Time-varying Parameter Vector Autoregressions II

Mike Ellington

Haindorf Seminar 2020; Hejnice, Czech Republic

January 22, 2020

Roadmap

Pitfalls of Conventional TVP VARs

Primiceri (2005), Cogley and Sargent (2005) and all other papers using these methods are subject to a number of pitfalls.

Can you tell me any?
You can respond in either manner:

1. Go to PollEv.com/mte00 and respond via your laptop
2. Text MTE00 to +44(0)7480781235 to join the session, then send a follow up text with your response

Pitfalls of Conventional TVP VARs

Primiceri (2005), Cogley and Sargent (2005) and all other papers using these methods are subject to a number of pitfalls.

Can you tell me any?
You can respond in either manner:

1. Go to PollEv.com/mte00 and respond via your laptop
2. Text MTE00 to +44(0)7480781235 to join the session, then send a follow up text with your response

Pitfalls of Conventional TVP VARs

- How do we know model parameters follow random walks?
- **Incorrectly specifying parameter process as RW means Kalman filter is not optimal which means inference may be invalid.**
- How long do these models take to compute?
- **Ellington et al. (2017), Ellington (2018a), and Ellington (2018b) take around 2-3 days to estimate reduced-form parameters for 1 model; structural analysis can take weeks!**
- **How many variables can we use in these models?**
- **The curse of dimensionality!**

Pitfalls of Conventional TVP VARs

- How do we know model parameters follow random walks?
- **Incorrectly specifying parameter process as RW means Kalman filter is not optimal which means inference may be invalid.**
- How long do these models take to compute?
- **Ellington et al. (2017), Ellington (2018a), and Ellington (2018b) take around 2-3 days to estimate reduced-form parameters for 1 model; structural analysis can take weeks!**
- **How many variables can we use in these models?**
- **The curse of dimensionality!**

Pitfalls of Conventional TVP VARs

- How do we know model parameters follow random walks?
- **Incorrectly specifying parameter process as RW means Kalman filter is not optimal which means inference may be invalid.**
- How long do these models take to compute?
- **Ellington et al. (2017), Ellington (2018a), and Ellington (2018b) take around 2-3 days to estimate reduced-form parameters for 1 model; structural analysis can take weeks!**
- **How many variables can we use in these models?**
- **The curse of dimensionality!**

Pitfalls of Conventional TVP VARs

- How do we know model parameters follow random walks?
- **Incorrectly specifying parameter process as RW means Kalman filter is not optimal which means inference may be invalid.**
- How long do these models take to compute?
- **Ellington et al. (2017), Ellington (2018a), and Ellington (2018b) take around 2-3 days to estimate reduced-form parameters for 1 model; structural analysis can take weeks!**
- How many variables can we use in these models?
- The curse of dimensionality!

Pitfalls of Conventional TVP VARs

- How do we know model parameters follow random walks?
- **Incorrectly specifying parameter process as RW means Kalman filter is not optimal which means inference may be invalid.**
- How long do these models take to compute?
- **Ellington et al. (2017), Ellington (2018a), and Ellington (2018b) take around 2-3 days to estimate reduced-form parameters for 1 model; structural analysis can take weeks!**
- **How many variables can we use in these models?**
- **The curse of dimensionality!**

Pitfalls of Conventional TVP VARs

- How do we know model parameters follow random walks?
- **Incorrectly specifying parameter process as RW means Kalman filter is not optimal which means inference may be invalid.**
- How long do these models take to compute?
- **Ellington et al. (2017), Ellington (2018a), and Ellington (2018b) take around 2-3 days to estimate reduced-form parameters for 1 model; structural analysis can take weeks!**
- **How many variables can we use in these models?**
- **The curse of dimensionality!**

The Curse of Dimensionality I

The magnitude of time-variation in these models is dictated by the matrix $\mathbf{Q}$.

In Ellington (2019) we have N=3 variables and p=2 lags. The vector $\mathbf{B}_t$ is of length $N(1+Np)$=21, which means $\mathbf{Q}$ is a 21×21 matrix.

Ellington (2018b) uses an N=4 p=2 TVP VAR which means $\mathbf{B}_t$ is of length $N(1+Np)$=36 and $\mathbf{Q}$ is a 36×36 matrix.

The typical algorithm requires inverting $\mathbf{Q}$; as the dimension of $\mathbf{Q}$ rises the less likely is that a draw of $\mathbf{Q}$ will be positive definite. This is heavily influenced by the prior specification for $\mathbf{Q}$.

The Curse of Dimensionality II

It is very difficult to estimate a conventional TVP VAR model with $N > 4$ variables; even if you have large T!

The Curse of Dimensionality III

In fact, an N dimensional TVP VAR with p lags requires:

- $N(3/2 + N(p + 1/2))$ state equations
- $N(N + Np)$ state variables
- $\mathbf{B}_t$ to be of length $N(1 + Np)$

The Curse of Dimensionality IV

N	p	# state variables	length state vector	# state equations
1	2	3	3	4
2	2	12	10	13
3	2	27	21	27
4	2	48	36	46
10	2	300	210	265
25	2	1875	1275	1600
50	2	7500	5050	6325
100	2	30000	20100	25150

Current Directions the Literature is taking us I

1. Joshua Chan develops hybrid TVP VARs by developing a method where the algorithm automatically decides whether VAR coefficients and error variance are time-varying or constant equation by equation. Large defined as N=20.
2. Chan, Eisenstat, and Strachan develop a new approach that exploits the near singularity of estimated coavriance matrices in large systems combined with parameter expansions. Large defined as N=15.
3. Koop & Korobilis use forgetting factors and rely on Kalman filter estimation to track TVP. Large defined as N=25. Korobilis & Yilmaz use this method to span network connections of N=35 stock return volatilities.
4. Kalli & Griffin develop Bayesian nonparametric VAR models that allow for parameter change, breaks in error variances, and non-Gaussian error terms. We currently have some work that uses N=9 variables (using MCMC).

Current Directions the Literature is taking us II

5. Amisano, Giannone, & Lenza use VAR Kronecker structure to reduce number of free parameters in state equation(s). Large defined as N=20.
6. Kapetanios, Marcellino, & Venditti propose a non-parametric frequentist approach that uses kernel regressions to impose time-variation and shrinkage. Large defined as N=78.
7. Eisenstat, Hou, & Koop develop composite likelihood methods for large VAR models with stochastic volatility. These methods involve estimating a large number of parsimonious models and taking a weighted average across models. In a high-dimensional setting, these methods are computationally feasible and produce better forecasts than standard VARs. Large is defined as N=196.
8. **Petrova proposes Quasi Bayesian Local Likelihood Methods to estimate TVP VARs. Empirical application defines large as N=87.**

Quasi Bayesian Local Likelihood Methods (QBLL) I

Petrova (2019) establishes Quasi Bayesian Local Likelihood (QBLL) methods for general multivariate models with TVP.

Advantages

- Time-variation enters in a non-parametric manner $\rightarrow$ no need to specify law of motion!
- Closed form solutions for quasi posterior $\rightarrow$ permits prior shrinkage and use for truly large models!
- Direct estimation of time-varying covariance matrices $\rightarrow$ no need for restrictions as in state space setting!
- Computationally efficient $\rightarrow$ time is treated independently which allows us to use parallel computing!

Quasi Bayesian Local Likelihood Methods (QBLL) II

Disadvantage

- Gaussian error variances $\rightarrow$ Potential to extend the special case in Petrova (2019) to allow for non-Gaussian error distributions (e.g. $\backsim t$ errors).
- Conjugate prior and posterior distributions.

Quasi Bayesian Local Likelihood Methods (QBLL) III

The method essentially re-weights the likelihood function at each time period j using a kernel weighting scheme to induce time-variation.

The weighting function provides higher proportions to observations surrounding the time period whose parameter values are of interest!

Predictive Regressions using QBLL I

We now consider a linear regression with TVP that has a Normal-Gamma quasi-posterior distribution. Let

$$
\begin{aligned}
y_t &= \beta_{0,t} + x_{1,t}\beta_{1,t} + \cdots + x_{k,t}\beta_{k,t} + \epsilon_t, \quad \epsilon \backsim \text{NID}\left(0, \sigma_t^2\right) \\
y_t &= x_t\beta_t + \epsilon_t
\end{aligned}
$$

where $x_t \equiv (1, x_{1,t}, \ldots, x_{k,t})$ and $\beta_t \equiv (\beta_{0,t}, \beta_{1,t}, \ldots, \beta_{k,t})^\top$.

Predictive Regressions using QBLL II

Now let $\lambda_t \equiv \sigma_t^{-2}$. The weighted likelihood function of the sample $Y \equiv (y_1, \ldots, y_T)$, using $\mathbf{X} \equiv \left(x_1^\top, \ldots, x_T^\top\right)^\top$ as a $T \times k$ matrix, at each time point j is given by

$$
\begin{aligned}
L_j\left(Y|\beta_j, \lambda_j, \mathbf{X}\right) &= (2\pi)^{-\mathrm{tr}(\mathbf{D}_j)/2} \lambda_j^{\mathrm{tr}(\mathbf{D}_j)/2} \exp\left\{-\frac{\lambda_j}{2}(Y - \mathbf{X}\beta_j)^\top \mathbf{D}_j (Y - \mathbf{X}\beta_j)\right\} \\
\mathbf{D}_j &= \mathrm{diag}\left(\vartheta_{j,1}, \ldots, \vartheta_{j,T}\right) \\
\vartheta_{j,t} &= \phi_{T,j} w_{j,t} / \sum_{t=1}^{T} w_{j,t} \\
w_{j,t} &= \left(1/\sqrt{2\pi}\right) \exp\left((-1/2)((j-t)/H)^2\right), \ \forall j, t \in \{1, \ldots, T\} \\
\phi_{T,j} &= \left(\sum_{t=1}^{T} w_{j,t}^2\right)^{-1}
\end{aligned}
$$

Predictive Regressions using QBLL III

Now assuming β_j, λ_j have a Normal-Gamma prior distribution $\forall j \in \{1, \ldots, T\}$

$$\begin{aligned} \beta_j|\lambda_j &\backsim \mathrm{N}\left(\beta_{0,j}, \left(\lambda_j \kappa_{0,j}\right)^{-1}\right) \\ \lambda_j &\backsim \mathrm{Ga}\left(\alpha_{0,j}, \gamma_{0,j}\right) \end{aligned}$$

Predictive Regressions using QBLL IV

We can combine L_j with the above priors such that β_j, λ_j have Normal-Gamma quasi-posterior distribution $\forall j \in \{1, \dots, T\}$

$$\begin{aligned} \beta_j|\lambda_j &\backsim \mathrm{N}\left(\bar{\beta}_j, (\lambda_j \bar{\kappa}_j)^{-1}\right) \\ \lambda_j &\backsim \mathrm{Ga}\left(\bar{\alpha}_j, \bar{\gamma}_j\right) \end{aligned}$$

with (quasi) posterior parameters:

$$\begin{aligned} \bar{\beta}_j &= \bar{\kappa}_j^{-1}\left(\mathbf{X}^\top \mathbf{D}_j \mathbf{X} \widehat{\beta}_j + \kappa_{0,j}\beta_{0,j}\right), \ \widehat{\beta}_j = \left(\mathbf{X}^\top \mathbf{D}_j \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{D}_j y \\ \bar{\kappa}_j &= \kappa_{0,j} + \mathbf{X}^\top \mathbf{D}_j \mathbf{X} \\ \bar{\alpha}_j &= \alpha_{0,j} + \sum_{t=1}^{T} \vartheta_j \\ \bar{\gamma}_j &= \gamma_{0,j} + \frac{1}{2}\left(Y^\top \mathbf{D}_j Y - \bar{\beta}_j^\top \bar{\kappa}_j \bar{\beta}_j + \beta_{0,j}^\top \kappa_{0,j} \beta_{0,j}\right)^3 \end{aligned}$$

Predictive Regressions using QBLL V

The Algorithm

1. Initialise $\beta_{0,j},\ \kappa_{0,j}, \alpha_{0,j}, \gamma_{0,j}$ and compute kernel weights. Then repeat steps 2-3 $i = 1, 2, \ldots, K$ times.
2. For every $j \in \{1, 2, \ldots, T\}$, draw $\lambda_j^k | \mathbf{X}, y, \beta_j^{k-1}$ from $\backsim \text{Ga}\left(\bar{\alpha}_j, \bar{\gamma}_j\right)$.
3. For every $j \in \{1, 2, \ldots, T\}$, draw $\beta_j^k | y, \mathbf{X}, \lambda_j^k$ from $\backsim \text{N}\left(\bar{\beta}_j, (\lambda_j \bar{\kappa}_j)^{-1}\right)$.

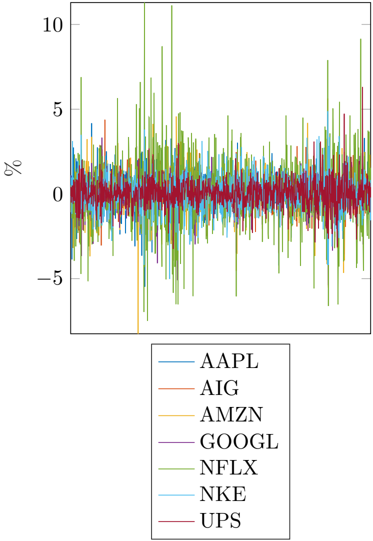
Time-varying Betas I

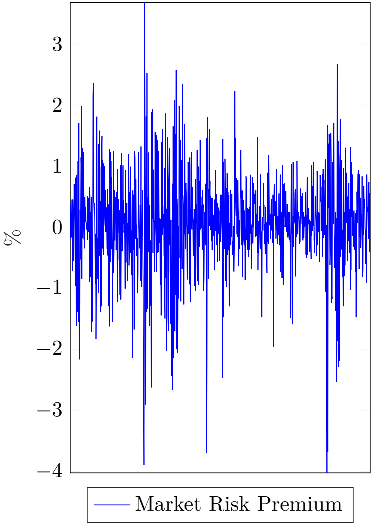
Suppose we wish to examine how a bundle of excess stock returns respond to changes in the market risk premium over time. We can use our predictive regression using QBLL to obtain time-varying betas, and indeed time-varying volatility.

$$r_{i,t} = \beta_{0,i,t} + \beta_{M,i,t} r_{M,t} + \epsilon_{i,t}, \;\; \epsilon_{i,t} \backsim \text{NID}\left(0, \sigma^2_{i,t}\right)$$

We use excess return data on 7 stocks: AAPL, AIG, AMZN, GOOGL, NFLX, NKE, UPS from 27/08/2014-31/08/2018, and the market risk premium from Ken French's Data Library.

Time-varying Betas II





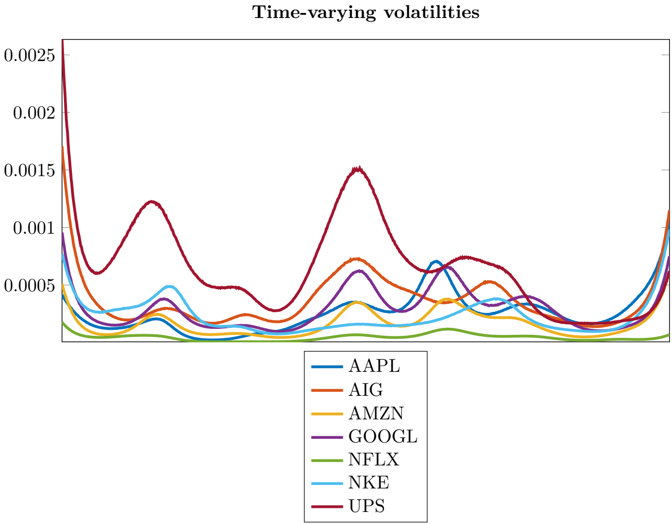
Time-varying Betas III

We run EXAMPLE4.m **which follows the below procedure:**

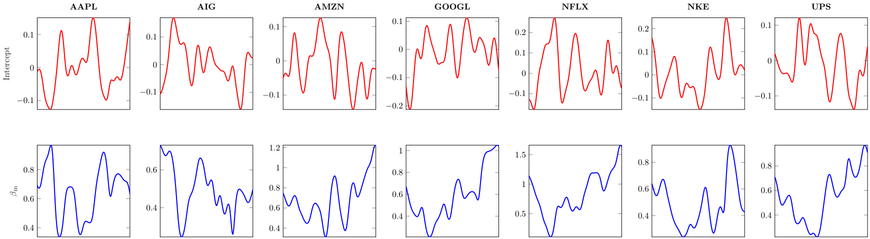
1. Initialise $\beta_{0,j} = \widehat{\beta}_{OLS}$, $\kappa_{0,j} = \widehat{\sigma}^2_{i,OLS}\left(\mathbf{X}^\top \mathbf{X}\right)^{-1}$, $\alpha_{0,j} = N + 2$, $\gamma_{0,j} = \widehat{\sigma}^{-2}_{i,OLS}$ and compute kernel weights. Then repeat steps 2-3 $i = 1, 2, \ldots, K = 500$ times.
2. For every $j \in \{1, 2, \ldots, T = 1000\}$, draw $\lambda_j^k | \mathbf{X}, y, \beta_j^{k-1}$ from $\backsim \mathrm{Ga}\left(\bar{\alpha}_j, \bar{\gamma}_j\right)$.
3. For every $j \in \{1, 2, \ldots, T = 1000\}$, draw $\beta_j^k | y, \mathbf{X}, \lambda_j^k$ from $\backsim \mathrm{N}\left(\bar{\beta}_j, \left(\lambda_j \bar{\kappa}_j\right)^{-1}\right)$.

Note: Requires parallel computing toolbox to run! If you do not have this on your laptop, change “parfor” loop on line 40 in QBLL_univariate.m function to “for” loop!

Time-varying Betas IV



Time-varying Betas V



Time-varying Betas VI

Table: OLS Estimates 27/08/2014–31/08/2018

	AAPL	AIG	AMZN	GOOGL	NFLX	NKE	UPS
$\widehat{\beta}_0^{OLS}$	-0.0022	0.0035	-0.0091	-0.0018	0.0330	0.0103	-0.0017
$\widehat{\beta}_M^{OLS}$	0.6300	0.4864	0.6563	0.5386	0.7300	0.4804	0.4931

TVP VARs using QBLL I

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

$$
\begin{aligned}
y_t &= \mathbf{B}_{0,t} + \sum_{p=1}^{P} \mathbf{B}_{p,t} y_{t-p} + \varepsilon_t, \;\; \varepsilon_t = \mathbf{\Xi}_t^{-\frac{1}{2}} \kappa_t, \;\; \kappa_t \backsim \mathrm{NID}(0, \mathbf{I}_N) \\
y_t &= \mathbf{X}_t' \theta_t + \mathbf{\Xi}_t^{-\frac{1}{2}} \kappa_t
\end{aligned}
$$

where $\mathbf{B}_{0,t}, \mathbf{B}_{p,t}$ contain the time-varying intercepts and autoregressive matrices, respectively. Note that all roots of the polynomial, $\psi(z) = \det\left(\mathbf{I}_N - \sum_{p=1}^{L} z^p \mathbf{B}_{p,t}\right)$, lie outside the unit circle, and $\mathbf{\Xi}_t^{-1}$ is a positive definite time-varying covariance matrix. θ_t stacks time varying coefficient matrices and intercepts, and $\mathbf{X}_t' = (\mathbf{I}_N \otimes x_t)\,,\; x_t = \left(1, y_{t-1}', \ldots, y_{t-p}'\right)$ and $\otimes$ denotes the Kronecker product.

TVP VARs using QBLL II

The local likelihood function at time period j is given by:

$$\begin{aligned}
\mathsf{L}_j(y|\theta_j, \mathbf{\Xi}_j, \mathbf{X}) &\propto |\mathbf{\Xi}_j|^{\text{trace}(\mathbf{D}_j)/2} \exp\{-\frac{1}{2}(y - \mathbf{X}'\theta_j)' (\mathbf{\Xi}_j \otimes \mathbf{D}_j)(y - \mathbf{X}'\theta_j)\} \\
\mathbf{D}_j &= \text{diag}(\vartheta_{j1}, \ldots, \vartheta_{jT}) \\
\vartheta_{jt} &= \phi_{T,j} w_{jt} / \sum_{t=1}^{T} w_{jt} \\
w_{jt} &= (1/\sqrt{2\pi}) \exp((-1/2)((j-t)/H)^2), \text{ for } j, t \in \{1, \ldots, T\} \\
\phi_{Tj} &= \left(\sum_{t=1}^{T} w_{jt}^2 \right)^{-1}
\end{aligned}$$

where ϑ_{jt} is a normalised kernel function. w_{jt} uses a Normal kernel weighting function. ϕ_{Tj} gives the rate of convergence and behaves like the bandwidth parameter H, and it is the kernel function that provides greater weight to observations surrounding the parameter estimates at time j relative to more distant observations.

TVP VARs using QBLL III

Using a Normal-Wishart prior distribution for $\theta_j | \mathbf{\Xi}_j$ for $j \in \{1, \ldots, T\}$:

$$\begin{aligned} \theta_j | \mathbf{\Xi}_j &\backsim \mathrm{N}\left(\theta_{0j}, (\mathbf{\Xi}_j \otimes \mathbf{\Omega}_{0j})^{-1}\right) \\ \mathbf{\Xi}_j &\backsim \mathrm{W}\left(\alpha_{0j}, \mathbf{\Gamma}_{0j}\right) \end{aligned}$$

where θ_{0j} is a vector of prior means, $\mathbf{\Omega}_{0j}$ is a positive definite matrix, α_{0j} is a scale parameter of the Wishart distribution (W), and $\mathbf{\Gamma}_{0j}$ is a positive definite matrix.

TVP VARs using QBLL IV

The prior and weighted likelihood function implies a Normal-Wishart quasi posterior distribution. Let $\tilde{\mathbf{X}} = (x_1', \ldots, x_T')'$ and $\tilde{\mathbf{Y}} = (y1, \ldots, y_T)'$ then:

$$\begin{aligned} \theta_j | \mathbf{\Xi}_j, \tilde{\mathbf{X}}, \tilde{\mathbf{Y}} &\backsim \mathsf{N}\left(\tilde{\theta}_j, \left(\mathbf{\Xi}_j \otimes \tilde{\mathbf{\Omega}}_j\right)^{-1}\right) \\ \mathbf{\Xi}_j &\backsim \mathsf{W}\left(\tilde{\alpha}_j, \tilde{\mathbf{\Gamma}}_j^{-1}\right) \end{aligned}$$

TVP VARs using QBLL V

The quasi posterior parameters are:

$$
\begin{aligned}
\tilde{\theta}_j &= \left(\mathbf{I}_N \otimes \tilde{\mathbf{\Omega}}_j^{-1}\right)\left[\left(\mathbf{I}_N \otimes \tilde{\mathbf{X}}'\mathbf{D}_j\tilde{\mathbf{X}}\right)\hat{\theta}_j + \left(\mathbf{I}_N \otimes \mathbf{\Omega}_{0j}\right)\theta_{0j}\right] \\
\tilde{\mathbf{\Omega}}_j &= \tilde{\mathbf{\Omega}}_{0j} + \tilde{\mathbf{X}}'\mathbf{D}_j\tilde{\mathbf{X}} \\
\tilde{\alpha}_j &= \alpha_{0j} + \sum_{t=1}^{T}\vartheta_{jt} \\
\tilde{\mathbf{\Gamma}}_j &= \mathbf{\Gamma}_{0j} + \tilde{\mathbf{Y}}'\mathbf{D}_j\tilde{\mathbf{Y}} + \mathbf{\Theta}_{0j}\mathbf{\Gamma}_{0j}\mathbf{\Theta}_{0j}' - \tilde{\mathbf{\Theta}}_j\tilde{\mathbf{\Gamma}}_j\tilde{\mathbf{\Theta}}_j'
\end{aligned}
$$

where $\hat{\theta}_j = \left(\mathbf{I}_N \otimes \tilde{\mathbf{X}}'\mathbf{D}_j\tilde{\mathbf{X}}\right)^{-1}\left(\mathbf{I}_N \otimes \tilde{\mathbf{X}}'\mathbf{D}_j\right)y$ is the local likelihood estimator for θ_j. The matrices $\mathbf{\Theta}_{0j}$, $\tilde{\mathbf{\Theta}}_j$ are conformable matrices from the vector of prior means, θ_{0j}, and a draw from the quasi posterior distribution, $\tilde{\theta}_j$, respectively.

TVP VARs using QBLL VI

The Algorithm

1. Initialise $\theta_{0,j},\ \mathbf{\Omega}_{0,j}, \alpha_{0,j}, \mathbf{\Gamma}_{0,j}$ and compute kernel weights. Then repeat steps 2-3 $i = 1, 2, \ldots, K$ times.
2. For every $j \in \{1, 2, \ldots, T\}$, draw $\mathbf{\Xi}_j^k | \tilde{\mathbf{X}}, \tilde{\mathbf{Y}}, \theta_j^{k-1}$ from $\backsim \mathrm{W}\left(\tilde{\alpha}_j, \tilde{\mathbf{\Gamma}}_j\right)$.
3. For every $j \in \{1, 2, \ldots, T\}$, draw $\theta_j^k | \tilde{\mathbf{Y}}, \tilde{\mathbf{X}}, \mathbf{\Xi}_j^k$ from $\backsim \mathrm{N}\left(\tilde{\theta}_j, \left(\mathbf{\Xi}_j \otimes \tilde{\mathbf{\Omega}}_j\right)^{-1}\right)$.

Network Connections using TVP VARs I

In order to track systemic and inter-sectoral spillovers, we decompose the H-step ahead forecast error variance of variable x_j that is due to shocks in $x_k, \forall j \neq k$ (Diebold and Yılmaz, 2014). Exploiting the results in Pesaran and Shin (1998), measures of connectedness work in a generalised framework.

Network Connections using TVP VARs II

In particular, variable k's contribution to j's H-step ahead generalised forecast error variance at time t, for $H = 1, \ldots, H$ is:

$$(\mathbf{\Lambda}_{H,t})_{j,k} = \frac{\xi_{kk,t}^{-1} \sum_{h=0}^{H} \left((\mathbf{\Psi}_{h,t} \mathbf{\Xi}_t)_{j,k} \right)^2}{\sum_{h=0}^{H} \left(\mathbf{\Psi}_{h,t} \mathbf{\Xi}_t \mathbf{\Psi}'_{h,t} \right)_{j,j}}$$

with $\mathbf{\Psi}_{h,t}$ is the time t $N \times N$ matrix of moving average coefficients at horizon h, $\xi_{kk,t}$ is the time t k^{th} diagonal element of $\mathbf{\Xi}_t$, $\xi_{kk,t} = (\mathbf{\Xi})_{kk,t}$. $(\mathbf{\Lambda}_{H,t})_{j,k}$ denotes the time t contribution of variable k to the variance of forecast error of variable j at horizon H.

Network Connections using TVP VARs III

Denote the normalised matrix as $\tilde{\boldsymbol{\Lambda}}_{H,t}$, the sum of all elements are equal to N by construction. Time t system-wide connectedness is:

$$\mathcal{C}_{H,t} = 100 \cdot \left(1 - \frac{\mathrm{Tr}\{\tilde{\boldsymbol{\Lambda}}_{H,t}\}}{\sum \tilde{\boldsymbol{\Lambda}}_{H,t}}\right)$$

System-wide connectedness is interpreted as the relative contribution of forecast error variances from other variables in the system. Therefore the higher the connectedness measure, the more central network is, and ultimately the riskier the system is.

Network Connections using TVP VARs IV

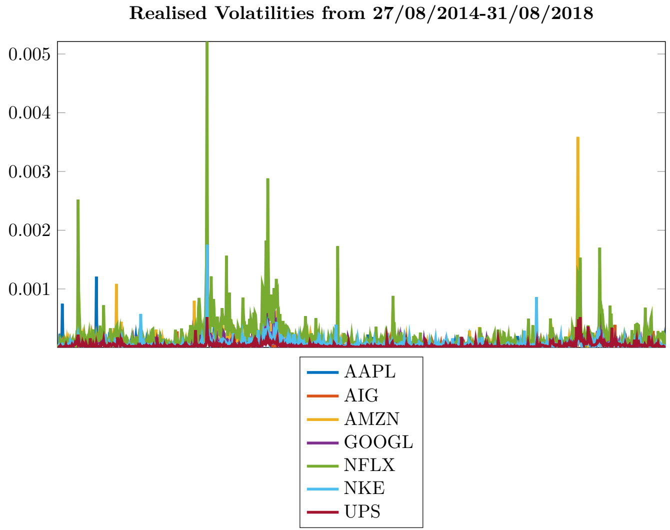
Note also that $\left(\tilde{\mathbf{\Lambda}}_{H,t}\right)_{j,k}$ provides measures of time t pairwise connectedness at horizon H.

$$
\begin{aligned}
\mathcal{C}_{\bullet\leftarrow k,t} &= 100\cdot\left(\frac{\sum_{k=1,k\neq j}^{N}\left(\tilde{\mathbf{\Lambda}}_{H,t}\right)_{k,j}}{\sum\tilde{\mathbf{\Lambda}}_{H,t}}\right)\\
\mathcal{C}_{k\leftarrow\bullet,t} &= 100\cdot\left(\frac{\sum_{k=1,k\neq j}^{N}\left(\tilde{\mathbf{\Lambda}}_{H,t}\right)_{j,k}}{\sum\tilde{\mathbf{\Lambda}}_{H,t}}\right)\\
\mathcal{C}_{k,t,H} &= \mathcal{C}_{\bullet\leftarrow k,t}-\mathcal{C}_{k\leftarrow\bullet,t}
\end{aligned}
$$

Empirical Application: TVP VAR QBLL I

Going back to our stocks: AAPL, AIG, AMZN, GOOGL, NFLX, NKE, UPS. We now wish to track connectedness between volatilities using our QBLL-TVP VAR.

Empirical Application: TVP VAR QBLL II



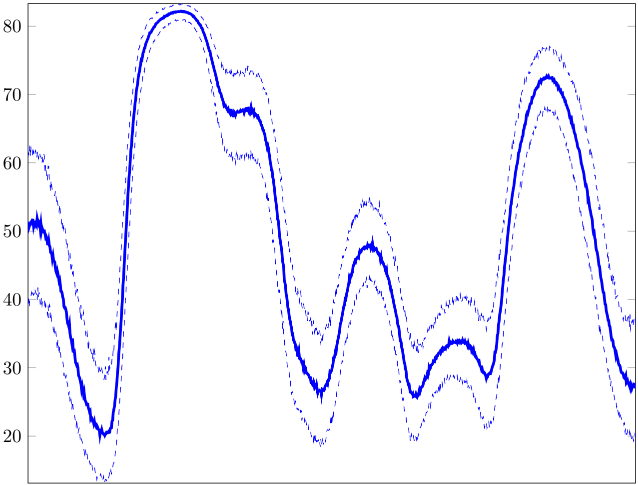
Empirical Application: Network Connections of Stock Return Volatilities I

We now run Estimate_QBLL_Connect.m **which implements the following:**

1. Initialise $\theta_{0,j}$, $\mathbf{\Omega}_{0,j}, \alpha_{0,j}, \mathbf{\Gamma}_{0,j}$ using a Minnesota Normal-Wishart prior and compute kernel weights. Then repeat steps 2-3 $i = 1, 2, \ldots, K$ times.
2. For every $j \in \{1, 2, \ldots, T\}$, draw $\mathbf{\Xi}_j^k | \tilde{\mathbf{X}}, \tilde{\mathbf{Y}}, \theta_j^{k-1}$ from $\backsim \mathrm{W}\left(\tilde{\alpha}_j, \tilde{\mathbf{\Gamma}}_j\right)$.
3. For every $j \in \{1, 2, \ldots, T\}$, draw $\theta_j^k | \tilde{\mathbf{Y}}, \tilde{\mathbf{X}}, \mathbf{\Xi}_j^k$ from $\backsim \mathrm{N}\left(\tilde{\theta}_j, \left(\mathbf{\Xi}_j \otimes \tilde{\mathbf{\Omega}}_j\right)^{-1}\right)$.

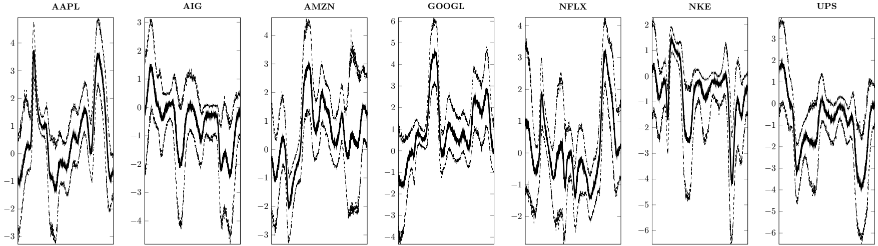
Empirical Application: Network Connections of Stock Return Volatilities II

Total Connectedness



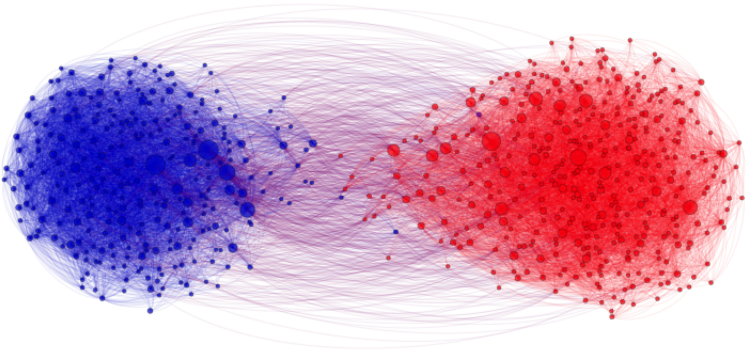
Empirical Application: Network Connections of Stock Return Volatilities III

Net-Directional Connectedness



Asset Pricing using Time-Frequency Dependent Network Centrality

Jozef Baruník & Michael Ellington



Network origins of factor structure

CONNECTEDNESS OF NETWORK

$\downarrow$

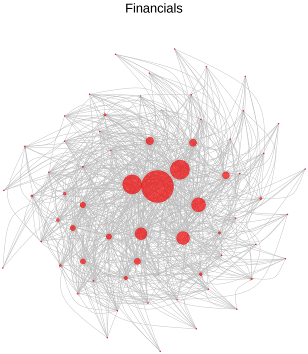
RISK

$\downarrow$

RISK PREMIUM

How to measure network risk?

Aggregate measures largely insufficient



Linkages may be horizon and time specific

Contribution I

- We provide a horizon specific time varying network origins of factor structure in a theoretical asset pricing model.
- **New measures for horizon specific time-varying network connectedness using all S&P500 constituents.**
- **Horizon specific time-varying factors are priced in the cross section of returns.**

Horizon specific time varying network connectedness I

Let's have locally stationary TVP-VAR of lag order p as

$$\mathbf{X}_{t,T} = \mathbf{\Phi}_1(t/T)\mathbf{X}_{t-1,T} + \ldots + \mathbf{\Phi}_p(t/T)\mathbf{X}_{t-p,T} + \epsilon_{t.T}, \quad (1)$$

In a neighborhood of a fixed time point $u_0 = t_0/T$ the process $\mathbf{X}_{t,T}$ can be approximated by a stationary process $\widetilde{\mathbf{X}}_t(u_0)$ as

$$\widetilde{\mathbf{X}}_t(u_0) = \mathbf{\Phi}_1(u_0)\widetilde{\mathbf{X}}_{t-1}(u_0) + \ldots + \mathbf{\Phi}_p(u_0)\widetilde{\mathbf{X}}_{t-p}(u_0) + \epsilon_t,$$

with $t \in \mathbb{Z}$ and under suitable regularity conditions $|\mathbf{X}_{t,T} - \widetilde{\mathbf{X}}_t(u_0)| = O_p\big(|t/T - u_0| + 1/T\big)$ which justifies the notation "locally stationary process."

Horizon specific time varying network connectedness II

Importantly, the process has time varying VMA(∞) representation

$$\mathbf{X}_{t,T} = \sum_{h=-\infty}^{\infty} \mathbf{\Psi}_{t,T}(h)\epsilon_{t-h}$$

where $\mathbf{\Psi}_{t,T}(h) \approx \mathbf{\Psi}(t/T, h)$ is a stochastic process satisfying $\sup_{\ell} ||\mathbf{\Psi}_t - \mathbf{\Psi}_{\ell}||^2 = O_p(h/t)$ for $1 \leq h \leq t$ as $t \to \infty$.

Horizon specific time varying network connectedness III

The time-varying spectral density is a key quantity for understanding frequency dynamics. Using the spectral representation for the local covariance that is associated with local spectral density,

$$\mathbb{E}\left[\widetilde{\mathbf{X}}_{t+h}(u)\widetilde{\mathbf{X}}_t^\top(u)\right] = \int_{-\pi}^{\pi} \boldsymbol{S}_{\mathbf{X}}(u,\omega)e^{i\omega h}d\omega$$

Suppose $\mathbf{X_{t,T}}$ is a weakly locally stationary process with

$$\sigma_{kk}^{-1}\sum_{h=0}^{\infty}\left|\left[\boldsymbol{\Psi}(u,h)\boldsymbol{\Sigma}(u)\right]_{j,k}\right| < +\infty, \forall j,k.$$

Horizon specific time varying network connectedness I

Proposition (FTVP Network Connectedness)

The ***time-frequency variance decompositions*** *of the jth variable at a rescaled time $u = t_0/T$ due to shocks in the kth variable on the frequency band $d = (a, b) : a, b \in (-\pi, \pi), a < b$ is defined as*

$$\left[\boldsymbol{\theta}(u,d)\right]_{j,k} = \frac{\sigma_{kk}^{-1} \int_a^b \left| \left[\boldsymbol{\Psi}(u) e^{-i\omega} \boldsymbol{\Sigma}(u) \right]_{j,k} \right|^2 d\omega}{\int_{-\pi}^{\pi} \left[\left\{ \boldsymbol{\Psi}(u) e^{-i\omega} \right\} \boldsymbol{\Sigma}(u) \left\{ \boldsymbol{\Psi}(u) e^{+i\omega} \right\}^{\top} \right]_{j,j} d\omega}$$

where $\boldsymbol{\Psi}(u)e^{-i\omega} = \sum_h e^{-i\omega h} \boldsymbol{\Psi}(u,h)$ is local impulse transfer function or frequency response function computed as the Fourier transform of the local impulse response $\Psi(u,h)$

Horizon specific time varying network connectedness II

Note that $\left[\boldsymbol{\theta}(u,d)\right]$ holds local horizon specific adjacency matrix and forms local weighted directed network at a given frequency

Horizon specific time varying network connectedness III

Remark

Denote by d_s an interval on the real line from the set of intervals D that form a partition of the interval $(-\pi, \pi)$, such that $\cap_{d_s \in D} d_s = \emptyset$, and $\cup_{d_s \in D} d_s = (-\pi, \pi)$. Due to the linearity of integral and the construction of d_s, we have

$$\Big[\boldsymbol{\theta}(u, \infty)\Big]_{j,k} = \sum_{d_s \in D} \Big[\boldsymbol{\theta}(u, d_s)\Big]_{j,k}.$$

Horizon specific time varying network connectedness IV

The row sum of time-frequency network connectedness do not necessarily sum to one, so we normalise elements in each row by the corresponding row sum

$$\left[\widetilde{\boldsymbol{\theta}}(u,d)\right]_{j,k} = \left[\boldsymbol{\theta}(u,d)\right]_{j,k} \Big/ \sum_{k=1}^{N} \left[\boldsymbol{\theta}(u,\infty)\right]_{j,k}$$

Horizon specific time varying network connectedness V

Local network connectedness at a given frequency band is the ratio of the off diagonal elements to the sum of the entire matrix

$$\mathcal{C}(u,d) = 100 \times \sum_{\substack{j,k=1 \\ j \neq k}}^{N} \left[\widetilde{\boldsymbol{\theta}}(u,d)\right]_{j,k} \Bigg/ \sum_{j,k=1}^{N} \left[\widetilde{\boldsymbol{\theta}}(u,\infty)\right]_{j,k}$$

Horizon specific time varying network connectedness VI

We define FROM connectedness and TO connectedness that signal whether an asset in the economy is a transmitter and receiver of shocks.

$$\mathcal{C}_{j\leftarrow\bullet}(u,d) = 100 \times \sum_{\substack{k=1\\k\neq j}}^{N} \left[\widetilde{\boldsymbol{\theta}}(u,d)\right]_{j,k} \Big/ \sum_{j,k=1}^{N} \left[\widetilde{\boldsymbol{\theta}}(u,\infty)\right]_{j,k}$$

$$\mathcal{C}_{j\rightarrow\bullet}(u,d) = 100 \times \sum_{\substack{k=1\\k\neq j}}^{N} \left[\widetilde{\boldsymbol{\theta}}(u,d)\right]_{k,j} \Big/ \sum_{k,j=1}^{N} \left[\widetilde{\boldsymbol{\theta}}(u,\infty)\right]_{k,j}$$

Horizon specific time varying network connectedness VII

Proposition (Reconstruction of the time-frequency network connectedness)

Denote by d_s an interval on the real line from the set of intervals D that form a partition of the interval $(-\pi, \pi)$, such that $\cap_{d_s \in D} d_s = \emptyset$, and $\cup_{d_s \in D} d_s = (-\pi, \pi)$. We then have that

$$\mathcal{C}(u, \infty) = \sum_{d_s \in D} \mathcal{C}(u, d_s)$$

$$\mathcal{C}_{j \leftarrow \bullet}(u, \infty) = \sum_{d_s \in D} \mathcal{C}_{j \leftarrow \bullet}(u, d_s)$$

$$\mathcal{C}_{j \rightarrow \bullet}(u, \infty) = \sum_{d_s \in D} \mathcal{C}_{j \rightarrow \bullet}(u, d_s)$$

where $\mathcal{C}(u, \infty)$ are local network connectedness measures aggregated over frequencies with $H \rightarrow \infty$.

Obtaining our Measures I

We use Quasi Bayesian Local Likelihood methods of Petrova (2019) to obtain our measures of network connections. We can re-write (1) as:

$$\mathbf{X}_{t,T} = \bar{\mathbf{X}}'_{t,T}\phi_{t,T} + \mathbf{\Sigma}_{t/T}^{-\frac{1}{2}}\boldsymbol{\eta}_{t,T} \tag{2}$$

The local likelihood function at time period k is given by:

$$\mathsf{L}_k\left(\mathbf{X}|\theta_k, \mathbf{\Sigma}_k, \bar{\mathbf{X}}\right) \propto$$
$$|\mathbf{\Sigma}_k|^{\text{trace}(\mathbf{D}_k)/2} \exp\left\{-\frac{1}{2}(\mathbf{X} - \bar{\mathbf{X}}'\phi_k)'\left(\mathbf{\Sigma}_k \otimes \mathbf{D}_k\right)(\mathbf{X} - \bar{\mathbf{X}}'\phi_k)\right\}$$

Obtaining our Measures II

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

$$
\begin{aligned}
\phi_k | \boldsymbol{\Sigma}_k &\backsim \mathrm{N}\left(\phi_{0k}, (\boldsymbol{\Sigma}_k \otimes \boldsymbol{\Xi}_{0k})^{-1}\right) \\
\boldsymbol{\Sigma}_k &\backsim \mathrm{W}\left(\alpha_{0k}, \boldsymbol{\Gamma}_{0k}\right)
\end{aligned}
$$

The prior and weighted likelihood function implies a Normal-Wishart quasi posterior distribution for $\phi_k | \, \boldsymbol{\Sigma}_k$ for $k = \{1, \ldots, T\}$.

Obtaining our Measures III

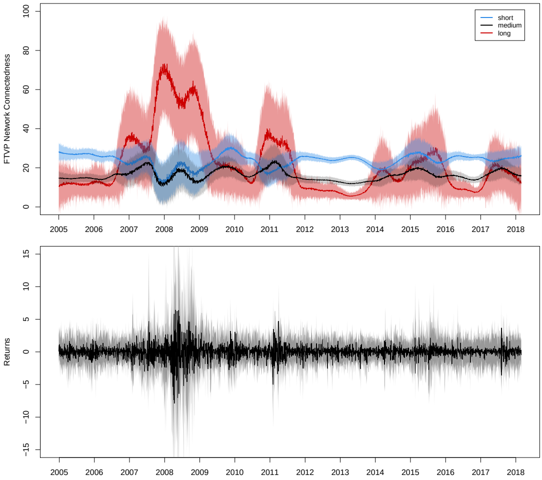
Formally let $\mathbf{A} = (\bar{x}_1', \ldots, \bar{x}_T')'$ and $\mathbf{Y} = (x_1, \ldots, x_T)'$ then:

$$\begin{aligned} \phi_k | \mathbf{\Sigma}_k, \mathbf{A}, \mathbf{Y} &\backsim \mathrm{N}\left(\tilde{\phi}_k, \left(\mathbf{\Sigma}_k \otimes \tilde{\mathbf{\Xi}}_k\right)^{-1}\right) \\ \mathbf{\Sigma}_k &\backsim \mathrm{W}\left(\tilde{\alpha}_k, \tilde{\mathbf{\Gamma}}_k^{-1}\right) \end{aligned}$$

with quasi posterior parameters

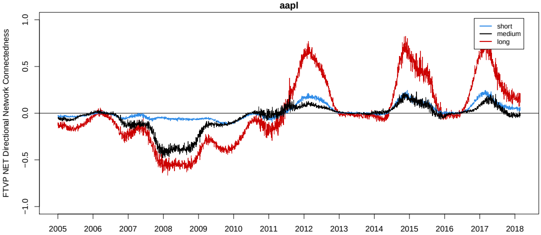
$$\begin{aligned} \tilde{\phi}_k &= \left(\mathbf{I}_N \otimes \tilde{\mathbf{\Xi}}_k^{-1}\right)\left[\left(\mathbf{I}_N \otimes \mathbf{A}'\mathbf{D}_k\mathbf{A}\right)\hat{\phi}_k + \left(\mathbf{I}_N \otimes \mathbf{\Xi}_{0k}\right)\phi_{0k}\right] \\ \tilde{\mathbf{\Xi}}_k &= \tilde{\mathbf{\Xi}}_{0k} + \mathbf{A}'\mathbf{D}_k\mathbf{A} \\ \tilde{\alpha}_k &= \alpha_{0k} + \sum_{t=1}^{T} \varrho_{kt} \\ \tilde{\mathbf{\Gamma}}_k &= \mathbf{\Gamma}_{0k} + \mathbf{Y}'\mathbf{D}_k\mathbf{Y} + \mathbf{\Phi}_{0k}\mathbf{\Gamma}_{0k}\mathbf{\Phi}_{0k}' - \tilde{\mathbf{\Phi}}_k\tilde{\mathbf{\Gamma}}_k\tilde{\mathbf{\Phi}}_k' \end{aligned}$$

TFVP Connectedness as determinant of cross-section of returns

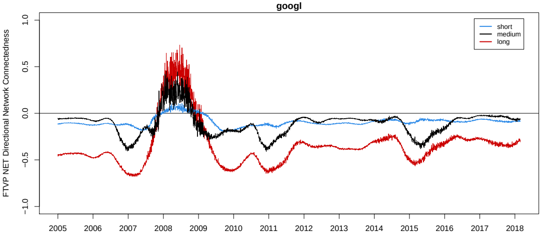


TFVP Directional Connectedness

- AAPL net transmitter in long run: 2012, 2015, 2017



- GOOG almost net receiver



Horizon Specific Factors and Horizon Specific Network Betas I

- We sort S&P500 assets based on horizon specific NET DIRECTIONAL CONNECTEDNESS for short-term, medium-term and long-term.
- Factors are based on long-short "portfolios" of horizon specific transmitters and recipients. We call these TMR(d) factors, for d $\in$ {S,M,L,}.
- We also construct a TMR(∞) factor that uses network connections summed over frequency bands d.
- We run a univariate version of Petrova (2019) on all stocks accounting for our TMR(∞) factors as well as the Fama French 5 Factor model:

$$R_t^i - R_{f,t} = \alpha_t^i + \beta_t^{i,\text{TMR}(\infty)} \text{TMR}(\infty)_t + \text{Fama French 5 Factors}$$

Horizon Specific Factors and Horizon Specific Network Betas II

- We then conduct portfolio sorts based on the $\beta_t^{i,\mathrm{TMR}(\infty)}$ estimates and estimate (quintile) portfolio excess returns on our TMR(d) factors:

$$
\begin{aligned}
R_{p,t} - R_{f,t} &= \alpha_{p,t} + \beta_{p,t}^{\mathrm{TMR(S)}}\mathrm{TMR(S)}_t + \beta_{p,t}^{\mathrm{TMR(M)}}\mathrm{TMR(M)}_t \\
&+ \beta_{p,t}^{\mathrm{TMR(L)}}\mathrm{TMR(L)}_t + \text{Fama French 5 Factors}
\end{aligned}
$$

Portfolio Sorts

A: VW Portfolios	$E(R_p)$	α_t^{FF5}	TMR(S)	TMR(M)	TMR(L)
1 Low TMR(T)	0.025	-0.003	0.070	-0.296	-0.370
2	0.026	0.001	0.267	0.253	0.037
3	0.030	0.005	0.517	0.467	0.269
4	0.021	-0.003	0.752	0.658	0.489
5 High TMR(T)	-0.007	-0.035	1.176	0.942	1.221
High-Low	-0.032	-0.03145	1.10633	1.23875	1.59154
t-stat/ [95% post. coverage]	-1.856	[-0.0315 -0.0314]	[1.1057 1.107]	[1.238 1.239]	[1.591 1.592]
B: EW Portfolios	$E(R_p)$	α_t^{FF5}	TMR(S)	TMR(M)	TMR(L)
1 Low TMR(T)	0.014	-0.014	0.124	-0.212	-0.242
2	0.021	-0.003	0.323	0.256	0.081
3	0.019	-0.005	0.554	0.470	0.320
4	0.007	-0.018	0.811	0.693	0.506
5 High TMR(T)	-0.036	-0.061	1.319	1.032	1.310
High-Low	-0.050	-0.04695	1.19571	1.24361	1.55202
t-stat/ [95% post. coverage]	-2.700	[-0.04698 -0.04692]	[1.19517 1.1963]	[1.24301 1.2442]	[1.5515 1.553]

Robustness Checks

- Alternative definitions of factor construction
- Simple OLS and rolling regression estimates
- Univariate QBLL regressions and portfolio sorts confirm baseline results
- Rolling Fama MacBeth Regressions to get prices of risk (confirm what we see in portfolio sorts and full sample results)
- Extensions to include a battery of other factors (e.g. Market Liquidity, VIX, Momentum, Conditional Skewness, Conditional Kurtosis)

Conclusions I

- We provide a theoretical framework where a factor decomposition allows us to price horizon specific network risk for assets on the market.
- **We propose new time-frequency varying connectedness measures that permit one to decompose networks into horizon specific bands of interest to investors and economists.**
- **We re-define the meaning of big in ”Big Data” and conduct empirical analysis using all constituents listed on the S&P500 and show that network based factors are priced in the cross-section of asset returns.**

Thank you for your attention.

References I

Cogley, T. and Sargent, T. J. (2005). Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US. *Review of Economic Dynamics*, 8(2):262–302.

Diebold, F. X. and Yılmaz, K. (2014). On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms. *Journal of Econometrics*, 182(1):119–134.

Ellington, M. (2018a). The Case for Divisia Monetary Statistics: A Bayesian Time-varying Approach. *Journal of Economic Dynamics and Control*, 96:26–41.

Ellington, M. (2018b). Financial Market Illiquidity Shocks and Macroeconomic Dynamics: Evidence from the UK. *Journal of Banking & Finance*, 89:225–236.

Ellington, M. (2019). The Empirical Relevance of the Shadow Rate and the Zero Lower Bound. *Available from: SSRN*.

References II

Ellington, M., Florackis, C., and Milas, C. (2017). Liquidity Shocks and Real GDP Growth: Evidence from a Bayesian Time-varying Parameter VAR. *Journal of International Money and Finance*, 72:93–117.

Pesaran, H. H. and Shin, Y. (1998). Generalized Impulse Response Analysis in Linear Multivariate Models. *Economics Letters*, 58(1):17–29.

Petrova, K. (2019). A Quasi-Bayesian Local Likelihood Approach to Time Varying Parameter VAR Models. *Journal of Econometrics*, 212(1):286–306.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies*, 72(3):821–852.