Appendix for

"Uncertainty Network Risk and Currency Returns"

Abstract

This appendix presents supplementary details not included in the main body of the paper.

Contents

A. Estimation of the time-varying parameter VAR model **49**

B. Asset Pricing Tests **51**

C. Transaction Costs **53**

A Estimation of the time-varying parameter VAR model

Let $\mathbf{CIV}_t$ be an $N \times 1$ vector generated by a stable time-varying parameter (TVP) heteroskedastic VAR model with p lags:

$$\mathbf{CIV}_{t,T} = \mathbf{\Phi}_1(t/T)\mathbf{CIV}_{t-1,T} + \ldots + \mathbf{\Phi}_p(t/T)\mathbf{CIV}_{t-p,T} + \boldsymbol{\epsilon}_{t,T}, \tag{A.1}$$

where $\boldsymbol{\epsilon}_{t,T} = \mathbf{\Sigma}^{-1/2}(t/T)\boldsymbol{\eta}_{t,T}, \boldsymbol{\eta}_{t,T} \sim NID(0, \boldsymbol{I}_M)$ and $\mathbf{\Phi}(t/T) = (\mathbf{\Phi}_1(t/T), \ldots, \mathbf{\Phi}_p(t/T))^\top$ are the time varying autoregressive coefficients. Note that all roots of the polynomial $\chi(z) = \det\left(\mathbf{I}_N - \sum_{p=1}^{L} z^p \mathbf{B}_{p,t}\right)$ lie outside the unit circle, and $\mathbf{\Sigma}_t^{-1}$ is a positive definite time-varying covariance matrix. Stacking the time-varying intercepts and autoregressive matrices in the vector $\phi_{t,T}$ with $\overline{\mathbf{CIV}}_t^\top = (\mathbf{I}_N \otimes x_t)$, $x_t = \left(1, x_{t-1}^\top, \ldots, x_{t-p}^\top\right)$ and denoting the Kronecker product by $\otimes$, the model can be written as:

$$\mathbf{CIV}_{t,T} = \overline{\mathbf{CIV}}_{t,T}^\top \phi_{t,T} + \mathbf{\Sigma}_{t/T}^{-\frac{1}{2}} \boldsymbol{\eta}_{t,T} \tag{A.2}$$

We obtain the time-varying parameters of the model by employing the Quasi-Bayesian Local-Likelihood (QBLL) approach of Petrova (2019). The estimation of Equation (A.1) requires re-weighting the likelihood function. The weighting function gives higher proportions to observations surrounding the time period whose parameter values are of interest. The local likelihood function at time period k is given by:

$$\begin{aligned} &\mathrm{L}_k\left(\mathbf{CIV}|\theta_k, \mathbf{\Sigma}_k, \overline{\mathbf{CIV}}\right) \propto \\ &|\mathbf{\Sigma}_k|^{\mathrm{trace}(\mathbf{D}_k)/2} \exp\left\{-\frac{1}{2}(\mathbf{CIV} - \overline{\mathbf{CIV}}^\top \phi_k)^\top (\mathbf{\Sigma}_k \otimes \mathbf{D}_k)(\mathbf{CIV} - \overline{\mathbf{CIV}}^\top \phi_k)\right\} \end{aligned} \tag{A.3}$$

The $\mathbf{D}_k$ is a diagonal matrix whose elements hold the weights:

$$\mathbf{D}_k = \mathrm{diag}(\varrho_{k1}, \ldots, \varrho_{kT}) \tag{A.4}$$

$$\varrho_{kt} = \phi_{T,k} w_{kt} / \sum_{t=1}^{T} w_{kt} \tag{A.5}$$

$$w_{kt} = (1/\sqrt{2\pi}) \exp((-1/2)((k-t)/H)^2), \quad \text{for } k, t \in \{1, \ldots, T\} \tag{A.6}$$

$$\zeta_{Tk} = \left(\left(\sum_{t=1}^{T} w_{kt}\right)^2\right)^{-1} \tag{A.7}$$

where ϱ_{kt} is a normalised kernel function. w_{kt} uses a Normal kernel weighting function.

ζ_{Tk} gives the rate of convergence and behaves like the bandwidth parameter H in (A.6). The kernel function puts a greater weight on the observations surrounding the parameter estimates at time k relative to more distant observations.

We use a Normal-Wishart prior distribution for $\phi_k|\,\boldsymbol{\Sigma}_k$ for $k \in \{1,\ldots,T\}$:

$$\phi_k|\boldsymbol{\Sigma}_k \backsim \mathcal{N}\left(\phi_{0k}, (\boldsymbol{\Sigma}_k \otimes \Xi_{0k})^{-1}\right) \tag{A.8}$$

$$\boldsymbol{\Sigma}_k \backsim \mathcal{W}\left(\alpha_{0k}, \boldsymbol{\Gamma}_{0k}\right) \tag{A.9}$$

where ϕ_{0k} is a vector of prior means, Ξ_{0k} is a positive definite matrix, α_{0k} is a scale parameter of the Wishart distribution ($\mathcal{W}$), and $\boldsymbol{\Gamma}_{0k}$ is a positive definite matrix.

The prior and weighted likelihood function implies a Normal-Wishart quasi posterior distribution for $\phi_k|\,\boldsymbol{\Sigma}_k$ for $k = \{1,\ldots,T\}$. Formally, let $\mathbf{A} = (\bar{x}_1^\top,\ldots,\bar{x}_T^\top)^\top$ and $\mathbf{Y} = (x_1,\ldots,x_T)^\top$, then:

$$\phi_k|\boldsymbol{\Sigma}_k, \mathbf{A}, \mathbf{Y} \;\backsim\; \mathcal{N}\left(\widetilde{\theta}_k, \left(\boldsymbol{\Sigma}_k \otimes \widetilde{\Xi}_k\right)^{-1}\right) \tag{A.10}$$

$$\boldsymbol{\Sigma}_k \;\backsim\; \mathcal{W}\left(\widetilde{\alpha}_k, \widetilde{\boldsymbol{\Gamma}}_k^{-1}\right) \tag{A.11}$$

with quasi posterior parameters

$$\widetilde{\phi}_k = \left(\mathbf{I}_N \otimes \widetilde{\Xi}_k^{-1}\right)\left[\left(\mathbf{I}_N \otimes \mathbf{A}^\top \mathbf{D}_k \mathbf{A}\right)\hat{\phi}_k + \left(\mathbf{I}_N \otimes \Xi_{0k}\right)\phi_{0k}\right] \tag{A.12}$$

$$\widetilde{\Xi}_k = \widetilde{\Xi}_{0k} + \mathbf{A}^\top \mathbf{D}_k \mathbf{A} \tag{A.13}$$

$$\widetilde{\alpha}_k = \alpha_{0k} + \sum_{t=1}^{T} \varrho_{kt} \tag{A.14}$$

$$\widetilde{\boldsymbol{\Gamma}}_k = \boldsymbol{\Gamma}_{0k} + \mathbf{Y}'\mathbf{D}_k\mathbf{Y} + \boldsymbol{\Phi}_{0k}\boldsymbol{\Gamma}_{0k}\boldsymbol{\Phi}_{0k}^\top - \widetilde{\boldsymbol{\Phi}}_k\widetilde{\boldsymbol{\Gamma}}_k\widetilde{\boldsymbol{\Phi}}_k^\top \tag{A.15}$$

where $\widehat{\phi}_k = \left(\mathbf{I}_N \otimes \mathbf{A}^\top \mathbf{D}_k \mathbf{A}\right)^{-1}\left(\mathbf{I}_N \otimes \mathbf{A}^\top \mathbf{D}_k\right) y$ is the local likelihood estimator for ϕ_k. The matrices $\boldsymbol{\Phi}_{0k}$, $\widetilde{\boldsymbol{\Phi}}_k$ are conformable matrices from the vector of prior means, ϕ_{0k}, and a draw from the quasi posterior distribution, $\widetilde{\phi}_k$, respectively.

The motivation for employing these methods are threefold. First, we are able to estimate large systems that conventional Bayesian estimation methods do not permit. This is typically because the state-space representation of an N-dimensional TVP VAR (p) requires an additional $N(3/2 + N(p + 1/2))$ state equations for every additional variable. Conventional Markov Chain Monte Carlo (MCMC) methods fail to estimate larger models, which in general confine one to (usually) fewer than 6 variables in the system. Second, the standard approach is fully parametric and requires a law of motion. This can distort inference if the true law of motion is misspecified. Third, the methods used here permit direct estimation of the VAR's time-varying covariance matrix, which has an inverse-Wishart density and is symmetric positive definite at every point in time.

In estimating the model, we use p=2 and a Minnesota Normal-Wishart prior with a shrinkage value $\varphi = 0.05$ and centre the coefficient on the first lag of each variable to 0.1 in each respective equation. The prior for the Wishart parameters are set following Kadiyala and Karlsson (1997). For each point in time, we run 500 simulations of the model to generate the (quasi) posterior distribution of parameter estimates. Note we experiment with various lag lengths, $p = \{2, 3, 4, 5\}$; shrinkage values, $\varphi = \{0.01, 0.25, 0.5\}$; and values to centre the coefficient on the first lag of each variable, $\{0, 0.05, 0.2, 0.5\}$. Network measures from these experiments are qualitatively similar. Notably, adding lags to the VAR and increasing the persistence in the prior value of the first lagged dependent variable in each equation increases computation time.

B Asset Pricing Tests

The standard Euler equation implies that the excess returns rx^j_{t+1} of a portfolio j satisfy the equation:

$$\mathbb{E}_t \left(M_{t+1} rx^j_{t+1} \right) = 0, \tag{B.16}$$

in which M_{t+1} is the stochastic discount factor (SDF). We assume that the SDF is a linear function of a set of risk factors f_{t+1} and is defined as follows:

$$M_{t+1} = 1 - b'(f_{t+1} - \mu_f). \tag{B.17}$$

Notice that we employ a de-meaned version of the SDF to avoid the issue related to an affine transformation of the factors (Kan and Robotti, 2008).

We are interested in testing the perfomance of the linear pricing models defined by Equations (B.16)-(B.17). To do so, we estimate factor loadings using the generalized method of moments (GMM) (Hansen, 1982). Substituting (B.17) into (B.16), we obtain the following N moment conditions $\mathbb{E}_t \left([1 - b'(f_{t+1} - \mu_f)] rx_{t+1} \right) = 0_N$, where rx_{t+1} is the N-

dimensional vector of test asset excess returns. We simultaneously estimate the unknown vector of factor means μ_f. Thus, GMM moment conditions also include the set of k restrictions $\mathbb{E}_t\left(f_{t+1}-\mu_f\right)=0_k$, where k denotes the number of factors in the SDF specification. Therefore, we have the following population moment conditions:

$$\mathbb{E}_t\left[g_{t+1}(\theta)\right]=\mathbb{E}_t\begin{bmatrix}[1-b'(f_{t+1}-\mu_f)]rx_{t+1}\\ f_{t+1}-\mu_f\end{bmatrix}=0_{N+k},$$

where $\theta=(b',\mu')'$ is the vector of parameters. The sample moment conditions are then defined as:

$$\bar{g}_T(\theta)=\begin{bmatrix}\bar{g}_T^1(\theta)\\ \bar{g}_T^2(\theta)\end{bmatrix}=\begin{bmatrix}\frac{1}{T}\sum_{t=1}^{T}\left[1-b'(f_{t+1}-\mu_f)\right]rx_{t+1}\\ \frac{1}{T}\sum_{t=1}^{T}\left[f_{t+1}-\mu_f\right]\end{bmatrix}.$$

We implement a one-stage GMM estimation with the prespecified weighting matrix consisting of the identity matrix I_N for the first moment conditions and a large weight assigned to the remaining restrictions. Standard errors are computed based on a heteroscedasticity and autocorrelation consistent (HAC) estimate of the long-run covariance matrix $S=\sum_{j=-\infty}^{\infty}\mathbb{E}[g(\theta)g(\theta)']$ by the Newey and West (1987) procedure with Andrews (1991) optimal lag selection.

We now evaluate the performance of linear pricing models in explaining the cross-section of network portfolios. We construct the cross-sectional R^2, root mean squared pricing error (RMSE), and the Hansen and Jagannathan (1997) distance (HJ_{dist}). Hansen and Jagannathan (1997) provide two nice illustrations of HJ_{dist}. First, it is the maximum pricing error of a portfolio with a unit second moment. Second, it measures the minimum distance between the proposed SDF and the set of admissible SDFs. Thus, tests of the linear SDFs defined by Equation (B.17) boil down to testing the null hypothesis that the pricing errors equal zero, i.e. HJ_{dist} equals zero. Formally, the Hansen and Jagannathan (1997) distance is defined as:

$$\text{HJ}_{\text{dist}}=\sqrt{\min_{\theta}\bar{g}_T(\theta)'G_T^{-1}\bar{g}_T(\theta)},\tag{B.18}$$

in which G_T is the sample second moment matrix of the test excess returns, that is, $G_T=\frac{1}{T}\sum_{t=1}^{T}rx_{t+1}rx_{t+1}'$. One can obtain HJ_{dist} by applying the one-stage GMM estimation with the weighting matrix equal to G_T^{-1}. The advantage of this definition is that G_T^{-1} is independent of the optimal parameters and hence this allows the comparison between different SDF specifications (Hansen and Jagannathan, 1997). The disadvantage of this approach is that G_T^{-1} is not optimal in the sense of Hansen (1982) and hence HJ_{dist} is not asymptotically a random variable of $\chi^2(N-k)$ distribution. Instead, the sample HJ_{dist} follows a weighted sum of $\chi^2(1)$ random variables (see Jagannathan and Wang (1996) and Kan and Robotti (2008) for specification tests using gross and excess returns, respectively). Therefore, we calculate the simulated p-values for HJ_{dist} based on this statistic.

C Transaction Costs

We use time-varying quoted bid-ask spreads to compute the currency excess returns adjusted for transaction costs. Following Menkhoff, Sarno, Schmeling, and Schrimpf (2012b), we take into account the whole cycle of each currency in the short or long positions from $t-1$ to $t+1$. When the investor buys the currency at time t and sells at time $t+1$, he pays the corresponding bid-ask costs each period. In our notations, the excess returns of long (l) and short (s) positions are respectively $rx_{t+1}^l = f_t^b - s_{t+1}^a$ and $rx_{t+1}^s = -f_t^a + s_{t+1}^b$. If the investor buys the currency at time t but decides to keep it in the portfolio at time $t+1$, then the net excess returns are computed as $rx_{t+1}^l = f_t^b - s_{t+1}$ and $rx_{t+1}^s = -f_t^a + s_{t+1}$. If the currency, which belongs to the portfolio at time t and is sold at time $t+1$, was already in the current portfolio at time $t-1$, then the excess returns $rx_{t+1}^l = f_t^b - s_{t+1}^a$ and $rx_{t+1}^s = -f_t^a + s_{t+1}^b$, that is, the investor must still initiate a position in the one-month forward contract. At the start (January 1996) and at the end (December 2013) of the sample, the investor is assumed to start and close positions in all foreign currencies.