Appendix A

Marginal distribution model:

We specify a model for the conditional mean of the stock return series X_t as follows:

$$X_t = \mu + \sum_{i=1}^{m} \phi_i X_{t-i} + \sum_{j=1}^{n} \theta_j \varepsilon_{t-j} + \varepsilon_t, \tag{A.1}$$

where μ is a vector of constants, and ϕ_i and θ_j are the AR and the MA components with m and n lags, respectively. We assume that the white noise process ε_t follows the Student's *t*-distribution with v degrees of freedom given as:

$$\sqrt{\frac{v}{\sigma_t^2(v-2)}}\,\varepsilon_t \overset{\text{i.i.d.}}{\sim} t_v, \tag{A.2}$$

and a Glosten–Jagannathan–Runkle GARCH model of Glosten et al. (1993), GJR-GARCH(p,q), for the conditional variance of X_t, σ_t^2, is represented by the following equation:

$$\sigma_t^2 = \kappa + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 + \sum_{j=1}^{q} \left(\alpha_j \varepsilon_{t-j}^2 + \xi_j 1\left(\varepsilon_{t-j} < 0\right) \varepsilon_{t-j}^2\right), \tag{A.3}$$

where α_j and β_i are the ARCH and GARCH parameters, respectively, κ is the intercept of the conditional variance equation, $1\left(\varepsilon_{t-j} < 0\right)$ is an indicator function that takes one if ε_{t-j} is negative and zero otherwise, and ξ_j captures the leverage effect in the underlying series. The GJR-GARCH(p,q) captures the asymmetric effects of negative and positive shocks on the returns' conditional volatility. When $\xi_j < 0$, a negative shock has a greater impact on the future conditional variance than a positive shock of equal absolute magnitude. Thus, an ARMA(m,n)-GJR-GARCH(p,q) model is appropriate as investors react differently to positive and negative shocks in stock returns.

Appendix B

Copula models:

For n stock return series of FIs, the generalized form of a Gaussian copula is given by:

$$\begin{aligned} C(u_1, \dots, u_n) &= \phi_\rho\left(\phi^{-1}(u_1), \dots, \phi^{-1}(u_n)\right) \\ &= \int_{-\infty}^{\phi^{-1}(u_1)} \dots \int_{-\infty}^{\phi^{-1}(u_n)} \frac{1}{2(\pi)^{n/2}|\rho|^{1/2}} \exp\left(-\frac{1}{2}\, z^T \rho^{-1} z\right) dz_1, \dots, dz_n, \end{aligned} \quad \text{(B.1)}$$

for all $(u_1, \dots, u_n) \in \mathbb{R}^n$, where $\phi_\rho(\cdot, \dots, \cdot)$ is the multivariate Gaussian distribution function, ρ is the correlation matrix, u_i is the marginal distribution function of stock return i, and $\phi^{-1}(\cdot)$ is the inverse of the univariate Gaussian distribution. We also employ the Student's *t* copula as follows:

$$\begin{aligned} &C(u_1, \dots, u_n) = t_{\rho,v}\left(t_v^{-1}(u_1), \dots, t_v^{-1}(u_n)\right) \\ &= \int_{-\infty}^{t^{-1}(u_1)} \dots \int_{-\infty}^{t^{-1}(u_n)} \frac{1}{\Gamma\left(\frac{v}{2}\right)(v\pi)^{n/2}|\rho|^{1/2}} \left(1 + \frac{1}{v}\, z^T \rho^{-1} z\right)^{-\frac{v+n}{2}} dz_1, \dots, dz_n, \end{aligned} \quad \text{(B.2)}$$

for all $(u_1, \dots, u_n) \in \mathbb{R}^n$, where $t_{\rho,v}(\cdot, \dots, \cdot)$ is the multivariate Student's *t*-distribution, ρ is the correlation matrix, v is the number of degrees of freedom, and $t_v^{-1}(\cdot)$ is the inverse of the univariate Student's *t*-distribution with v degrees of freedom. The Student's *t* copula captures the variations in the distribution tails, and it accounts for possible joint extreme movements that characterize the financial return series. The Student's *t*-distribution converges to a Gaussian distribution as $v \to \infty$.

Following Patton (2006), we implement the Clayton copula, formulated as follows:

$$C_t^C(u_1, u_2;\, \tau_t) = \left(u_1^{-\tau_t} + u_2^{-\tau_t} - 1\right)^{-\frac{1}{\tau_t}}, \;\; \tau_t \in (0, \infty), \quad \text{(B.3)}$$

for all $(u_1, u_2) \in \mathbb{R}^2$, where τ_t follows the process

$$\tau_t = \Lambda\left(\omega + \beta\,\tau_{t-1} + \alpha \cdot \tfrac{1}{10}\textstyle\sum_{i=1}^{10}\left|u_{1,t-i} - u_{2,t-i}\right|\right), \tag{B.4}$$

where $\Lambda(z) = (1+e^{-z})^{-1}$ is the logistic function that guarantees that $\tau_t \in (0,1)$ for all t. To consider asymmetric tail dependence between the return series, we apply the Symmetrized Joe-Clayton (SJC) (Patton, 2006), which is specified as:

$$C^{SJC}(u_1,u_2|\tau_U,\tau_L) = 0.5[C^{JC}(u_1,u_2|\tau_U,\tau_L) + C^{JC}(1-u_1,1-u_2|\tau_U,\tau_L) + u_1 + u_2 - 1], \tag{B.5}$$

where $C^{JC}(u_1,u_2|\tau_U,\tau_L) = 1 - \left\{1 - [(1-(1-u_1)^{\kappa})^{-\gamma} + (1-(1-u_2)^{\kappa})^{-\gamma} - 1]^{-1/\gamma}\right\}^{1/k}$ is the Joe-Clayton copula with $\kappa = 1/\log_2(2-\tau_U)$, $\gamma = -1/\log_2(\tau_L)$, $\tau_U \in (0,1)$, and $\tau_L \in (0,1)$. The parameters τ_U and τ_L assess the dependence at the upper and lower tails of the distribution, respectively. For $\tau_U = \tau_L$, the SJC dependence structure is symmetric; otherwise, it is asymmetric.

Patton (2006) defined the evolution of dependence parameters of the SJC copula as:

$$\tau_{j,t} = \Lambda\left(\omega_j + \beta_j\,\tau_{j,t-1} + \alpha_j \cdot \tfrac{1}{10}\textstyle\sum_{i=1}^{10}\left|u_{1,t-i} - u_{2,t-i}\right|\right), \tag{B.6}$$

with $j = \{U, L\}$, and $\Lambda(z) = (1+e^{-z})^{-1}$ is the logistic function that guarantees that $\tau_{j,t} \in (0,1)$ for all t.

Let $H_t = E_{t-1}(\boldsymbol{u}_t\boldsymbol{u}_t')$, with $\boldsymbol{u}_t = (u_{1,t}, u_{2,t})'$, be the $(T \times T)$ matrix of conditional variance-covariance of returns, which can also be written as:

$$H_t = D_t R_t D_t,$$

where R_t is the $(T \times T)$ conditional correlation matrix, $D_t = diag\left[\sqrt{h_{1,t}}, \sqrt{h_{2,t}}\right]$ with $h_{i,t} = E_{t-1}(u_{i,t}^2)$ and $u_{i,t} = \sqrt{h_{i,t}}\varepsilon_{i,t}$, for i=1, 2. Then, $\varepsilon_{i,t}$ is a standardized error with mean zero and variance one. Let $\boldsymbol{\epsilon}_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ be a vector of standardized disturbances. We estimate time-varying linear correlations $\rho_{i,j,t}$ by applying the dynamic conditional correlation (DCC) method proposed by Engle (2002) as follows:

$$R_t \equiv \mathrm{diag}(Q_t)^{-1} H_t \, \mathrm{diag}(Q_t)^{-1} = E_{t-1}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t'),$$

$$Q_t \equiv [q_{i,j,t}] = \bar{R}(1-a-b) + a\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}' + bQ_{t-1}, \tag{B.7}$$

where R_t has elements $\rho_{i,j,t} = q_{i,j,t}/\sqrt{q_{ii,t}q_{jj,t}}$, $\bar{R} = E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t{}')$ is the matrix of unconditional correlation of the returns, and a and b satisfy the restrictions $a, b \in (0,1)$ with $a + b < 1$. Following Joe (1997), we apply the two-step maximum likelihood procedure to estimate the marginal models and the copula density. First, we estimate the GARCH marginal parameters $\widehat{\boldsymbol{\theta}}_1$ by fitting univariate marginal distributions that solve:

$$\widehat{\boldsymbol{\theta}}_1 = \underset{\boldsymbol{\theta}_1}{\mathrm{argmax}} \sum_{t=1}^{T} \sum_{j=1}^{n} \ln f_j(u_{j,t};\, \boldsymbol{\theta}_1), \tag{B.8}$$

where $\ln f_j(u_{j,t};\, \boldsymbol{\theta}_1)$ is the log-likelihood of the j-th FI stock return, and $\widehat{\boldsymbol{\theta}}_1$ is a $n \times 1$ vector of maximum likelihood estimates of the GARCH marginal parameters. In the second step, given the vector $\widehat{\boldsymbol{\theta}}_1$ from Eq. (B.8), we compute the DCC copula parameters, $\widehat{\boldsymbol{\theta}}_2$, as follows:

$$\widehat{\boldsymbol{\theta}}_2 = \underset{\boldsymbol{\theta}_2}{\mathrm{argmax}} \sum_{t=1}^{T} \ln c\big(F_1(u_{1,t}), F_2(u_{2,t}), \dots, F_n(u_{n,t});\, \boldsymbol{\theta}_2, \widehat{\boldsymbol{\theta}}_1\big). \tag{B.9}$$

Appendix C

Systemic risk across different frequencies:

We decompose the underlying return series into wavelet components to evaluate the VaR and ΔCoVaR across different investment horizons. The wavelet method is based on a Fourier representation of a series on its frequencies. Since the Fourier transform loses the time information, a Fourier representation can be implemented on a rolling window, a wavelet, to recover both time and scale information (Percival & Walden, 2000). We can write a wavelet transform $\psi_{\tau,s}(t)$ with time translation τ and scale s as:

$$\psi_{\tau,s}(t) = s^{-1/2}\psi\frac{(t-s)}{s},$$

for $s, \tau \in \mathbb{R}$, $s \neq 0$, and a mother wavelet $\psi(t)$. Since our return series are discrete, we employ a discrete wavelet transform on the data with length 2^J. We can represent a discrete wavelet transform with time translation $\tau = 2n$ and scale $s = 2J$ as:

$$\psi_{J,n}(t) = (2^J)^{-1/2}\psi\frac{(t-2^n)}{2^n},$$

for $J, n \in \mathbb{Z}$. The discrete wavelet transform (DWT) is a band-pass filter that recovers the frequencies around its main frequency, generating a scaling filter ϕ. We define the vector of J wavelet filter ψ elements as $(h_0, \dots, h_{J-1})$ such that $(i)\sum_{j=0}^{J-1} h_j = 0$, $(ii)\sum_{j=0}^{J-1} h_j^2 = 1$, and $(iii)\sum_{j=0}^{J-1} h_j h_{j+2k} = 0$, with $k \neq 0$. The properties (i)-(iii) of the wavelet filter elements imply an orthogonal matrix of the discrete wavelet transform. The scaling filter ϕ elements $(g_0, \dots, g_{J-1})$ also satisfy the properties (i)-(iii).

For a FI stock returns series $\mathbf{x} = \{X_t\}_{t=1}^{T}$, we can apply the DWT to obtain the vector of wavelet coefficients as $\mathbf{w} = \mathbf{W}\mathbf{x}$, where $\mathbf{w} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_J, \mathbf{v}_J]'$ is a $(J+1)T \times 1$ vector of wavelet coefficients with $T = 2^J$, and the DWT matrix $\mathbf{W}_{(J+1)T\times T}$ specifies the transform (Gençay et al., 2005). Nevertheless, the DWT is restricted to the number of observations of $\mathbf{x}$ being a multiple of 2^J.

A maximum overlap discrete wavelet transform (MODWT) is an extension of the wavelet transform that is indifferent to the number of observations of **x**, and the estimator of the MODWT is asymptotically more efficient than that of the DWT (Percival & Walden, 2000). Let $h_{j,k}$ and $g_{j,k}$ be an element of the wavelet filter ψ and scaling filter ϕ, respectively. We can write an element of the vector of K wavelet filter and scaling filter elements of the MODWT as $(\tilde{h}_0, \dots, \tilde{h}_{J-1})$ and $(\tilde{g}_0, \dots, \tilde{g}_{J-1})$, with

$$\tilde{h}_{j,k} = \left(2^j\right)^{-1/2} h_{j,k} \quad \text{and} \quad \tilde{g}_{j,k} = \left(2^j\right)^{-1/2} g_{j,k}, \tag{C.1}$$

such that $(i) \sum_{j=0}^{J-1} \tilde{h}_{j,k} = 0$, $(ii) \sum_{j=-\infty}^{J-1} \tilde{h}^2{}_{J,k} = (1/2^J)$, and $(iii) \sum_{j=0}^{J-1} \tilde{h}_{j,k} \tilde{h}_{j,k+2} = 0$. We can apply the MODWT to obtain the vector of wavelet coefficients as $\tilde{\mathbf{w}} = \widetilde{\mathbf{W}}\mathbf{x}$, where $\tilde{\mathbf{w}} = [\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \dots, \tilde{\mathbf{w}}_J, \tilde{\mathbf{v}}_J]'$ is a $(J+1)T \times 1$ vector of wavelet coefficients, $\tilde{\mathbf{w}}_j$ are vectors of wavelet coefficients with length $T/2^j$ of the scale of width $a_j = 2^{j-1}$, $\tilde{\mathbf{v}}_j$ are scaling vectors with length 2^j, and the DWT matrix $\widetilde{\mathbf{W}}_{(J+1)T \times T}$ specifies the transform.

We can employ the MODWT to obtain an additive wavelet approximation of the return series. We define $\widetilde{\mathbf{D}}_j = \widetilde{\mathbf{W}}_j{}' \tilde{\boldsymbol{w}}_j$ as the wavelet detail for the MODWT of variations in the returns **x** at the scale a_j with levels $j = 1, 2, \dots, J$. Then, we can write the multiscale decomposition for each return X_t as

$$X_t = \sum_{j=1}^{J+1} \widetilde{\mathrm{D}}_{j,t}\, , t = 1, 2, \dots, T-1,$$

where $\widetilde{\mathrm{D}}_{j,t}$ is the *t*-th element of $\widetilde{\mathbf{D}}_j$. Let $\widetilde{\mathbf{A}}_j = \sum_{k=j+1}^{J+1} \widetilde{\mathbf{D}}_k$ be the MODWT wavelet approximation for $0 \leq j \leq J$. Then, we can decompose the vector of returns **x** into as follows:

$$\mathbf{x} = \widetilde{\mathbf{A}}_j + \sum_{k=1}^{j} \widetilde{\mathbf{D}}_k, \tag{C.2}$$

where $\widetilde{\mathbf{D}}_k$ are the decomposed signals used for further analysis to evaluate the risk spillover across varying frequencies.

The choice of the wavelet filter class in the MODWT is important to determine the frequency variation between scales in the data since wavelet basis functions need to represent the stylized features of the data. Gençay et al. (2001) suggest using a wavelet filter with a balanced length (such as length eight) that recovers the main characteristics of financial returns. We employ the Daubechies (1992)'s least-asymmetric wavelet filter with length eight, LA(8), because it counterbalances length, symmetry, and smoothness (Gençay et al., 2001). Besides, the LA(8) wavelet filter of Daubechies (1992) has been adopted in many empirical applications in finance and economics (Gençay et al., 2005; Bekiros and Marcellino, 2013).

Following Bekiros and Marcellino (2013), we implement a periodic extension pattern of the MODWT to consider boundary estimation problems. We employ the MODWT wavelet approximation on the underlying returns to evaluate the VaR and ΔCoVaR for various investment horizons. More specifically, due to heterogeneous investor's behavior and time-horizon of investment, we transform the return series into the short-, medium-, and long-term horizons that correspond to variations over 2-4 days, 32-64 days, and 256-512 days, respectively. Then, we estimate the VaR and ΔCoVaR for each subsequent wavelet.

Table A.1. Details of the Australian financial institutions

Financial institutions	Acronym	DataStream industry classification	DataStream code
	Major banks		
Australia and New Zealand	ANZ	Banks	A:ANZX
Commonwealth Bank of Australia	CBA	Banks	A:CBAX
National Australia Bank	NAB	Banks	A:NABX
WestPac Banking	WBC	Banks	A:WBCX
	Regional banks		
Auswide Bank	ABA	Banks	A:ABAX
Bendigo and Adelaide Bank	BEN	Banks	A:BENX
Bank of Queensland	BOQ	Banks	A:BOQX
	Insurance companies		
Suncorp Group	SUN	Life insurance	A:SUNX
QBE Insurance Group	QBE	Nonlife insurance	A:QBEX
	Other financial service providers		
AMP	AMP	Financial services	A:AMPX
ASX	ASX	Financial services	A:ASXX
Challenger	CGF	Financial services	A:CGFX
Computershare	CPU	Financial services	A:CPUX
EQT Holdings	EQT	Financial services	A:EQTX
Euroz	EZL	Financial services	A:EZLX
First Graphene	FGR	Financial services	A:FGRX
Macquarie Group	MQG	Financial services	A:MQG
Pacific Current Group	PAC	Financial services	A:PACX
Perpetual	PPT	Financial services	A:PPTX
Premier Investment	PMV	Financial services	A:PMVX

References

Bekiros, S., & Marcellino, M. (2013). The multiscale causal dynamics of foreign exchange markets. *Journal of International Money and Finance*, *33*, 282–305.

Bekiros, S., Nguyen, D. K., Uddin, G. S., & Sjö, B. (2016). On the time scale behavior of equity-commodity links: Implications for portfolio management. *Journal of International Financial Markets, Institutions and Money*, *41*, 30–46.

Daubechies, I. (1992). Ten Lectures of Wavelets. In *Siam*. Society for Industrial and Applied Mathematics.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, *20*, 339–350.

Gençay, R., Selçuk, F., & Whitcher, B. (2005). Multiscale systematic risk. *Journal of International Money and Finance*, *24*, 55–70.

Gençay, R., Selçuk, F., & Whitcher, B. J. (2001). *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Elsevier.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*, 1779–1801.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. Chapman & Hall.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, *47*, 527–556.

Percival, D. B., & Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge.