1 Appendix A: Link Matrices

Careful construction of the link matrices used in (2.7) is critical in the development of the GVAR. For our purposes, and employing the country ordering presented in Table 1, the $\boldsymbol{W}_i$'s are given by:

$$\underset{10\times 177}{\boldsymbol{W}_0} = \begin{pmatrix} \boldsymbol{R}_{00} & \boldsymbol{0}_{7\times 7} & \cdots & \boldsymbol{0}_{7\times 7} & \boldsymbol{0}_{7\times 6} & \cdots & \boldsymbol{0}_{7\times 6} & \boldsymbol{0}_{7\times 5} \\ \boldsymbol{0}_{3\times 8} & \boldsymbol{W}_{01} & \cdots & \boldsymbol{W}_{0,20} & \boldsymbol{W}_{0,21} & \cdots & \boldsymbol{W}_{0,24} & \boldsymbol{W}_{0,25} \end{pmatrix},$$

$$\underset{12\times 177}{\boldsymbol{W}_i} = \begin{pmatrix} \boldsymbol{R}_{i0} & \boldsymbol{R}_{i1} & \boldsymbol{R}_{i2} & \cdots & \boldsymbol{R}_{i,25} \\ \boldsymbol{W}_{i0} & \boldsymbol{W}_{i1} & \boldsymbol{W}_{i2} & \cdots & \boldsymbol{W}_{i,25} \end{pmatrix}, \ i = 1, ..., 25,$$

where

$$\boldsymbol{R}_{00} = \begin{bmatrix} \boldsymbol{0}_{7\times 1} & \boldsymbol{I}_7 \end{bmatrix}, \ \boldsymbol{R}_{i0} = \begin{bmatrix} -w_{i0} & \boldsymbol{0}_{1\times 7} \\ \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 7} \end{bmatrix}, \ i = 1, ..., 25,$$

$$\{\boldsymbol{R}_{ij}\}_{j=1}^{20} = \left\{ \begin{array}{cc} \begin{bmatrix} -w_{ij} & \boldsymbol{0}_{1\times 6} \\ \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 6} \end{bmatrix} & \text{if } j \neq i \\ \boldsymbol{I}_7 & \text{if } j = i \end{array} \right\}, \ i = 1, ..., 25,$$

$$\{\boldsymbol{R}_{ij}\}_{j=21}^{24} = \left\{ \begin{array}{cc} \begin{bmatrix} -w_{ij} & \boldsymbol{0}_{1\times 5} \\ \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 5} \end{bmatrix} & \text{if } j \neq i \\ \boldsymbol{I}_6 & \text{if } j = i \end{array} \right\}, \ i = 1, ..., 25,$$

$$\boldsymbol{R}_{i,25} = \left\{ \begin{array}{cc} \begin{bmatrix} -w_{i,25} & \boldsymbol{0}_{1\times 4} \\ \boldsymbol{0}_{6\times 1} & \boldsymbol{0}_{6\times 4} \end{bmatrix} & \text{if } i \neq 25 \\ \boldsymbol{I}_5 & \text{if } i = 25 \end{array} \right\},$$

$$\{\boldsymbol{W}_{0j}\}_{j=1}^{20} = \begin{bmatrix} w_{0j} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{0j} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{0j} \end{bmatrix},$$

$$\{\boldsymbol{W}_{0j}\}_{j=21}^{24} = \begin{bmatrix} w_{0j} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{0j} & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{0j} \end{bmatrix}, \ \boldsymbol{W}_{0,25} = \begin{bmatrix} w_{0,25} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{0,25} & 0 \\ 0 & 0 & 0 & 0 & w_{0,25} \end{bmatrix},$$

and for $i = 1, ..., 25$,

$$\boldsymbol{W}_{i0} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & w_{i0}^* & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{i0}^{**} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{i0} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{i0} \end{bmatrix}, \ \{\boldsymbol{W}_{ij}\}_{j=1}^{20} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{ij}^* & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{ij}^{**} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{ij} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{ij} \end{bmatrix},$$

$$\{\boldsymbol{W}_{ij}\}_{j=21}^{24} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{ij}^* & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{ij} & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{ij} \end{bmatrix}, \ \boldsymbol{W}_{i,25} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{i,25} & 0 \\ 0 & 0 & 0 & 0 & w_{i,25} \end{bmatrix}.$$

where w_{ij} is the weight of country i in the trade of country j, w_{ij}^* is the ith country's adjusted trade-weight with the jth country after allowing for the lack of Saudi interest rate data, and w_{ij}^{**} is the ith country's trade-weight with the jth country adjusted to accommodate the lack of reliable stock market data for China, Indonesia, Peru, Turkey and Saudi Arabia. Notice that $\sum_{j=0}^{N} w_{ij} = \sum_{j=0}^{N} w_{ij}^* = \sum_{j=0}^{N} w_{ij}^{**} = 1$, and $w_{ii} = w_{ii}^* = w_{ii}^* = 0$ for all i.

We follow DdPS in our construction of the 26×26 trade-weighted link matrices based on trade averages over the period 1999-2001. Initial experimentation with the model using more recent windows including 2001-2003 yielded qualitatively similar results. We also considered the use of time-varying trade weights, although as DdPS note, one must be careful not to introduce an undesirable element of

randomness into estimation in this manner. In general, the results subject to the time-varying structure were somewhat similar to those derived from the static link matrices and we could find no firm basis on which to discriminate positively between them. Ultimately, our decision to use the same fixed trade-weighting scheme as DdPS is motivated by a desire to maintain comparability with their results.

2 Appendix B: Derivations of PP, GIRF and GFEVD in the GVAR Model

The reduced-form GVAR model in our paper is written as:

$$\boldsymbol{x}_t = \boldsymbol{g}_0^* + \boldsymbol{g}_1 t + \boldsymbol{G}_1 \boldsymbol{x}_{t-1} + \boldsymbol{G}_2 \boldsymbol{x}_{t-2} + \boldsymbol{G}_3 \boldsymbol{x}_{t-3} + \boldsymbol{\varepsilon}_t, \tag{1}$$

Derivations of PP, GIRF and GFEVD are based on an infinite order MA representation of the GVAR model, (1):

$$\boldsymbol{x}_t = \boldsymbol{d}_t + \sum_{j=0}^{\infty} \boldsymbol{B}_j \boldsymbol{\varepsilon}_{t-j}, \tag{2}$$

where $\boldsymbol{d}_t$ represents the (perfectly forecastable) deterministic component of $\boldsymbol{x}_t$ and $\boldsymbol{B}_j$ can be evaluated recursively as

$$\boldsymbol{B}_j = \boldsymbol{G}_1 \boldsymbol{B}_{j-1} + \boldsymbol{G}_2 \boldsymbol{B}_{j-2} + \boldsymbol{G}_3 \boldsymbol{B}_{j-3},\ j = 1, 2, \text{ with } \boldsymbol{B}_0 = \boldsymbol{I}_m,\ \boldsymbol{B}_j = \boldsymbol{0} \text{ for } j < 0.$$

2.1 Persistence Profiles

The cointegrating relationships are estimated and given in terms of the country–specific variables, $\boldsymbol{\beta}_i' \boldsymbol{z}_{it}$, whereas the variables in the GVAR are given by $\boldsymbol{x}_t$. Hence, appropriate mapping between $\boldsymbol{z}_{it}$ and $\boldsymbol{x}_t$ is necessary. Using (2.7) and (2.9) we have

$$\begin{aligned} \boldsymbol{z}_{it} &= \boldsymbol{W}_i \tilde{\boldsymbol{x}}_t = \boldsymbol{W}_i \left(\boldsymbol{S}_0 \boldsymbol{x}_t - \boldsymbol{S}_1 \boldsymbol{x}_{t-1} \right) \\ &= \boldsymbol{W}_i \left(\boldsymbol{S}_0 \boldsymbol{d}_t - \boldsymbol{S}_1 \boldsymbol{d}_{t-1} \right) + \boldsymbol{W}_i \boldsymbol{S}_0 \boldsymbol{\varepsilon}_t + \sum_{j=0}^{\infty} \boldsymbol{W}_i \left(\boldsymbol{S}_0 \boldsymbol{B}_j - \boldsymbol{S}_1 \boldsymbol{B}_{j-1} \right) \boldsymbol{\varepsilon}_{t-j}. \end{aligned}$$

Therefore, the PP of $\boldsymbol{\beta}_{ji}' \boldsymbol{z}_{it}$ with respect to a system–wide shock to $\boldsymbol{\varepsilon}_t$ is given by

$$PP\left(\boldsymbol{\beta}_{ji}' \boldsymbol{z}_{it}; \boldsymbol{\varepsilon}_t, n \right) = \frac{\boldsymbol{\beta}_{ji}' \boldsymbol{W}_i \boldsymbol{C}_n \boldsymbol{\Sigma}_\varepsilon \boldsymbol{C}_n' \boldsymbol{W}_i' \boldsymbol{\beta}_{ji}}{\boldsymbol{\beta}_{ji}' \boldsymbol{W}_i \boldsymbol{C}_0 \boldsymbol{\Sigma}_\varepsilon \boldsymbol{C}_0' \boldsymbol{W}_i' \boldsymbol{\beta}_{ji}},\ n = 0, 1, 2, ... \tag{3}$$

where $\boldsymbol{\beta}_{ji}$ is the jth cointegrating relationship in the ith country $(j = 1, ..., r_i)$, n is the forecasting horizon, $\boldsymbol{\Sigma}_\varepsilon$ is the covariance matrix of $\boldsymbol{\varepsilon}_t$ and

$$\boldsymbol{C}_0 = \boldsymbol{S}_0 \boldsymbol{B}_0 = \boldsymbol{S}_0 \text{ and } \boldsymbol{C}_n = \boldsymbol{S}_0 \boldsymbol{B}_n - \boldsymbol{S}_1 \boldsymbol{B}_{n-1},\ n = 1, 2, ...$$

Similarly, the PP of $\boldsymbol{\beta}_{ji}' \boldsymbol{z}_{it}$ with respect to a variable–specific shock, say the ℓth element of $\boldsymbol{x}_t$, is given by

$$PP\left(\boldsymbol{\beta}_{ji}' \boldsymbol{z}_{it}; \varepsilon_{\ell t}, n \right) = \frac{\boldsymbol{\beta}_{ji}' \boldsymbol{W}_i \boldsymbol{C}_n \boldsymbol{\Sigma}_\varepsilon \boldsymbol{e}_\ell}{\sqrt{\sigma_{\ell\ell}}},\ n = 0, 1, 2, ... \tag{4}$$

where $\sigma_{\ell\ell}$ is the ℓth diagonal element of $\boldsymbol{\Sigma}_\varepsilon$ and $\boldsymbol{e}_\ell$ is an $m \times 1$ vector with its ℓth in $\boldsymbol{x}_t$ being unity and zeros elsewhere.

2.2 Impulse Responses

The generalised impulse response function (GIRF) of a unit (one standard error) shock to the ℓth element of $\boldsymbol{x}_t$ on the jth element of $\boldsymbol{x}_t$ is given by

$$GIRF\left(x_{jt};u_{\ell t},n\right)=\frac{\boldsymbol{e}_j'\boldsymbol{B}_n\boldsymbol{F}_0^{-1}\boldsymbol{\Sigma}_u\boldsymbol{e}_\ell}{\sqrt{\boldsymbol{e}_\ell'\boldsymbol{\Sigma}_u\boldsymbol{e}_\ell}},\ n=0,1,2,...,j,\ell=1,...,m. \tag{5}$$

The above expression can be used to compute the effects of shocking (displacing) a given endogenous variable in country i on all the variables in the global economy at different horizons.[1] For a structurally identified shock, $v_{\ell t}$, (such as a US monetary policy shock) the GIRF is given by

$$SGIRF\left(x_{jt};v_{\ell t},n\right)=\frac{\boldsymbol{e}_j'\boldsymbol{B}_n\left(\boldsymbol{P}\boldsymbol{F}_0\right)^{-1}\boldsymbol{\Sigma}_v\boldsymbol{e}_\ell}{\sqrt{\boldsymbol{e}_\ell'\boldsymbol{\Sigma}_v\boldsymbol{e}_\ell}},\ n=0,1,2,...,j,\ell=1,...,m, \tag{6}$$

where $\boldsymbol{\Sigma}_v$ is the covariance matrix of the structural shocks and $\boldsymbol{P}\boldsymbol{F}_0$ is defined by the identification scheme used to identify the shocks. For example, for identification of a US monetary policy shock using the triangular approach of Sims (1980), starting with the US model,

$$\boldsymbol{x}_{0t}=\boldsymbol{h}_{00}^{*}+\boldsymbol{h}_{01}t+\boldsymbol{\Phi}_{01}\boldsymbol{x}_{0,t-1}+\boldsymbol{\Phi}_{02}\boldsymbol{x}_{0,t-2}+\boldsymbol{\Psi}_{00}\boldsymbol{x}_{0t}^{*}+\boldsymbol{\Psi}_{01}\boldsymbol{x}_{0,t-1}^{*}+\boldsymbol{\Psi}_{02}\boldsymbol{x}_{0,t-2}^{*}+\boldsymbol{u}_{0t}, \tag{7}$$

the structural shocks are identified by

$$\boldsymbol{v}_{0t}=\boldsymbol{P}_0\boldsymbol{u}_{0t},$$

where $\boldsymbol{P}_0$ is a lower triangular matrix obtained as the $m_0\times m_0$ Choleski factor of $\boldsymbol{\Sigma}_{u_0}$, such that $\boldsymbol{\Sigma}_{u_0}=\boldsymbol{P}_0\boldsymbol{P}_0'$. Pre-multiplying the GVAR model (1) by

$$\boldsymbol{P}=\begin{bmatrix}\boldsymbol{P}_0 & \boldsymbol{0} & & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}_{m_1} & & \boldsymbol{0}\\ & & \ddots & \\ \boldsymbol{0} & \boldsymbol{0} & & \boldsymbol{I}_{m_N}\end{bmatrix},$$

it follows that (abstracting from deterministic elements)

$$\boldsymbol{P}\boldsymbol{F}_0\boldsymbol{x}_t=\boldsymbol{P}\boldsymbol{F}_1\boldsymbol{x}_{t-1}+\boldsymbol{P}\boldsymbol{F}_2\boldsymbol{x}_{t-2}+\boldsymbol{P}\boldsymbol{F}_3\boldsymbol{x}_{t-3}+\boldsymbol{v}_t,$$

where

$$\boldsymbol{v}_t=\begin{bmatrix}\boldsymbol{v}_{0t}\\ \boldsymbol{u}_{1t}\\ \\ \boldsymbol{u}_{Nt}\end{bmatrix},\ \boldsymbol{\Sigma}_v=\begin{pmatrix}Var\left(\boldsymbol{v}_{0t}\right) & Cov\left(\boldsymbol{v}_{0t},\boldsymbol{u}_{1t}\right) & & Cov\left(\boldsymbol{v}_{0t},\boldsymbol{u}_{Nt}\right)\\ Cov\left(\boldsymbol{u}_{1t},\boldsymbol{v}_{0t}\right) & Var\left(\boldsymbol{u}_{1t}\right) & & Cov\left(\boldsymbol{u}_{1t},\boldsymbol{u}_{Nt}\right)\\ & & \ddots & \\ Cov\left(\boldsymbol{u}_{Nt},\boldsymbol{v}_{0t}\right) & Cov\left(\boldsymbol{u}_{Nt},\boldsymbol{u}_{1t}\right) & & Var\left(\boldsymbol{u}_{Nt}\right)\end{pmatrix}.$$

By using the definition of GIRF with respect to the structural shocks, we have

$$SGIRF\left(\boldsymbol{x}_t;v_{\ell t},n\right)=E\left(\boldsymbol{x}_{t+n}|\Omega_{t-1},\boldsymbol{v}_{\ell t}=\sqrt{\boldsymbol{e}_\ell'\boldsymbol{\Sigma}_v\boldsymbol{e}_\ell}\right)-E\left(\boldsymbol{x}_{t+n}|\Omega_{t-1}\right),$$

then (6) readily follows (see also DdPS).

2.3 Forecast Error Variance Decomposition

The generalised FEVD of shocks to specific variables is then given by

$$GFEVD\left(x_{\ell t};u_{jt},n\right)=\frac{\sigma_{u,jj}^{-1}\sum_{h=0}^{n}\left(\boldsymbol{e}_\ell'\boldsymbol{B}_h\boldsymbol{F}_0^{-1}\boldsymbol{\Sigma}_u\boldsymbol{e}_j\right)^2}{\sum_{h=0}^{n}\boldsymbol{e}_\ell'\boldsymbol{B}_h\boldsymbol{F}_0^{-1}\boldsymbol{\Sigma}_u\boldsymbol{F}_0^{-1\prime}\boldsymbol{B}_h'\boldsymbol{e}_\ell},\ n=0,1,2,...,\ \ell=1,...,m, \tag{8}$$

[1]Note that the PP or GIRF of a unit shock to the US price level are the same as those of a shock to US inflation.

which gives the proportion of the n–step ahead forecast error variance of the ℓth element of $\boldsymbol{x}_t$ accounted for by the innovation in the jth element of $\boldsymbol{x}_t$. Notice that, due to the non–diagonal form of $\boldsymbol{\Sigma}_u$, the elements of $GFEVD\left(x_{\ell t};u_{jt},n\right)$ need not sum to unity across j.

In the case of structurally identified shocks, we have

$$SGFEVD\left(x_{\ell t};u_{jt},n\right)=\frac{\sigma_{u,jj}^{-1}\sum_{h=0}^{n}\left(\boldsymbol{e}_{\ell}'\boldsymbol{B}_h\left(\boldsymbol{PF}_0\right)^{-1}\boldsymbol{\Sigma}_v\boldsymbol{e}_j\right)^2}{\sum_{h=0}^{n}\boldsymbol{e}_{\ell}'\boldsymbol{B}_h\left(\boldsymbol{PF}_0\right)^{-1}\boldsymbol{\Sigma}_u\left(\boldsymbol{PF}_0\right)^{-1\prime}\boldsymbol{B}_h'\boldsymbol{e}_{\ell}},\ n=0,1,2,...,\ \ell=1,...,m, \tag{9}$$

3 Appendix C: Central Forecasts and Probability Event Forecasts in the GVAR Model

We will now describe in detail the estimation and construction of the probability event forecasts where the underlying model is the following VAR(p) model in m global variables in $\boldsymbol{x}_t$:

$$\boldsymbol{x}_t=\sum_{j=1}^{p}\boldsymbol{G}_j\boldsymbol{x}_{t-j}+\boldsymbol{g}_0^*+\boldsymbol{g}_1t+\boldsymbol{\varepsilon}_t,\ t=1,2,...,T, \tag{10}$$

where $\boldsymbol{G}_j$, $j=1,\ldots,p$, are $m\times m$ matrices of coefficients, $\boldsymbol{g}_0^*$ and $\boldsymbol{g}_1$ are $m\times 1$ vectors of coefficients on deterministics and $\boldsymbol{\varepsilon}_t$ is assumed to be iid with zero means, a positive semi–definite covariance matrix, $\boldsymbol{\Sigma}_{\varepsilon}$, and serially uncorrelated.

3.1 Case 1: Absence of Parameter Uncertainty

Suppose that the consistent estimators of $\boldsymbol{G}_j$, $j=1,\ldots,p$, $\boldsymbol{g}_0^*$, $\boldsymbol{g}_1$ and $\boldsymbol{\Sigma}_{\varepsilon}$ are given by $\hat{\boldsymbol{G}}_j$, $j=1,\ldots,p$, $\hat{\boldsymbol{g}}_0^*$, $\hat{\boldsymbol{g}}_1$ and $\hat{\boldsymbol{\Sigma}}_{\varepsilon}$, respectively. Then, the estimates of the n–step ahead (central) forecasts of $\boldsymbol{x}_T$ conditional on the information set, Ω_T, are obtained recursively by

$$\hat{\boldsymbol{x}}_{t+h}=\sum_{j=1}^{p}\hat{\boldsymbol{G}}_j\boldsymbol{x}_{t+h-j}+\hat{\boldsymbol{g}}_0^*+\hat{\boldsymbol{g}}_1\left(t+h\right),\ n=1,2,\ldots, \tag{11}$$

where the initial values, $\boldsymbol{x}_t$, $\boldsymbol{x}_{t-1},\ldots,\boldsymbol{x}_{t-p+1}$, are given.

3.1.1 Analytic method

In this case we need to derive the n–step–ahead forecast innovations, denoted by $\boldsymbol{\zeta}_{T+n}$, and its covariance matrix. To this end it is more convenient to work with the following canonical representation of (10):

$$\boldsymbol{X}_t=\boldsymbol{G}\boldsymbol{X}_{t-1}+\boldsymbol{D}_t+\boldsymbol{E}_t,\ t=1,...,T,T+1,...,T+n, \tag{12}$$

where

$$\underset{mp\times 1}{\boldsymbol{X}_t}=\begin{bmatrix}\boldsymbol{x}_t\\ \boldsymbol{x}_{t-1}\\ \vdots\\ \boldsymbol{x}_{t-p+1}\end{bmatrix},\ \underset{mp\times 1}{\boldsymbol{D}_t}=\begin{bmatrix}\boldsymbol{g}_0+\boldsymbol{g}_1^*t\\ \boldsymbol{0}\\ \vdots\\ \boldsymbol{0}\end{bmatrix},\ \underset{mp\times 1}{\boldsymbol{E}_t}=\begin{bmatrix}\boldsymbol{\varepsilon}_t\\ \boldsymbol{0}\\ \vdots\\ \boldsymbol{0}\end{bmatrix},$$

$$\underset{mp\times mp}{\boldsymbol{G}}=\begin{bmatrix}\boldsymbol{G}_1 & \boldsymbol{G}_2 & \boldsymbol{G}_3 & \cdots & \boldsymbol{G}_{p-1} & \boldsymbol{G}_p\\ \boldsymbol{I}_m & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}_m & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{I}_m & \boldsymbol{0}\end{bmatrix}.$$

Then, it is straightforward to show that

$$\boldsymbol{x}_{T+n} = \boldsymbol{J}\boldsymbol{G}^n\boldsymbol{X}_T + \boldsymbol{D}^*_{T+n} + \boldsymbol{\zeta}_{T+n},\ n = 1, 2, \ldots, \tag{13}$$

where $\underset{m\times mp}{\boldsymbol{J}} = (\boldsymbol{I}_m, \boldsymbol{0}, ..., \boldsymbol{0})$,

$$\boldsymbol{D}^*_{T+n} = \sum_{j=0}^{n-1} \boldsymbol{B}_j\boldsymbol{D}_{T+n-j},\ \boldsymbol{\zeta}_{T+n} = \sum_{j=0}^{n-1} \boldsymbol{B}_j\boldsymbol{\varepsilon}_{T+n-j}, \tag{14}$$

$$\boldsymbol{B}_j = \boldsymbol{J}\boldsymbol{G}^j\boldsymbol{J}',\ j = 1, 2, \ldots,\ \boldsymbol{B}_0 = \boldsymbol{I}_m,\ \boldsymbol{B}_j = \boldsymbol{0},\ j < 0. \tag{15}$$

Then it is easily seen that

$$Var\left(\boldsymbol{\zeta}_{T+n}\right) = \sum_{j=0}^{n-1} \boldsymbol{B}_j\boldsymbol{\Sigma}_\varepsilon\boldsymbol{B}'_j. \tag{16}$$

Having obtained the estimates of $\boldsymbol{x}_{T+n}$ and the covariance matrix, $Var\left(\boldsymbol{\zeta}_{T+n}\right)$, the probability forecast for event(s) can be readily evaluated via the analytic method, assuming that $\boldsymbol{\varepsilon}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_\varepsilon)$. Notice that the single event of interest can be written as

$$x_{j,T+n} = \boldsymbol{\tau}'_j\boldsymbol{x}_{T+n} < a,$$

where $\boldsymbol{\tau}_j$ is an $m \times 1$ selection vector with unity on the j–th row and zeros elsewhere.[2] Then, the associated probability forecast (PF) is given by $\Pr\left(\boldsymbol{\tau}'_j\boldsymbol{x}_{T+n} < a|\Omega_T; M(\boldsymbol{\theta})\right)$,[3] where $\boldsymbol{\theta}$ simply denotes the set of all parameters in the model. This PF can be evaluated analytically by the standard normal cumulative distribution function,[4]

$$\Phi\left(\frac{a - \boldsymbol{\tau}'_j\boldsymbol{x}_{T+n}}{\sqrt{\boldsymbol{\tau}'_j\widehat{Var\left(\boldsymbol{\zeta}_{T+n}\right)}\boldsymbol{\tau}_j}}\right), \tag{17}$$

where $\widehat{Var\left(\boldsymbol{\zeta}_{T+n}\right)}$ is the consistent estimator of the covariance matrix of the forecast innovations, $\boldsymbol{\zeta}_{T+n}$, given by (16).

Now consider the case where the event of interest is that the n–quarter–ahead forecast of the quarterly change in the jth variable in $\boldsymbol{x}_T$ conditional on Ω_T is less than a. In this case, the associated PF is given by $\Pr\left(\boldsymbol{\tau}'_j\Delta\boldsymbol{x}_{T+n} < a\right)$, which can be evaluated analytically by

$$\Phi\left(\frac{a - \boldsymbol{\tau}'_j\Delta\boldsymbol{x}_{T+n}}{\sqrt{\boldsymbol{\tau}'_j\widehat{Var\left(\Delta\boldsymbol{\zeta}_{T+n}\right)}\boldsymbol{\tau}_j}}\right), \tag{18}$$

where $\widehat{Var\left(\Delta\boldsymbol{\zeta}_{T+n}\right)}$ is the consistent estimator of

$$Var\left(\Delta\boldsymbol{\zeta}_{T+n}\right) = \sum_{j=0}^{n-1} \boldsymbol{B}^*_j\boldsymbol{\Sigma}_\varepsilon\boldsymbol{B}^{*'}_j,\ \boldsymbol{B}^*_j = \boldsymbol{B}_j - \boldsymbol{B}_{j-1},\ j = 0, 1, 2, \ldots, \tag{19}$$

In the case where we evaluate the probability forecast of the event that the n–quarter–ahead forecast of the 4 quarter moving average growth rate of the jth variable in $\boldsymbol{x}_T$ conditional on Ω_T is less than a, this can be expressed as

$$\frac{1}{4}\sum_{i=n-3}^{n} \Delta x_{j,T+i} = \frac{1}{4}\boldsymbol{\tau}'_j\left(\boldsymbol{x}_{T+n} - \boldsymbol{x}_{T+n-4}\right) < a,$$

[2]For simplicity, but without loss of generality, we have chosen the simple selection vector, $\boldsymbol{\tau}_j$. In principle, the analysis is valid for any other selection vector (see the case of current account forecasts for an example).

[3]In what follows, we use the simplifying notation $\Pr\left(\boldsymbol{\tau}'_j\boldsymbol{x}_{T+h} < a\right)$.

[4]In fact, the probability of both single and joint events can be computed using the standard normal *cdf*. However, the evaluation of the multiple integrals in the case of joint events is analytically complex where the number of events exceeds about four.

the associated PF, $\Pr\left(\frac{1}{4}\boldsymbol{\tau}_j'\left(\boldsymbol{x}_{T+n}-\boldsymbol{x}_{T+n-4}\right)<a\right)$, is evaluated analytically by

$$\Phi\left(\frac{a-\frac{1}{4}\boldsymbol{\tau}_j'\left(\boldsymbol{x}_{T+n}-\boldsymbol{x}_{T+n-4}\right)}{\frac{1}{4}\sqrt{\boldsymbol{\tau}_j' Var\left(\frac{1}{4}\widehat{\sum_{i=n-3}^{n}\Delta\boldsymbol{\zeta}_{T+i}}\right)\boldsymbol{\tau}_j}}\right), \tag{20}$$

where

$$Var\left(\frac{1}{4}\sum_{i=n-3}^{n}\Delta\boldsymbol{\zeta}_{T+i}\right)=\frac{1}{16}\sum_{j=0}^{n-1}\boldsymbol{B}_j^{**}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{B}_j^{**\prime},\ \boldsymbol{B}_j^{**}=\boldsymbol{B}_j-\boldsymbol{B}_{j-4},\ j=0,1,2,\ldots,$$

Finally, consider the event that the n–quarter–ahead forecast of the n quarter average growth rate of the jth variable in $\boldsymbol{x}_T$ conditional on Ω_T is less than a. Now the associated PF is given by $\Pr\left(\frac{1}{n}\boldsymbol{\tau}_j'\left(\boldsymbol{x}_{T+n}-\boldsymbol{x}_T\right)<a\right)$ and evaluated analytically by

$$\Phi\left(\frac{a-\frac{1}{n}\boldsymbol{\tau}_j'\left(\boldsymbol{x}_{T+n}-\boldsymbol{x}_T\right)}{\frac{1}{n}\sqrt{\boldsymbol{\tau}_j' Var\left(\frac{1}{n}\widehat{\sum_{i=1}^{n}\Delta\boldsymbol{\zeta}_{T+i}}\right)\boldsymbol{\tau}_j}}\right), \tag{21}$$

where

$$Var\left(\frac{1}{n}\sum_{i=1}^{n}\Delta\boldsymbol{\zeta}_{T+i}\right)=\frac{1}{n^2}\sum_{j=0}^{n-1}\boldsymbol{B}_j^{**}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{B}_j^{**\prime},\ \boldsymbol{B}_j^{**}=\boldsymbol{B}_j-\boldsymbol{B}_{j-n},\ j=0,1,2,\ldots,$$

3.1.2 Stochastic Simulations

The probability of the single or joint events can be carried out using the standard normal *cdf*s as described above. However, the evaluation of the multiple integrals for the case of the joint events is, in general, analytically complex. Therefore, towards this end, we need to simulate the values of $\boldsymbol{x}_{T+n}$ by

$$\boldsymbol{x}_{T+n}^{(s)}=\sum_{j=1}^{p}\hat{\boldsymbol{G}}_j\boldsymbol{x}_{T+n-j}^{(s)}+\hat{\boldsymbol{g}}_0^{*}+\hat{\boldsymbol{g}}_1(t+n)+\boldsymbol{\varepsilon}_{T+n}^{(s)},\ n=1,2,\ldots,\ s=1,...,S, \tag{22}$$

where superscript '(s)' refers to the sth replication, and $\boldsymbol{x}_T^{(s)}=\boldsymbol{x}_T$, $\boldsymbol{x}_{T-1}^{(s)}=\boldsymbol{x}_{T-1},\ldots,$ $\boldsymbol{x}_{T-p+1}^{(s)}=\boldsymbol{x}_{T-p+1}$ for all s. The $\boldsymbol{\varepsilon}_{T+n}^{(s)}$'s can be drawn by either the parametric or nonparametric methods; that is, parametrically from a Normal distribution, $N(0,\hat{\boldsymbol{\Sigma}}_\varepsilon)$, or nonparametrically from the historic residuals, $\{\hat{\boldsymbol{\varepsilon}}_1,\ldots,\hat{\boldsymbol{\varepsilon}}_T\}$.

• *Probability Forecasts of Single Events:* The probability forecast of the single event, $\boldsymbol{\tau}_j'\boldsymbol{x}_{T+n}<a$, may be evaluated via stochastic simulations by

$$\frac{1}{S}\sum_{s=1}^{S}I\left(a-\boldsymbol{\tau}_j'\boldsymbol{x}_{T+n}^{(s)}\right), \tag{23}$$

where $I(w)$ is an indicator function which takes the value of unity if $w>0$ and zero otherwise. Similarly, the probability forecasts of other single events, $\boldsymbol{\tau}_j'\Delta\boldsymbol{x}_{T+n}<a$, $\frac{1}{4}\boldsymbol{\tau}_j'\left(\boldsymbol{x}_{T+n}-\boldsymbol{x}_{T+n-4}\right)<a$ or $\frac{1}{n}\boldsymbol{\tau}_j'\left(\boldsymbol{x}_{T+n}-\boldsymbol{x}_T\right)<a$ can be measured via stochastic simulations respectively by

$$\frac{1}{S}\sum_{s=1}^{S}I\left(a-\boldsymbol{\tau}_j'\Delta\boldsymbol{x}_{T+n}^{(s)}\right), \tag{24}$$

$$\frac{1}{S}\sum_{s=1}^{S}I\left(a-\frac{1}{4}\boldsymbol{\tau}_j'\left(\boldsymbol{x}_{T+n}^{(s)}-\boldsymbol{x}_{T+n-4}^{(s)}\right)\right), \tag{25}$$

$$\frac{1}{S}\sum_{s=1}^{S}I\left(a-\frac{1}{n}\boldsymbol{\tau}_j'\left(\boldsymbol{x}_{T+n}^{(s)}-\boldsymbol{x}_T^{(s)}\right)\right). \tag{26}$$

• *Probability Forecasts of Joint Events:* Suppose the joint event of interest, which consists of L single events, can be defined by

$$\boldsymbol{\varphi}\left(\boldsymbol{x}_{T+1,h}\right) \ < \ \boldsymbol{a}, \tag{27}$$

where $\boldsymbol{\varphi}\left(.\right)$ is the $L \times 1$ vectors $\boldsymbol{\varphi}\left(.\right) = \left(\varphi_1\left(.\right), \varphi_2\left(.\right), ..., \varphi_L\left(.\right)\right)'$, $\varphi_j\left(\boldsymbol{x}_{T+1,h}\right)$ is a scalar function of the variables over the forecasting horizon $T+1$, ..., $T+h$, $\boldsymbol{a} = \left(a_1, a_2, ..., a_L\right)'$ with a_j being the "threshold" value associated with $\varphi_j\left(.\right)$, and $\varphi_j\left(\boldsymbol{x}_{T+1,h}\right) < a_j$ can be defined by any single event considered above. Thus, the probability forecast of the joint event (27), $\Pr\left(\boldsymbol{\varphi}\left(\boldsymbol{x}_{T+1,h}\right) \ < \ \boldsymbol{a}\right)$, can be computed as

$$\frac{1}{S}\sum_{s=1}^{S} I\left(\boldsymbol{a} - \boldsymbol{\varphi}\left(\boldsymbol{x}_{T+1}^{(s)}, ..., \boldsymbol{x}_{T+h}^{(s)}\right)\right), \tag{28}$$

where $I\left(\boldsymbol{w}\right)$ is an indicator function with $\boldsymbol{w} = \left(w_1, w_2, ..., w_L\right)'$. $I\left(\boldsymbol{w}\right)$ takes the value of unity if $w_j > 0$ $\forall j$ and zero otherwise.

3.2 Case 2: Presence of Parameter Uncertainty

In the presence of parameter uncertainty, it is advisable to focus on bootstrap techniques, since the analytic methods are generally complicated even in the case of simple events. First, to allow for parameter uncertainty, we use the bootstrapping technique to obtain R simulated *within sample* values of $\boldsymbol{x}_t$, $t = 1, ..., T$, denoted by $\boldsymbol{x}_t^{(r)}$,

$$\boldsymbol{x}_t^{(r)} = \sum_{j=1}^{p} \hat{\boldsymbol{G}}_i \boldsymbol{x}_{t-j}^{(r)} + \hat{\boldsymbol{g}}_0^{*} + \hat{\boldsymbol{g}}_1 t + \boldsymbol{\varepsilon}_t^{(r)}, \ t = 1, 2, ..., T, \ r = 1, ..., R;$$

where the actual observations of initial values, $\boldsymbol{x}_{-1}, \ldots, \boldsymbol{x}_{-r}$ are used. The $\boldsymbol{\varepsilon}_t^{(r)}$'s can also be drawn by either the parametric or nonparametric method.

Having obtained the set of R simulated samples, $\left\{\boldsymbol{x}_1^{(r)}, \ldots, \boldsymbol{x}_T^{(r)}\right\}$, the $VAR(p)$ model (10) is re-estimated R times to obtain new estimates, $\hat{\boldsymbol{G}}_j^{(r)}$, $j = 1, \ldots, s$, $\hat{\boldsymbol{g}}_0^{*(r)}$, $\hat{\boldsymbol{g}}_1^{(r)}$ and $\hat{\boldsymbol{\Sigma}}_{\varepsilon}^{(r)}$, $r = 1, .., R$. On each occasion, we undertake similar exercises to those described above to obtain measures of the probability forecast, denoted simply by $\pi^{(r)}$, $r = 1, ..., R$.[5] The empirical mean of the probability forecast is obtained by

$$\bar{\pi} = \frac{1}{R}\sum_{r=1}^{R} \pi^{(r)},$$

and the associated $(100 - \alpha)\%$ lower and upper confidence bands computed as the $R\alpha$th smallest and largest values of $\pi^{(r)}$, $r = 1, .., R$, respectively.[6]

3.3 Generating Simulated Errors

Either of two methods may be used in the simulation of the in–sample and future errors so that the contemporaneous correlations that exist across the errors in the different equations of the GVAR model are taken into account. The first is the *parametric* method where the errors are drawn from an assumed probability distribution function. Alternatively, one could employ a *non–parametric* procedure based on re–sampling techniques.

[5] In the case of stochastic simulation in conjunction with the bootstrapping technique, there is another possibility: nonparametric draws from the pooled set of residuals, $\left\{\boldsymbol{\varepsilon}_1^{(r)}, \ldots, \boldsymbol{\varepsilon}_T^{(r)}\right\}_{r=1}^{R}$.

[6] We can also obtain the empirical mean of the central forecast(s) by

$$\bar{\boldsymbol{x}}_{T+n} = \frac{1}{R}\sum_{r=1}^{R}\frac{1}{S}\sum_{i=1}^{S} \boldsymbol{x}_{T+n}^{(i,r)}, \ n = 1, 2, ...,$$

and construct the associated 100α % lower and upper confidence bands as the $RS\alpha$–th smallest and largest values of $\boldsymbol{x}_{T+n}^{(i,r)}$, $i = 1, ..., S$, $r = 1, .., R$, respectively.

3.3.1 Parametric Approach

Under this approach, we assume that the errors are drawn from a multivariate distribution with zero means and the covariance matrix, $\hat{\boldsymbol{\Sigma}}_{\varepsilon}\left(=T^{-1}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\right)$. To obtain the simulated errors for m global variables over H periods we first generate mH draws from an assumed iid distribution which we denote $\boldsymbol{v}_{T+h}^{(s)}$, $h=1,...,H$. These are then used to obtain $\left\{\boldsymbol{\varepsilon}_{T+n}^{(s)}, n=1,...,H\right\}$ computed as $\boldsymbol{\varepsilon}_{T+n}^{(s)}=\hat{\boldsymbol{K}}\boldsymbol{v}_{T+n}^{(s)}$ for $s=1,...,S$, where $\hat{\boldsymbol{K}}=\hat{\boldsymbol{P}}\hat{\boldsymbol{\Lambda}}^{1/2}$, $\hat{\boldsymbol{\Lambda}}$ is a diagonal matrix with eigenvalues of $\hat{\boldsymbol{\Sigma}}_{\varepsilon}$ and $\hat{\boldsymbol{P}}$ is an orthogonal matrix consisting of the corresponding eigenvectors. These matrices are obtained using the singular value decomposition of $\hat{\boldsymbol{\Sigma}}_{\varepsilon}=\hat{\boldsymbol{P}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{P}}'$. Notice that the Choleski decomposition of $\hat{\boldsymbol{\Sigma}}_{\varepsilon}$ is not applicable here since $\hat{\boldsymbol{\Sigma}}_{\varepsilon}$ is positive semi–definite due to the underlying common factor structure of the GVAR model (see DHPS). In our applications, we generate, $\boldsymbol{v}_{T+n}^{(s)}$, $n=1,...,H$, as $IIN(0,\boldsymbol{I}_m)$, although other parametric distributions such as multivariate Student t can also be used.

3.3.2 Non–Parametric Approaches

The most obvious non–parametric approach to generating the simulated errors, $\boldsymbol{\varepsilon}_{T+n}^{(r)}$, is simply to take n random draws with replacement from the in–sample residual vectors $\{\hat{\boldsymbol{\varepsilon}}_1,\ldots,\hat{\boldsymbol{\varepsilon}}_T\}$. We refer to this as 'Method 1'. The simulated errors thus obtained clearly have the same distribution and covariance structure as that observed in the original sample. However, this procedure is subject to the criticism that it could introduce serial dependence at longer forecast horizons since the pseudo–random draws are made from the same relatively small T vector of residuals.

An alternative non–parametric method for generating simulated errors ('Method 2'), makes use of the singular value decomposition of the estimated covariance employed in the parametric approach. For a given choice of $\hat{\boldsymbol{K}}=\hat{\boldsymbol{P}}\hat{\boldsymbol{\Lambda}}^{1/2}$, a set of mT transformed error terms $\{\hat{\boldsymbol{v}}_1,\ldots,\hat{\boldsymbol{v}}_T\}$ are computed such that $\hat{\boldsymbol{v}}_t=\hat{\boldsymbol{K}}^{-}\hat{\boldsymbol{\varepsilon}}_t$, $t=1,...,T$, where $\hat{\boldsymbol{K}}^{-}$ is the generalised inverse of $\hat{\boldsymbol{K}}$. The mT individual error terms are mutually uncorrelated, but retain the distributional information (relating to extreme values and so on) contained in the original observed errors. A set of mh simulated errors are then obtained by drawing with replacement from these transformed residuals, denoted by $\left\{\boldsymbol{v}_{T+1}^{(s)},\ldots,\boldsymbol{v}_{T+H}^{(s)}\right\}$. These, in turn, are used to obtain $\left\{\boldsymbol{\varepsilon}_{T+1}^{(s)},\ldots,\boldsymbol{\varepsilon}_{T+H}^{(s)}\right\}$, recalling that $\boldsymbol{\varepsilon}_{T+h}^{(r)}=\hat{\boldsymbol{K}}\boldsymbol{v}_{T+h}^{(r)}$ for $s=1,...,S$. Given that the re–sampling occurs from the mT transformed error terms, Method 2 also has the advantage over Method 1 that the serial dependence introduced through sampling with replacement is likely to be less problematic.

3.3.3 Choice of Approach

The non–parametric approaches described above have the advantages that they make no distributional assumptions on the error terms and are better able to capture the uncertainties arising from (possibly rare) extreme observations. However, they suffer from the fact that they require random sampling *with replacement*. Replacement is essential as otherwise the draws at longer forecast horizons are effectively truncated and unrepresentative. On the other hand, for a given sample size, it is clear that re–sampling from the observed errors with replacement inevitably introduces serial dependence in the simulated forecast errors at longer horizons as the same observed errors are drawn repeatedly. When generating simulated errors over a forecast horizon, therefore, this provides an argument for the use of non–parametric methods over shorter forecast horizons, but suggests that a greater reliance might be placed on the parametric approach for the generation of probability forecasts at longer time horizons.

4 Appendix D: Structural Break Test, Co-Breaking Test and Weak Exogeneity Test in the GVAR Model

4.1 Test for Structural Breaks

We follow the recursive-residuals-based approach proposed by Brown, Durbin and Evans (1975) to test for unknown structural breaks in the countries/regions of interest. Our test is based on the following ARDL(1,1) specification as indicated in Section 3.2 (equation 3.12)

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + \boldsymbol{\theta}_i \boldsymbol{e}_{i,t-1} + \boldsymbol{\phi}_i \boldsymbol{x}^*_{i,t-1} + \sum_{j=1}^{p-1} \boldsymbol{\varrho}_{ij} \Delta \boldsymbol{y}_{i,t-j} + \sum_{j=0}^{p} \boldsymbol{\vartheta}_{ij} \Delta \boldsymbol{e}_{i,t-j} + \sum_{j=0}^{p} \boldsymbol{\varphi}_{ij} \Delta \boldsymbol{x}^*_{i,t-j} + \epsilon_{it} \quad (29)$$

Equation (29) can be re-written as

$$\Delta y_{it} = \boldsymbol{\beta}' \boldsymbol{w}_t + \epsilon_{it}, \quad (30)$$

where $\boldsymbol{w}_t$ and $\boldsymbol{\beta}$ are the $k \times 1$ vectors including all the regressors and all the coefficients of equation (29), respectively.

Equation (30) is estimated recursively from $k+1$ to T to obtain the recursive residuals computed as

$$r_t = \frac{\left(\Delta y_{it} - \boldsymbol{\beta}'_{t-1} \boldsymbol{w}_t\right)}{\left(\sqrt{1 + \boldsymbol{w}'_t \left(\boldsymbol{W}'_{t-1} \boldsymbol{W}_{t-1}\right)^{-1}} \boldsymbol{w}_t\right)}, \; t = k+1, ..., T, \quad (31)$$

where $\boldsymbol{W}'_{t-1} = [\boldsymbol{w}_1, \boldsymbol{w}_2, ..., \boldsymbol{w}_t]$, $\Delta \boldsymbol{Y}'_{t-1} = [\Delta y_1, \Delta y_2, ..., \Delta y_t]$, and $\boldsymbol{\beta}_t = \left(\boldsymbol{W}'_t \boldsymbol{W}_t\right)^{-1} \boldsymbol{W}'_t \boldsymbol{Y}_t$. The CUSUM test statistic is computed as

$$R_t = \sum_{t=k+1}^{T} r_t / s, \quad (32)$$

where s is the standard error of regression fitted to all T sample points, calculated as

$$s = \sqrt{\frac{1}{T-k} \sum_{t=1}^{T} \hat{\epsilon}_t^2} \; , \; \hat{\epsilon}_t = \Delta y_{it} \; - \; \hat{\boldsymbol{\beta}}' \boldsymbol{w}_t, \; t = 1, ..., T, \quad (33)$$

The (100 - α)% significance lines, upper line (UL) and lower line (LL), are found by connecting the following points:

$$\left[k \; , \; \pm a \left(T-k\right)^{1/2}\right] \quad and \quad \left[k \; , \; \pm 3 \times a \left(T-k\right)^{1/2}\right] \quad (34)$$

where

$$\alpha = 0.01, \; a = 1.143,$$
$$\alpha = 0.05, \; a = 0.948,$$
$$\alpha = 0.10, \; a = 0.850,$$

The Null hypothesis of no structural break is defined as:

$$\boldsymbol{H}_o : LL_t \; < \; R_t \; < \; UL_t,$$

4.2 Test for Co-Breaking

The co-breaking test is based on country-specific equation (2.2) in the paper:

$$\begin{aligned} \Delta \boldsymbol{x}_{it} \; &= \; \boldsymbol{c}_{i0} + \boldsymbol{c}^*_{i0} \Delta d_{it} + \boldsymbol{c}^*_{i1} \Delta d_{i,t-1} + \boldsymbol{\Lambda}_i \Delta \boldsymbol{x}^*_{it} + \boldsymbol{\Gamma}_i \Delta \boldsymbol{z}_{i,t-1} \\ &+ \; \boldsymbol{\alpha}_i \boldsymbol{\beta}'_i \left(\boldsymbol{z}_{i,t-1} - \boldsymbol{\mu}_i d_{i,t-1} - \boldsymbol{\gamma}_i \left(t-1\right)\right) + \boldsymbol{u}_{it}. \end{aligned} \quad (35)$$

Under the Null hypothesis of no co-breaking, r over-identified zero restrictions are imposed on the coefficients of the break dummies in $\boldsymbol{\beta}_i$ in addition to the r^2 exactly-identified restrictions (r denotes

the number of cointegrating vectors). Thus, the co-breaking test is simply the test for the validity of r over-identified zero restrictions imposed on $\boldsymbol{\beta}_i$. The log-likelihood ratio (LR) test statistic is computed as

$$LR = -2\left[\ell\left(\hat{\boldsymbol{\varphi}}_0\right) - \ell\left(\hat{\boldsymbol{\varphi}}_1\right)\right] \ \sim \ \chi^2_r, \tag{36}$$

where $\boldsymbol{\varphi}$ denotes the vector of all the parameters in (35), $\hat{\boldsymbol{\varphi}}_0$ and $\hat{\boldsymbol{\varphi}}_1$ are the ML estimators of $\boldsymbol{\varphi}$ obtained under the null and alternative (respectively), and $\ell\left(\hat{\boldsymbol{\varphi}}_0\right)$ and $\ell\left(\hat{\boldsymbol{\varphi}}_1\right)$ are the corresponding maximised values of the log-likelihood function.

4.3 Test for the Weak Exogeneity of Foreign Variables

In order to test the weak exogeneity of the country-specific foreign variables, we have recourse to the distinction between the marginal VAR and conditional VEC models employed by GLPS. It follows that we may write the VECMX* (p, p) generalisation of (35) for the ith country as follows:

$$\begin{aligned}\begin{pmatrix} \boldsymbol{A}_{xx,i} & \boldsymbol{A}_{xx^*,i} \\ \mathbf{0} & \boldsymbol{A}_{x^*x^*,i} \end{pmatrix} \begin{pmatrix} \Delta \boldsymbol{x}_{it} \\ \Delta \boldsymbol{x}^*_{it} \end{pmatrix} \ &= \ \boldsymbol{c}_{i0} + \boldsymbol{c}^*_{i0}\Delta d_{it} + \boldsymbol{c}^*_{i1}\Delta d_{i,t-1} + \sum_{j=1}^{p-1} \boldsymbol{\Gamma}_{ij} \begin{pmatrix} \Delta \boldsymbol{x}_{i,t-j} \\ \Delta \boldsymbol{x}^*_{i,t-j} \end{pmatrix} \\ &+ \ \boldsymbol{\Pi}_i \begin{pmatrix} \boldsymbol{x}_{i,t-1} \\ \boldsymbol{x}^*_{i,t-1} \\ \boldsymbol{\mu}_i d_{i,t-1} \\ \boldsymbol{\gamma}_i\left(t-1\right) \end{pmatrix} + \begin{pmatrix} \boldsymbol{\epsilon}_{x,it} \\ \boldsymbol{\epsilon}_{x^*,it} \end{pmatrix}\end{aligned} \tag{37}$$

where:

$$\boldsymbol{\Pi}_i = \begin{pmatrix} \boldsymbol{\Pi}_{x,i} \\ \mathbf{0} \\ \boldsymbol{\Pi}_{\mu,i} \\ \boldsymbol{\Pi}_{\gamma,i} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_{x,i} \\ \mathbf{0} \\ \boldsymbol{\alpha}_{\mu,i} \\ \boldsymbol{\alpha}_{\gamma,i} \end{pmatrix} \boldsymbol{\beta}'_i$$

The zero restrictions in the matrices $\boldsymbol{A}_i$ and $\boldsymbol{\Pi}_i$ jointly ensure that the foreign variables are $I(1)$ forcing for the system (*c.f.* Granger and Lin, 1995). Hence, the weak exogeneity of the variables in $\boldsymbol{x}^*_{it}$ may be evaluated simply by testing the zero restrictions relating to the foreign variables in the adjustment matrix, $\boldsymbol{\alpha}_{x^*,i} = \mathbf{0}$. In particular, $\boldsymbol{\alpha}_{x^*,i}$ is the $m^* \times r$ adjustment matrix of the marginal VAR model for the m^* foreign variables of the ith country. The test for the weak exogeneity of the jth foreign variable in the ith country, $x^*_{j,it}$, is simply the test for the validity of the following joint restriction

$$\alpha^{(j,1)}_{x^*,i} = \alpha^{(j,2)}_{x^*,i} = ... = \alpha^{(j,r)}_{x^*,i} = 0, \tag{38}$$

where the superscript (j, s), $(s = 1, 2, ..., r)$, $(j = 1, 2, ..., m^*)$ refers to the position of the coefficient $\alpha^{(j,s)}$ within matrix $\boldsymbol{\alpha}_{x^*,i}$. The F-test statistic is computed as

$$\boldsymbol{F} = \frac{\left(\sum_1^T \tilde{\epsilon}^j_{x^*,it}\tilde{\epsilon}^j_{x^*,it} - \sum_1^T \hat{\epsilon}^j_{x^*,it}\hat{\epsilon}^j_{x^*,it}\right)\backslash r}{\sum_1^T \hat{\epsilon}^j_{x^*,it}\hat{\epsilon}^j_{x^*,it}\backslash(T-k)} \ \sim \ F\left(r \ , \ T-k\right), \tag{39}$$

where $\tilde{\epsilon}^j_{x^*,it}$ and $\hat{\epsilon}^j_{x^*,it}$ are the residuals from the restricted and unrestricted models (respectively), r is the number of restrictions and k the total number of regressors.