296
ING	0.004	0.069	-1.319	13.358	-0.546	0.290

Table D.2: Sample summary statistics; 73 firms

The table reports sample moments across all 73 institutions included in the empirical application. For example, the row labeled '*skewness*' contains the mean and standard deviation of the skewness statistic across all 73 firms, followed by the minimum, 25th, 50th, and 75th percentile, and the maximum. The first row reports the number of time series observations per firm.

	Mean	Std. Dev.	Minimum	25%	50%	75%	Maximum
T_i	717	116	196	762	762	762	762
Std. Dev.	0.06	0.02	0.03	0.04	0.05	0.06	0.11
Skewness	-0.81	1.94	-12.06	-0.84	-0.41	-0.22	1.89
Kurtosis	18.05	38.47	3.69	6.77	10.13	14.26	266.26

Appendix E: The GHST vs symmetric-t model

Figure E.1 provides estimates of the joint tail risk probability and average conditional tail probability based on different copula model specifications. While the joint probability estimates are fairly invariant to moving from a GHST to symmetric-t dependence, the conditional probabilities are not. The negative skewness implies considerably higher conditional probabilities on average.

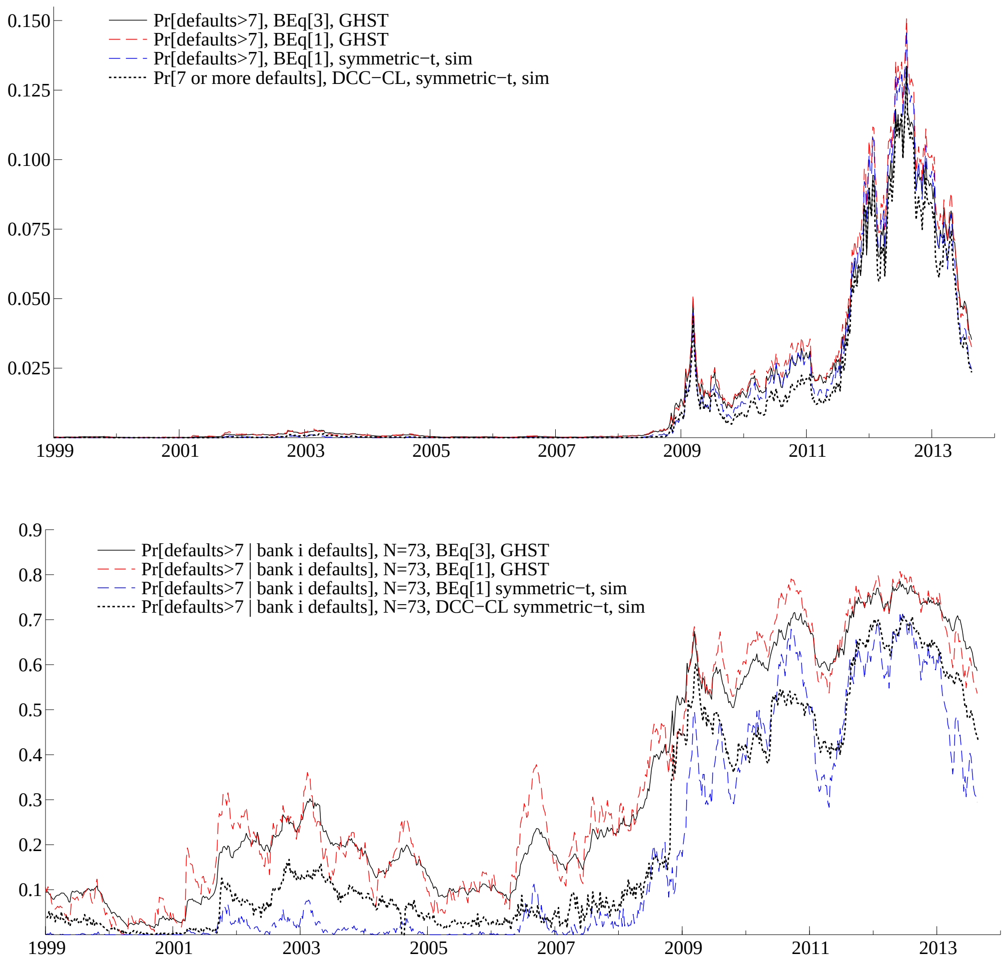

Figure E.1: GHST vs. symmetric-t estimates

Estimates of the joint tail risk probability (top) and average conditional tail probability (bottom) based on different copula model specifications. We provide two specifications of the GHST and symmetric-t copula, respectively. The estimation sample is from Jan 1999 – Oct 2013.

References

Blasques, F., S. J. Koopman, and A. Lucas (2015). Information theoretic optimality of observation driven time series models for continuous responses. *Biometrika 102*(2), 325–343.

Engle, R. and B. Kelly (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics 30*(2), 212–228.

Harville, D. (2008). *Matrix algebra from a statistician's perspective.* Springer Verlag.

Lucas, A., B. Schwaab, and X. Zhang (2015). Quantifying policy spillovers from unconventional monetary policies: Do central banks take or make their credit risks? *Unpublished working paper*.

Rodriguez, M. J. and E. Ruiz (2012). Revisiting several popular GARCH models with leverage effect: Differences and similarities. *Journal of Financial Econometrics 10(4)*, 637–668.