The impact of network connectivity on factor exposures, expected returns and portfolio diversification

Monica Billio[1] **Massimiliano Caporin**[2]
Roberto Panzica[3] **Loriana Pelizzon**[1,3]

[1]University Ca' Foscari Venezia (Italy)

[2]University of Padova (Italy)

[3]Goethe University Frankfurt (Germany)

Helsinki, September 2015

Introduction

- Linear factor models (CAPM, APT) have a central role in finance for both pricing and risk evaluation
- From a pricing perspective, factor risk premium is combined with asset exposure to risk factors to recover the equilibrium asset return
- From a risk analysis perspective, we decompose an asset (portfolio) total risk into systematic and idiosyncratic components
- Systematic risks comes from the dependence of returns on common factors, while idiosyncratic risks are asset-specific
- But...
- There is also a recent consensus on the existence of network interconnections/systemic risks that affect asset pricing and risk

(Challenging) Research questions

- **What is the impact of network interconnections on the asset return loading to common factors?**

- **How does network interconnections affect the expected returns in a factor model?**

- **Does network interconnections endanger the power of diversification?**

Literature review

- An increasing literature investigates the role of interconnections between different firms and sectors, functioning as a potential propagation mechanism of idiosyncratic shocks throughout the economy.
- Canonical idea: microeconomic shocks average out and thus, only have negligible aggregate effects (Lucas, 1977, among others). Similarly, these shocks have little impact on asset prices.
- However:
- Acemoglou et al. (2011) use network structure to show the possibility that aggregate fluctuations may originate from microeconomic shocks to firms. Such a possibility is discarded in standard macroeconomics models due to a "diversification argument".
- Shock propagation in static networks: Horvath (1998, 2000), Dupor (1999), Shea (2002), and Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2011).

Measuring systemic links

- Role of idiosyncratic risk in asset pricing: The CAPM predicts that only systematic risk is priced and expected excess returns satisfy two-fund separation. This prediction is contradicted by:
- Ozsoylev and Walden (2011): study a static model of asset price formation in large information networks, mainly centred on the relation between price volatility and network connectedness.
- Ang, Hodrick, Xing, and Zhang (2006): who show that idiosyncratic volatility risk is priced in the cross-section of expected stock returns, a regularity which is not subsumed by size, book-to-market, momentum, or liquidity effects.
- Buraschi and Porchia (2014) and Branger et al. (2014): dynamic network connectivity has implications on diversification and asset pricing

Pricing and diversification in multifactor models

- General multi-factor model for a $k-$dimensional vector of risky assets (m risk factors)

$$\mathbf{R}_t = \mathbb{E}\left[\mathbf{R}_t\right] + B\mathbf{F}_t + \epsilon_t$$

- Expected returns depend on the factor risk premiums Λ

$$\mathbb{E}\left[\mathbf{R}_t\right] = r_f + B\Lambda$$

- Standard total risk decomposition

$$\begin{aligned} VAR\left[\mathbf{R}_t\right] &= BVAR\left[\mathbf{F}_t\right]B' + VAR\left[\epsilon_t\right] \\ \Sigma_R &= B\Sigma_F B' + \Omega \end{aligned}$$

Pricing and diversification in multifactor models

- Portfolio risk decomposition (ω being a $k-$dimensional vector of portfolio weights)

$$\begin{aligned} VAR\left[\omega' \mathbf{R}_t\right] &= \omega' B VAR\left[\mathbf{F}_t\right] B' \omega + \omega' VAR\left[\epsilon_t\right] \omega \\ \omega' \Sigma_R \omega &= \omega' B \Sigma_F B' \omega + \omega' \Omega \omega \end{aligned}$$

- Diversification benefit

$$lim_{k\rightarrow\infty} \omega' \Omega \omega = \nu$$

- Special case: uncorrelated idiosyncratic shocks with average variance $\bar{\sigma}^2$

$$lim_{k\rightarrow\infty} \omega' \Omega \omega = \frac{1}{k} \bar{\sigma}^2 = 0$$

Adding network connections in multifactor models

- Network connections represent contemporaneous relations across k assets that co-exists with the dependence on $m << k$ common risk factors
- We are agnostic on the network estimation approach and simply condition the model to the availability of the network
- Network connections are contemporaneous relations across endogenous variables, and thus are included in the matrix A below

$$A\left(\mathbf{R}_t - \mathbb{E}\left[\mathbf{R}_t\right]\right) = B\mathbf{F}_t + \epsilon_t$$

- The simultaneous equation system is not identified unless we impose some restriction

Adding Network connections in multifactor models

- **Assumption 1:** the idiosyncratic shocks are uncorrelated, that is Ω is a diagonal matrix (assumption already taken into account in multi-factor models)
- **Assumption 2:** the matrix of contemporaneous relations has a structure driven by network links
- In networks, nodes are connected (in general) to a subset of the network total number of nodes, where connections represent links across assets; shocks are transmitted through the network
- Intuition: connected assets are *neighbours* assets $\Rightarrow$ Spatial Econometrics tools introduced to capture the spatial dependence across subjects

Adding Network connections in multifactor models

- Networks can be used to define spatial matrices
- An example of a spatial matrix

$$W = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

- Spatial matrices can be row normalized
- Spatial matrices might be generalized in such a way values are associated with the intensity of the link across assets

A general framework with Network connections and systematic risks

- By means of spatial matrices we can impose a structure on matrix A and rewrite the simultaneous equation system

$$
\begin{aligned}
A &= I - \rho W \\
(\mathbf{R}_t - \mathbb{E}[\mathbf{R}_t]) &= \rho W (\mathbf{R}_t - \mathbb{E}[\mathbf{R}_t]) + B\mathbf{F}_t + \epsilon_t
\end{aligned}
$$

- The coefficient ρ represents the impact coming from neighbours and by now we assume it is a scalar
- This simultaneous equation system corresponds to a Spatial Auto Regression Panel model where the covariates (risk factors) are common across all subjects (at least in a simplified representation)

A general framework with Network connections and systematic risks

- The model has an obvious reduced-form representation

$$\mathbf{R}_t = \mathbb{E}\left[\mathbf{R}_t\right] + A^{-1}B\mathbf{F}_t + A^{-1}\epsilon_t$$

$$\mathbf{R}_t = \mathbb{E}\left[\mathbf{R}_t\right] + \left(I - \rho W\right)^{-1}B\mathbf{F}_t + \left(I - \rho W\right)^{-1}\epsilon_t$$

- The introduction of network connections allows
  - **Capturing residual correlations in traditional linear factor models**
  - **Decomposing the exposure to common factors (reduced-form) into the structural exposure and in an exposure induced by network connections**

The impact of common factors and networks

- Reference model

$$(I - \rho W)\left(\mathbf{R}_t - \mathbb{E}\left[\mathbf{R}_t\right]\right) = B\mathbf{F}_t + \epsilon_t$$

- It holds that

$$(I - \rho W)^{-1} = I + \rho W + \rho^2 W^2 + \rho^3 W^3 \ldots$$

- Therefore the model corresponds to

$$\mathbf{R}_t = \mathbb{E}\left[\mathbf{R}_t\right] + B\mathbf{F}_t + \sum_{j=1}^{\infty} \rho^j W^j B\mathbf{F}_t + \sum_{j=1}^{\infty} \rho^j W^j \epsilon_t + \epsilon_t$$

The impact of common factors and networks

- Four components in the returns
  - Structural exposure to common factors: $B\mathbf{F}_t$
  - Network exposure to common factors: $\sum_{j=1}^{\infty} \rho^j W^j B\mathbf{F}_t$
  - Idiosyncratic shocks: ϵ_t
  - Network impact of idiosyncratic shocks: $\sum_{j=1}^{\infty} \rho^j W^j \epsilon_t$
- Focus on factor exposures: **if $\rho = 0$ back to standard linear factor models; network exposure acts as an inflating factor on structural exposures (assuming W and ρ positive)**

The impact of common factors and networks

- Network links might induce exposure of an assets to a common factor even if the given asset has a zero structural exposure to the factor
- Take a multifactor model and focus on asset i that is linked to asset $i+1$ only

$$W = \begin{bmatrix} & \vdots & \\ \mathbf{0}_i & 1 & \mathbf{0}_{k-i-1} \\ & \vdots & \end{bmatrix}$$

- Assume also that asset i depends only on factor 1, and asset $i+1$ depends only on factors 1 and 2

$$\beta = \begin{bmatrix} & \vdots & & \\ \beta_{1,i} & 0 & 0 & 0 \\ \beta_{1,i+1} & \beta_{1,i+1} & 0 & 0 \\ & \vdots & & \end{bmatrix},$$

The impact of common factors and networks

- Now consider the factor exposure of the $i-$th asset

$$\beta_{1,i}F_{1,t} + \rho\beta_{1,i+1}F_{1,t} + \rho\beta_{2,i+1}F_{2,t} + \sum_{j=2}^{\infty}\left(\rho^{j}W^{j}\beta F_{t}\right)|_{i}$$

- Asset i is also exposed to factor 2, the exposure is induced by the link existing between asset i and asset $i+1$

The impact of common factors and networks

- Focus now on pricing issues: under equilibrium expected returns equal (recover them from the reduced form representation)

$$\mathbb{E}\left[\mathbf{R}_t\right] = r_f + B\Lambda + \sum_{j=1}^{\infty} \rho^j W^j B\Lambda$$

- If $\rho > 0$ and elements in W are positive, the presence of network links inflates the loading to the factors with an impact on the asset expected returns
- If W (or even ρ) change over time we have a dynamic in the expected returns driven by the network links or by the impact of network exposure
- Expected returns increase as a consequence to an increase in ρ or a change in W with subsequent effects on prices

Risk decomposition

- A risk decomposition applies also in the presence of network exposure

$$
\begin{aligned}
VAR\left[\mathbf{R}_t\right] &= A^{-1}BVAR\left[\mathbf{F}_t\right]B'\left(A^{-1}\right)' + A^{-1}VAR\left[\epsilon_t\right]\left(A^{-1}\right)' \\
\Sigma_R &= A^{-1}B\Sigma_F B'\left(A^{-1}\right)' + A^{-1}\Omega\left(A^{-1}\right)'
\end{aligned}
$$

- In this case, the systematic and idiosyncratic contributions to total risk are also influenced by the existence of Network connections
- Obviously, if network links are absent, that is $\rho = 0$ we get back to the traditional risk decomposition

Risk decomposition

- We can play around this decomposition to recover a more insightful one

$$
\begin{aligned}
\Sigma_R &= A^{-1}B\Sigma_F B'\left(A^{-1}\right)' + A^{-1}\Omega\left(A^{-1}\right)' \\
&= \bar{B}\Sigma_F\bar{B}' + \mathcal{A}\Omega\mathcal{A}' \\
&= \bar{B}\Sigma_F\bar{B}' + \mathcal{A}\Omega\mathcal{A}' \pm B\Sigma_F B' \pm \Omega \\
&= \underbrace{B\Sigma_F B'}_{i} + \underbrace{\Omega}_{ii} + \underbrace{\left(\bar{B}\Sigma_F\bar{B}' - B\Sigma_F B'\right)}_{iii} + \underbrace{\left(\mathcal{A}\Omega\mathcal{A}' - \Omega\right)}_{iv}
\end{aligned}
$$

- We have thus four terms in the risk decomposition
  - i The systematic component
  - ii The idiosyncratic component
  - iii The network impact on the systematic component
  - iv The network impact on the idiosyncratic component

Risk decomposition

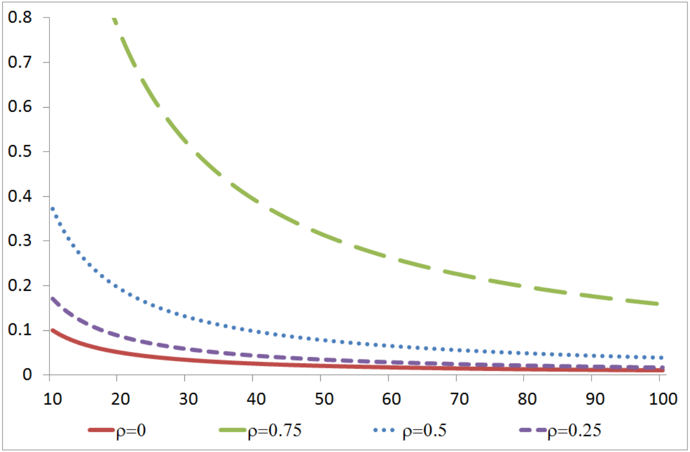
- This has effects on the diversification benefits which can be analytically derived in a special case
- Consider K uncorrelated idiosyncratic shocks with average variance $\bar{\sigma}^2$ and a W matrix where all assets are linked to each other, we have

$$lim_{K \to \infty} \omega' \mathcal{A} \Omega_\eta \mathcal{A}' \omega = \frac{K + \rho^2}{(K + \rho)^2 (\rho - 1)^2} \bar{\sigma}^2 = 0$$

- Diversification benefits still present but the decrease of the idiosyncratic component of the portfolio variance is much slower

Risk decomposition

- Portfolio idiosyncratic risk across different ρ levels and increasing number of assets. The case $\rho = 0$ corresponds to the absence of spatial links and is the standard result for diversification benefits.



Model estimation

- The use of spatial matrices in deriving the structure for the contemporaneous relations matrix A allows imposing a large number of restrictions on A
- From a different viewpoint, A is driven by a small number of parameters
- Nevertheless, the restrictions on A are not sufficient to achieve identification of the parameters driving A, and the assumption of uncorrelated idiosyncratic structural shocks is crucial
- This also impose the approach to parameter estimation which must be a Full Information Maximum Likelihood method
- Despite the number of parameters is large, the likelihood function can be easily concentrated with respect to the risk factor loadings as well as with respect to the idiosyncratic shocks variances

Simulations

- Questions: what is the effect of neglecting network links and estimating the reduced-form linear factor model?
- Simulations: 100 assets, one common factor with volatility equal 15% yearly, betas to the common factor $U \sim (0.8, 1.2)$, idiosyncratic volatilities $U \sim (10\%, 25\%)$
- Spatial matrix W is random with elements following $w_i \sim Bern\,(0.3)$
- Evaluate the effects of different ρ values (baseline 0.5), different density of W, different dispersion of betas, different volatility of the common factor, presence of dynamic in W

Simulations

- First, focus on the distance between the *true* structural betas and the estimated reduced form betas under model misspecification: we simulate from a model with W and estimate a standard linear factor model
- Second, consider the average correlation across the misspecified model residuals
- Simulated series have length equal to 200, 500 and 1000 observations (grey, dashed, bold)

Simulations

- Distortions and standard deviations across ρ values for estimated betas versus structural betas

T	$\rho = 0$		$\rho = 0.25$		$\rho = 0.5$		$\rho = 0.75$	
	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev
Linear factor model								
200	0.000	0.026	0.337	0.027	1.013	0.028	3.033	0.043
500	0.000	0.016	0.337	0.017	1.011	0.018	3.034	0.027
1000	0.000	0.011	0.337	0.012	1.012	0.012	3.035	0.019
Model with network exposure								
200	-0.072	0.091	-0.088	0.107	-0.096	0.107	-0.119	0.133
500	-0.032	0.043	-0.040	0.053	-0.045	0.061	-0.059	0.072
1000	-0.016	0.026	-0.020	0.032	-0.022	0.034	-0.031	0.045

- Some underestimation for the correct model, increasing with ρ

Simulations

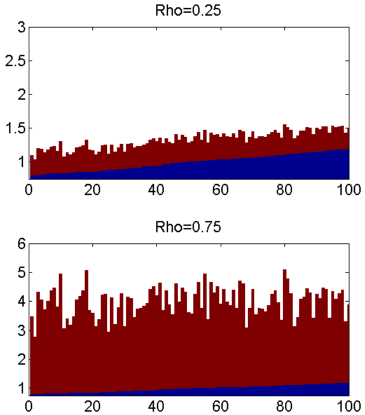
- Average correlation across residuals for different ρ values

T	$\rho = 0$		$\rho = 0.25$		$\rho = 0.5$		$\rho = 0.75$	
	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev	Mean	Std.dev
Linear factor model								
200	0.000	0.003	0.009	0.007	0.034	0.014	0.149	0.035
500	0.000	0.002	0.009	0.007	0.033	0.013	0.149	0.035
1000	0.000	0.001	0.009	0.006	0.033	0.013	0.149	0.035
Model with network exposure								
200	-0.001	0.004	-0.002	0.003	-0.002	0.003	-0.002	0.003
500	-0.001	0.002	-0.001	0.002	-0.001	0.002	-0.001	0.002
1000	0.000	0.001	0.000	0.001	0.000	0.001	-0.001	0.001

- Misspecified model show residual correlations increasing with ρ

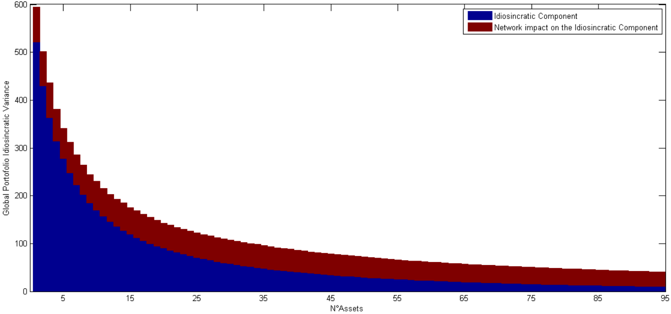
Simulations

- Beta values across assets: structural betas (in blue) and augmented betas (in red) across different values for ρ with the random matrix W.



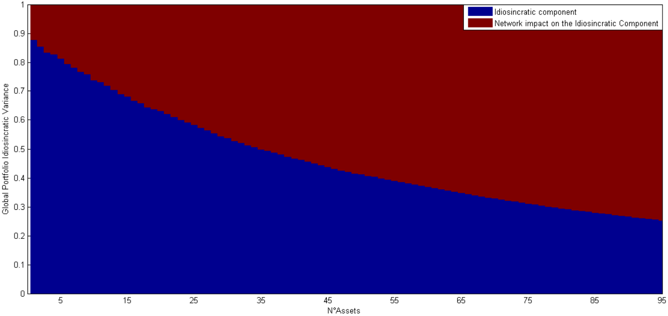
Simulations

- Order assets with respect to their reduced form risk; consider a $1/N$ portfolio and decompose the portfolio risk for $N = 5, 6, \ldots 100$
- Absolute decomposition



Simulations

- Relative decomposition



- Diversification benefits decrease and the role of network-induced risk increases with N

Model generalizations

- Heterogeneous network impact: ρ is asset specific,

$$\begin{aligned} A &= I - \mathcal{R}W \\ \mathcal{R} &= diag\left(\rho_1, \rho_2, \ldots, \rho_N\right) \end{aligned}$$

- Time-change in networks: W depends on the time

$$A_t = I - \mathcal{R}W_t$$

- Plurality of networks: the networks could be more than one

$$A = I - \sum_{j=1}^{p} \mathcal{R}_j W_j$$

Estimation

- We do not impose a method for estimating the systemic links, those can come from any approach or measurement framework; we do impose that the W is known before estimating the other model parameters
- In this work we estimate W by means of Granger causality, see Billio et al. (2012)
- Estimation of the factor loadings, spatial coefficients, and idiosyncratic shocks by Full Information Maximum Likelihood; computational advantages come from the possibility of concentrating out the factor loadings
- Impose technical assumptions guaranteeing the non singularity of A_t and the identification of the parameters in $\mathcal{R}$

Data description

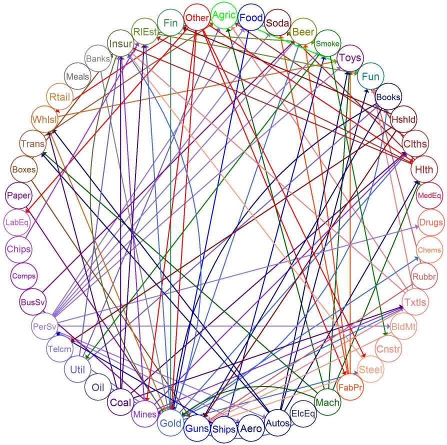
- We consider the industrial sector indices available from the Kenneth French website
- We take the decomposition of the market into 48 economic sectors/industries
- The factors we use are macroeconomic factors(Industrial Production, Inflation, Oil, US dollar Index, short term rate, Term Spread, Credit spread.
- Data are considered at the daily and monthly frequencies

Contemporaneous links

- Networks, monitoring connections across economic sectors, have been estimated by means of Granger Causality tests on daily data depurated by Garch(1,1)
- Estimation of the networks is based on a daily date over yearly samples
- We thus have a sequence of matrices W_t with time index evolving over years
- On spatial returns models estimated on frequencies higher than the year this induces a time variation in the model coefficients and mild heteroskedasticity over reduced-form innovations

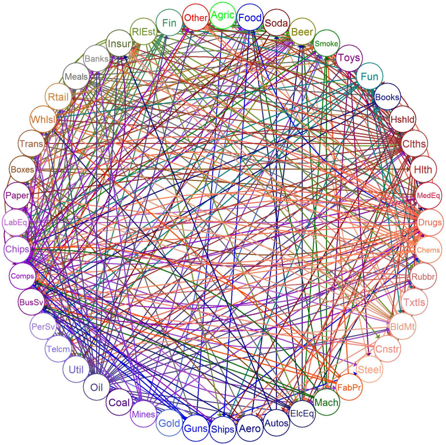
Contemporaneous links

- Estimated network for the year 2004



Contemporaneous links

- Estimated network for the year 2008

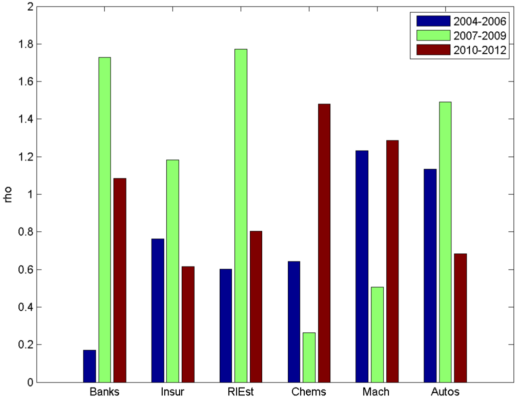


Spatial Model output

- The model has been estimated using monthly data and for different samples, in particular we distinguish three periods.
- the first period 2004-2006 before the crisis
- the second period 2007-2009 during the crisis
- the third period 2010-2012 after the crisis

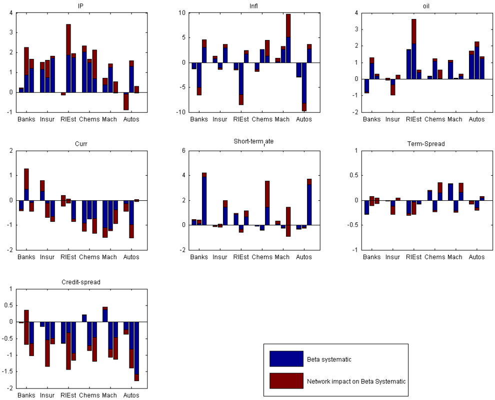
Model output

- Estimated ρ for selected sectors and over different periods



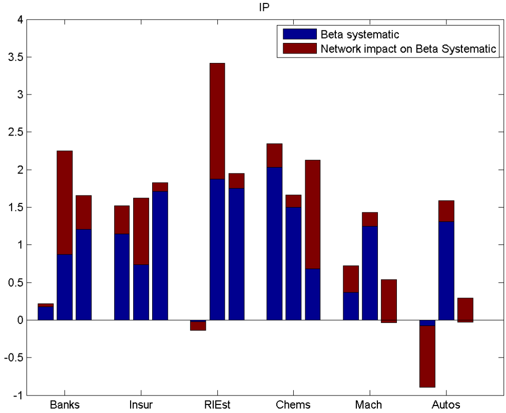
Model output

- Estimated β for selected sectors and over different periods 2004-2006, 2007-2009, 2010-2012.



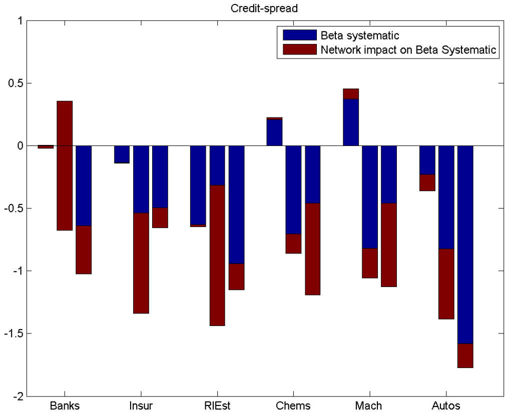
Model output

- Estimated β for selected sectors and over different periods 2004-2006, 2007-2009, 2010-2012.



Model output

- Estimated β for selected sectors and over different periods 2004-2006, 2007-2009, 2010-2012.



Future developments

- Empirical evaluation focusing the US equity market by economic sectors
- Two cases: networks from Granger causality and from across-sector sales
- Underlying risk factors associated with macroeconomic variables
- Comparison between US and Europe
- From the methodological point of view
- Time-variation in the ρ coefficients
- Different design of the W matrices taking into account the strength of the relation across subjects before max row normalization