Forecasting High Frequency Systemic Risk

Nikolaus Hautsch
Institute for Statistics and Econometrics
Humboldt-Universität zu Berlin and CASE, CFS, QPL

Lada M. Kyj
Quantitative Products Laboratory
Humboldt-Universität zu Berlin

Hajnice, February 2010

dbqpl quantitative products laboratory

Outline

Applicability I

Any problem with > 1 asset and risk considerations requires covariance forecasts

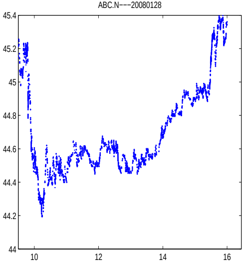
- Optimal trade scheduling for portfolio execution (OPX)
  - used by program trading desk
  - requires *short-horizon* covariance forecast (≈ 1 day)
  - requires *large universe* coverage (often > 500 assets)
- Structuring
  - optimal asset allocation in alpha and beta oriented strategies
  - tracking error minimization of ETFs
- Index arbitrage
  - selected (index) universe
  - very short term, possibly with intra-day updates

Status quo Risk Estimation

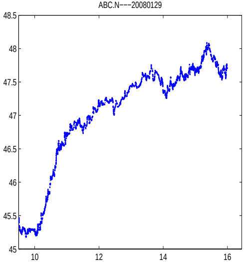
- BARRA risk models
  - factor structure with global coverage
  - forecast horizon between 1 week and 3 months
  - suitable for some structuring problems, but suboptimal for high frequency trading
- M-GARCH
  - CCC or DCC
  - short forecast horizons
  - cross-products of daily returns are noisy measures of covariance

growing demand for short-term & high dim. covariance forecasts
the aim is to leverage our high-frequency data resources

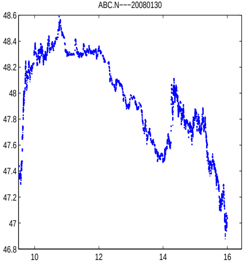
Why High Frequency Data?



$RV_t(1) = 0.70\%$
$RV_t(1/78) = 6.31\%$



$RV_t(1) = 71.53\%$
$RV_t(1/78) = 8.21\%$



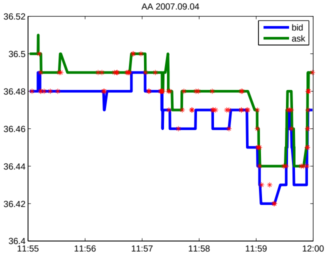
$RV_t(1) = 11.73\%$
$RV_t(1/78) = 6.18\%$

Econometrics of High Dimensional Covariance

- Challenges: numerically well-behaved
  - Is estimate invertible? (i.e. is it positive definite?)
  - Is estimate well-conditioned? (i.e are inversions numerically stable?)
- Requirements: As the number of assets increases need to
  - use more data (i.e. longer history)
    - *can use high frequency instead of longer history*
  - impose structure

High Frequency Challenges

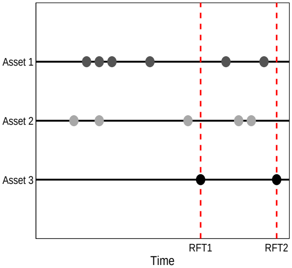
- Market microstructure effects
- Asynchronicity of observations in time, Epps (1979, JASA)
- Efficient sampling



Multivariate Realized Kernels

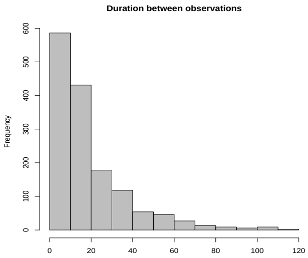
by Barndorff-Nielsen, Hansen, Lunde, Shephard (2008)

- Kernel based lead-lag estimator
- Prices are synchronized via Refresh Time Sampling (RTS).
- Addresses three features: synchronization, sparse sampling, and smoothing
- Sampling time-scale is driven by activity of least active process in the set



Hautsch, Kyj, Oomen (2009)

Instead of grouping according to fundamentals (i.e. energy, consumer goods, etc.), our approach groups according to characteristics of the observed price processes.



- Group assets into subsets according to duration time between observations

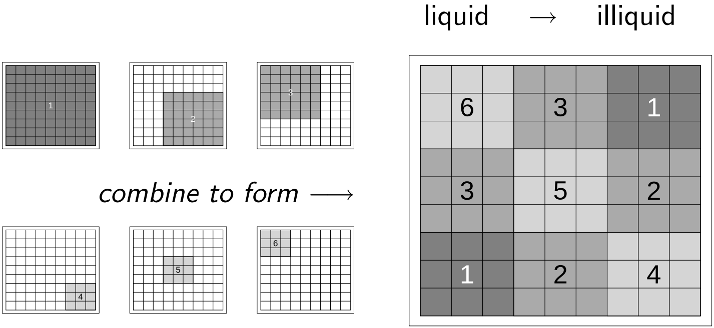
Blocking and Regularization Estimator (RnB)

1. Blocking: apply Multivariate Realized Kernel to blocks (RK)
   - Group similar assets into clusters
   - Compose blocks from similar groups
   - Different RTS time scale for each block
2. Regularization: impose invertibility/well-conditioned properties by "eigenvalue cleaning".

This framework provides a systematic method for computing high dimensional covariance matrices which addresses:

- asynchroneity
- market microstructure effects
- curse of dimensionality problem

Blocking



Mixed-Frequency Covariance Model

- Goal: Provide a short horizon (i.e. day) update to a long-horizon (i.e. 1 year), BARRA-like covariance estimates.
- Approach: A mixed frequency model is used to maintain the correlation structure of the long horizon while updating the volatility with the short horizon.
- Features:
  - Only a local stationarity constraint (i.e. long-horizon and short-horizon correlation structures are the same)
  - The approach is sufficiently flexible to allow for different estimates of long and short horizons.

Approach: Latent Factor Model

1. Factor model decomposition:
   - How? $\Sigma = Q\Lambda Q^T$
     $\Sigma = \sum_{i=1}^{r} \lambda_i q_i q_i^T + \sum_{i=r+1}^{p} \lambda_i q_i q_i^T$.
     $\Sigma_{F_r} = \sum_{i=1}^{r} \lambda_i q_i q_i^T + D$.
   - Why? Fan Fan, Lv (2008 JoEx) and Woodbury matrix identity.
2. Forecasting:
   - Fix eigenvectors and forecast eigenvalues.
   - like Diebold et. al. (2006 JoEx).

Mixed Frequency Model

This model allows the variance and correlation to evolve at two different frequencies,

$$\Sigma_t = D_{s(t)} R_{l(t)} D_{s(t)}$$

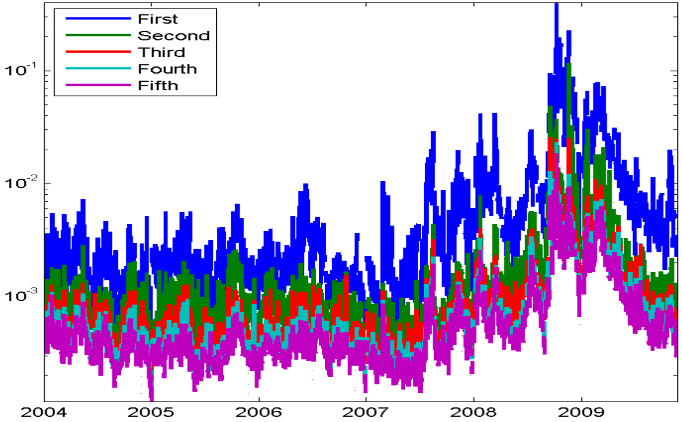
$$\text{where} \quad D_{s(t)} = \sqrt{diag(RC_{s(t)})} \quad \text{and} \quad R_{l(t)} = Corr(RC_{l(t)}).$$

Related Literature:

- CCC: Bollerslev (1990)
- DCC: Engle (2002)
- BARRA: Wang and Miller (2004)

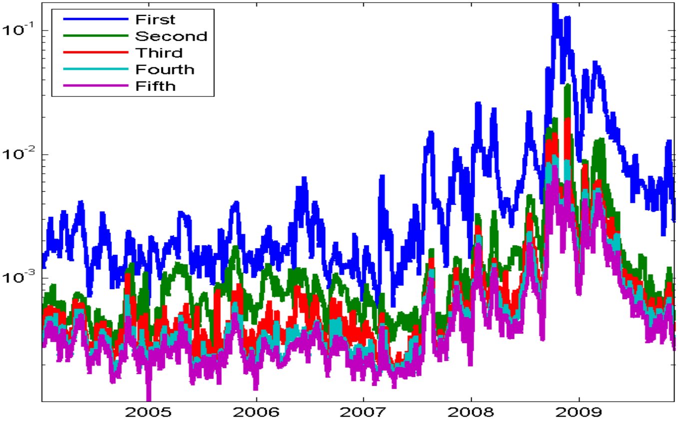
Evolution of PCA

$h = 1$



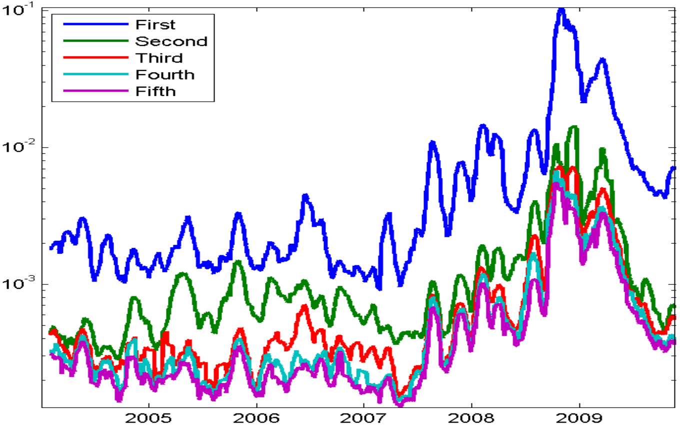
Evolution of PCA

$h = 5$



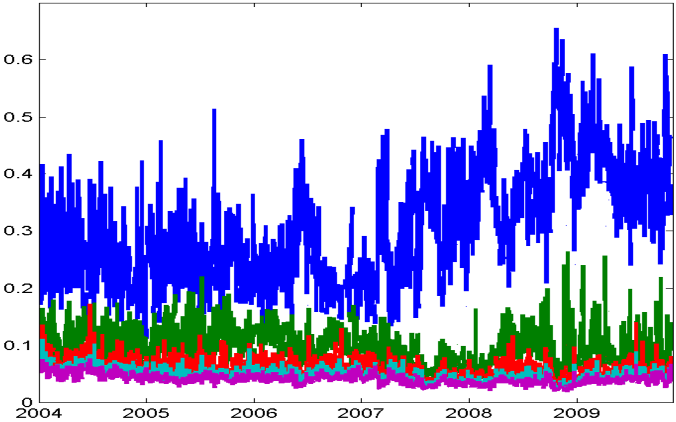
Evolution of PCA

$h = 22$



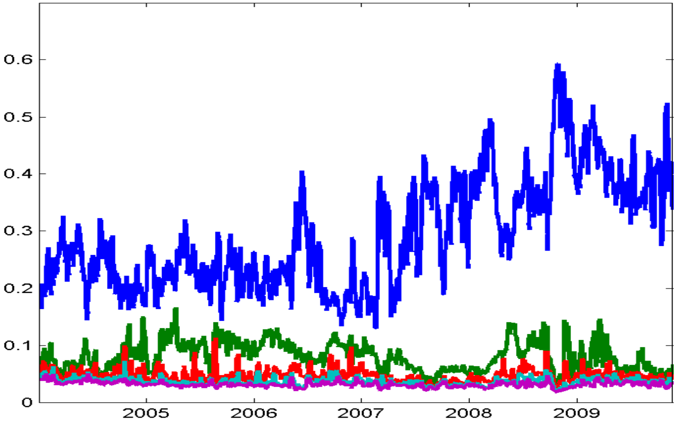
Percentage of Variation Explained

$h = 1$



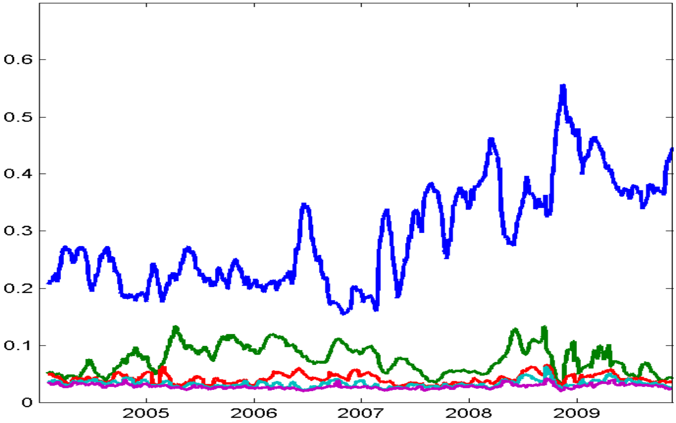
Percentage of Variation Explained

$h = 5$



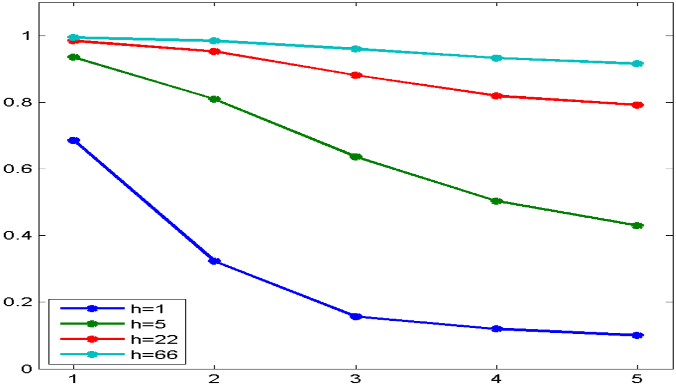
Percentage of Variation Explained

$h = 22$



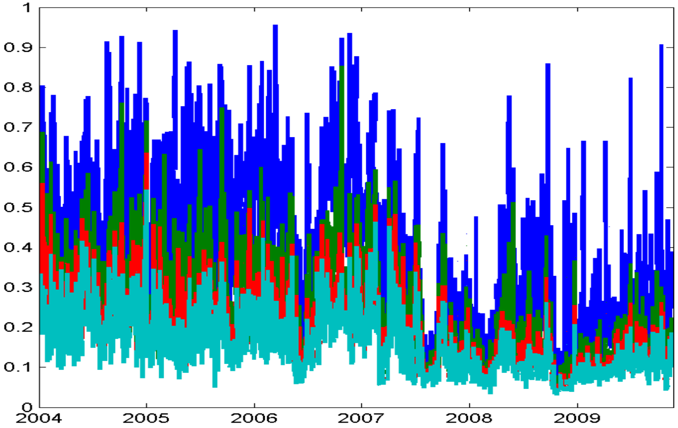
Vector Stability

$$1 - \overline{k_i} \quad \text{where} \quad k = \frac{\left\| q_{i,t} - q_{i,t-1} \right\|}{\left\| q_{i,t} - q_{j,t} \right\|}$$



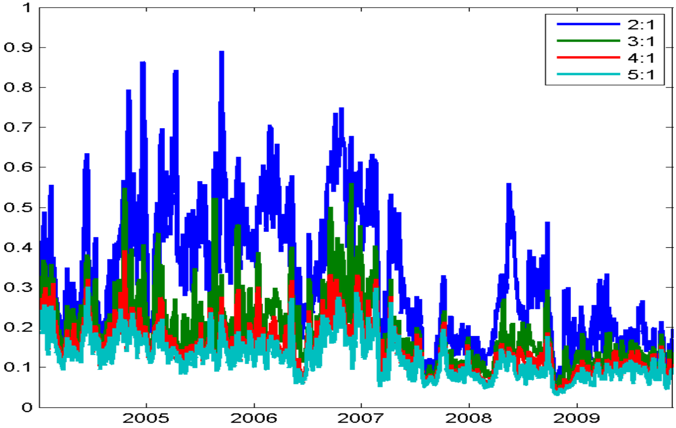
Eigenvalue Ratio

$$h = 1$$



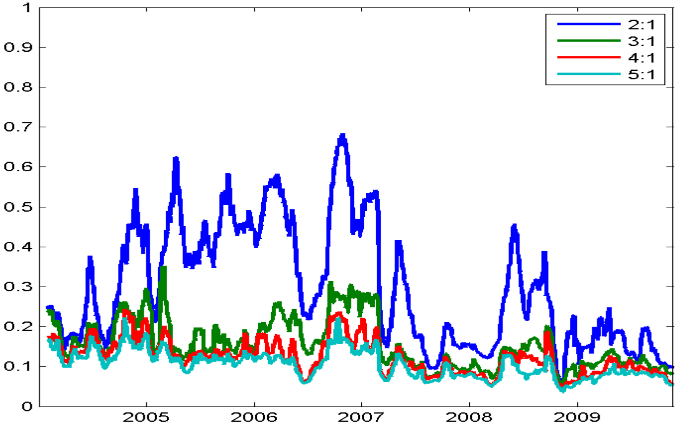
Eigenvalue Ratio

$h = 5$



Eigenvalue Ratio

$h = 22$



Global Minimum Variance Portfolios

We consider dynamics of GMV portfolios using various covariance estimation strategies.

- Solve the following optimization problem:

$$\begin{aligned} min_{w_t} \quad & w_t^{'}\Sigma_t w_t \\ s.t. \quad & w_t^{'}\iota = 1 \end{aligned}$$

where Σ is the $p \times p$ covariance matrix and ι is the unit vector. The GMV weights are

$$w_{t,GMV} = \frac{\Sigma_t^{-1}\iota}{\iota^{'}\Sigma_t^{-1}\iota}.$$

- Technique is used extensively in the literature:

GMV Portfolio Performance

	h=1		
	annualized σ		
	mean	Q(0.05)	Q(0.95)
EW	0.0916	0.0871	0.0962
DayF2	0.0892	0.0797	0.0989
WeekF2	0.0661	0.0630	0.0695
MonthF2	0.0668	0.0634	0.0704
LW22	0.0737	0.0702	0.0773
	h=5		
	annualized σ		
	mean	Q(0.05)	Q(0.95)
EW	0.0840	0.0807	0.0873
DayF2	0.1058	0.0808	0.1336
WeekF2	0.0695	0.0667	0.0720
MonthF2	0.0689	0.0663	0.0716
LW22	0.0779	0.0749	0.0811

Research Overview

- Developed and extensively tested in simulation and empirically the performance of a new high dimensional covariance estimation framework using high frequency data.
  - applicable to any class or universe of assets
  - computationally very tractable

- Have taken first steps toward forecasting high dimensional covariances.

Thank You

Vector Ratio Intuition