Dynamic Econometric Modeling of Volatility, Correlation and Connectedness in Financial Markets

Francis X. Diebold
University of Pennsylvania

July 23, 2012

Prologue: Reading

- Andersen, T.G., Bollerslev, T., Christoffersen, P.F. and Diebold, F.X. (2012), "Financial Risk Measurement for Financial Risk Management," in G. Constantinedes, M. Harris and Rene Stulz (eds.), *Handbook of the Economics of Finance*, Elsevier.
- Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2010), "Parametric and Nonparametric Volatility Measurement," in L.P. Hansen and Y. Ait-Sahalia (eds.), *Handbook of Financial Econometrics*. Amsterdam: North-Holland, 67-138.
- Andersen, T.G., Bollerslev, T., Christoffersen, P.F., and Diebold, F.X. (2006), "Volatility and Correlation Forecasting," in G. Elliott, C.W.J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting*. Amsterdam: North-Holland, 778-878.

Prologue: Reading

- Diebold, F.X. and Yilmaz, K. (2011), "On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms," Manuscript, University of Pennsylvania and Koc University.

- Aruoba, S.B. and Diebold, F.X. (2010), "Real-Time Macroeconomic Monitoring: Real Activity, Inflation, and Interactions," *American Economic Review*, 100, 20-24.

- Chen, F., Diebold, F.X. and Schorfheide, F. (2012), "A Markov-Switching Multi-Fractal Inter-Trade Duration Model, with Application to U.S. Equities," Manuscript, Huazhong University and University of Pennsylvania.
- Diebold, F.X. (2012), "100+ Years of Financial Risk Measurement and Management," in F.X. Diebold (ed.), *Financial Risk Measurement and Management*, Cheltenham, U.K. and Northampton, Mass.: Edward Elgar Publishing Ltd. (International Library of Critical Writings in Economics).

Penn
University of Pennsylvania

Prologue

- Throughout: Desirability of *conditional* risk measurement
- Aggregation level
  – Portfolio-level (aggregated) Risk measurement
  – Asset-level (disaggregated): Risk management
- Frequency of data observations
  – Low-frequency vs. high-frequency data
  – Parametric vs. nonparametric volatility measurement
- Object measured and modeled
  – Conditional variances, intervals, densities
- Dimensionality reduction in "vast" multivariate environments
  – Factor structure
- Beyond correlation: Network connectedness and systemic risk

Penn
University of Pennsylvania

What's in the Data?

Returns

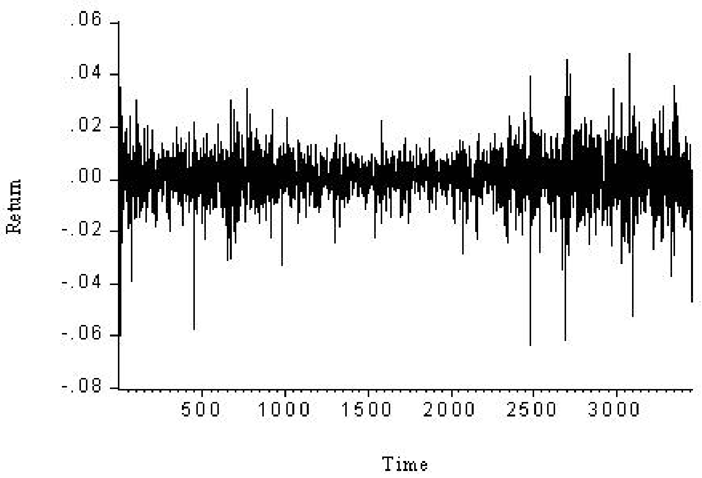


Figure: Time Series of Daily NYSE Returns.

Key Fact 1: Returns are Approximately Serially Uncorrelated

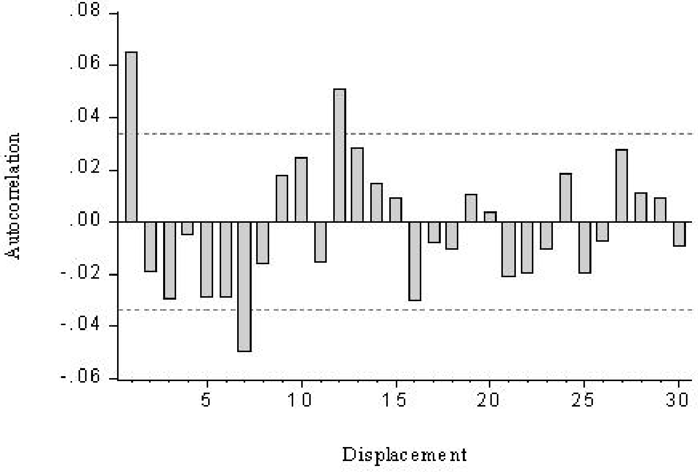
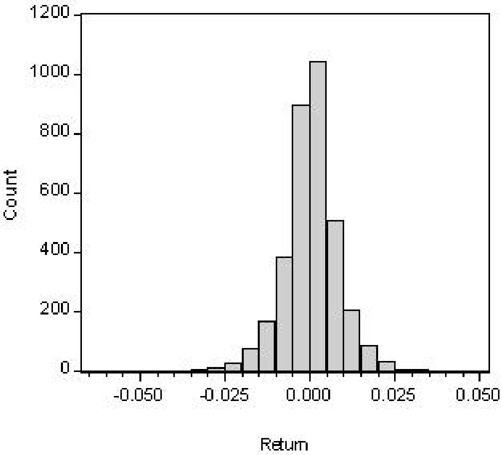


Figure: Correlogram of Daily NYSE Returns.

Key Fact 2: Returns are not Gaussian



Series: R	
Sample 1 3461	
Observations 3461	
Mean	0.000522
Median	0.000640
Maximum	0.047840
Minimum	-0.063910
Std. Dev.	0.008541
Skewness	-0.505540
Kurtosis	8.535016
Jarque-Bera	4565.446
Probability	0.000000

Figure: Histogram and Statistics for Daily NYSE Returns.

Key Fact 3: Returns are Conditionally Heteroskedastic I

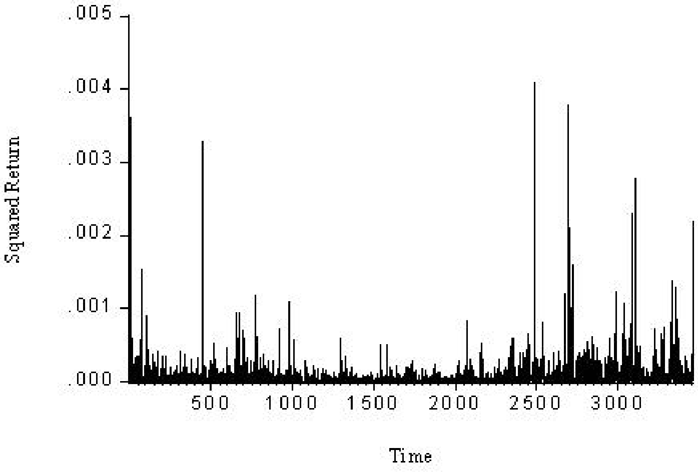


Figure: Time Series of Daily Squared NYSE Returns.

Key Fact 3: Returns are Conditionally Heteroskedastic II

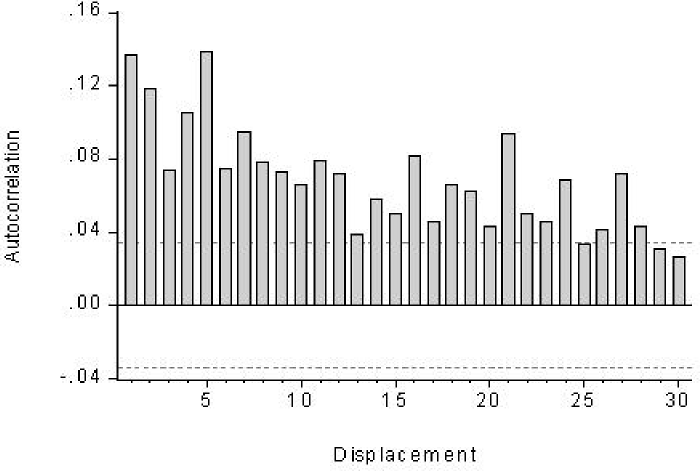


Figure: Correlogram of Daily Squared NYSE Returns.

Why Care About Volatility Dynamics?

Everything Changes when Volatility is Dynamic

- Risk management
- Portfolio allocation
- Asset pricing
- Hedging
- Trading

Risk Management

Individual asset returns:

$$r \sim (\mu, \Sigma)$$

Portfolio returns:

$$r_p = \lambda' r \sim (\lambda'\mu, \lambda'\Sigma\lambda)$$

If Σ varies, we need to track time-varying portfolio risk, $\lambda'\Sigma_t\lambda$

Portfolio Allocation

Optimal portfolio shares w^* solve:

$$min_w \ w'\Sigma w$$

$$\text{s.t. } w'\mu = \mu_p,$$

Importantly, $w^* = f(\Sigma)$

If Σ varies, we have $w_t^* = f(\Sigma_t)$

Penn
University of Pennsylvania

Asset Pricing I: Sharpe Ratios

Standard Sharpe:

$$\frac{E(r_{it} - r_{ft})}{\sigma}$$

Conditional Sharpe:

$$\frac{E(r_{it} - r_{ft})}{\sigma_t}$$

Penn
UNIVERSITY of PENNSYLVANIA

Asset Pricing II: CAPM

Standard CAPM:

$$(r_{it} - r_{ft}) = \alpha + \beta(r_{mt} - r_{ft})$$

$$\beta \;=\; \frac{cov((r_{it} - r_{ft}),\ (r_{mt} - r_{ft}))}{var(r_{mt} - r_{ft})}$$

Conditional CAPM:

$$\beta_t \;=\; \frac{cov_t((r_{it} - r_{ft}),\ (r_{mt} - r_{ft}))}{var_t(r_{mt} - r_{ft})}$$

Penn
UNIVERSITY of PENNSYLVANIA

Asset Pricing III: Derivatives

Black-Scholes:

$$C = N(d_1)S - N(d_2)Ke^{-r\tau}$$

$$d_1 = \frac{ln(S/K) + (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}}$$

$$d_2 = \frac{ln(S/K) + (r - \sigma^2/2)\tau}{\sigma\sqrt{\tau}}$$

$$P_C \;=\; BS(\sigma,\ ...)$$

(Standard Black-Scholes options pricing)

Completely different when σ varies!

Hedging

- Standard delta hedging

$$\Delta H_t \;=\; \delta\,\Delta S_t \;+\; u_t$$

$$\delta \;=\; \frac{cov(\Delta H_t,\;\Delta S_t)}{var(\Delta S_t)}$$

- Dynamic hedging

$$\Delta H_t \;=\; \delta_t\,\Delta S_t \;+\; u_t$$

$$\delta_t \;=\; \frac{cov_t(\Delta H_t,\;\Delta S_t)}{var_t(\Delta S_t)}$$

Penn
UNIVERSITY of PENNSYLVANIA

Trading

- Standard case: no way to trade on fixed volatility
- Time-varying volatility I: Options (indirect) Take position according to whether: $P_C > f(\sigma_{t+h,t}, \ldots)$, $P_C < f(\sigma_{t+h,t}, \ldots)$
- Time-varying volatility II: Volatility swaps (direct) Effectively futures contracts written on underlying "realized volatility"

Penn
University of Pennsylvania

Some Warm-Up

Unconditional Volatility Measures

Variance: $\sigma^2 \ = \ E(r_t - \mu)^2$ (or standard deviation: σ)

Mean Absolute Deviation: $MAD \ = \ E|r_t - \mu|$

Interquartile Range: $IQR \ = \ 75\% \ - \ 25\%$

$p\%$ Value at Risk (VaR^p)): x s.t. $P(r_t < x) = p$

Outlier probability: $P|r_t - \mu| \ > \ 5\sigma$

Tail index: γ s.t. $P(r_t > r) \ = \ k \ r^{-\gamma}$

Kurtosis: $K \ = \ E(r - \mu)^4/\sigma^4$

Dangers of HS-*VaR*, Take I

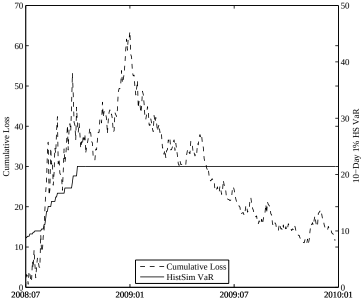


Figure: Cumulative S&P500 Loss (left-scale, dashed) and 1% 10-day HS-*VaR* (right scale, solid), July 1, 2008 - December 31, 2009. The dashed line shows the cumulative percentage loss on an S&P500 portfolio from July 2008 through December 2009. The solid line shows the daily 10-day 1% HS-VaR based on a 500-day moving window of historical returns.

Dangers of HS-*VaR*, Take II

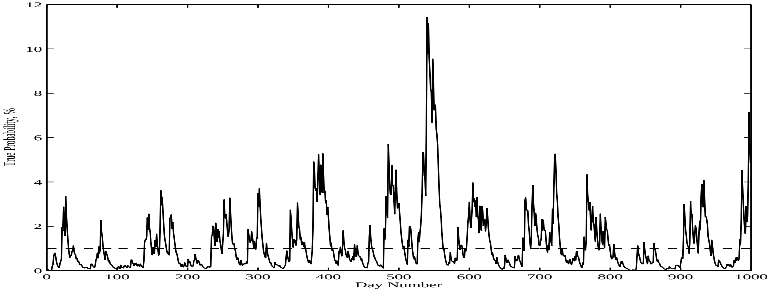


Figure: True Exceedance Probabilities of Nominal 1% HS-*VaR* When Volatility is Persistent. We simulate returns from a realistically-calibrated dynamic volatility model, after which we compute 1-day 1% HS-*VaR* using a rolling window of 500 observations. We plot the daily series of true conditional exceedance probabilities, which we infer from the model. For visual reference we include a horizontal line at the desired 1% probability level.

Dangers of HS-*VaR*, Take III

The unconditional HS-*VaR* perspective encourages incorrect rules of thumb, like scaling by $\sqrt{h}$ to convert 1-day VaR into h-day VaR.

Conditional VaR is Generally More Appropriate

Conditional VaR ($VaR^p_{T+1|T}$) solves:

$$p \; = \; Pr_T(r_{T+1} \le -VaR^p_{T+1|T}) \; = \; \int_{-\infty}^{-VaR^p_{T+1|T}} f_T(r_{T+1}) dr_{T+1}$$

($f_T(r_{T+1})$ is density of r_{T+1} *conditional* on time-T information)

But VaR of any Flavor has Issues

- *VaR* is silent regarding expected loss when *VaR* is exceeded (fails to assess the entire distributional tail)
- *VaR* fails to capture beneficial effects of portfolio diversification

Expected shortfall:

$$ES^{p}_{T+1|T} = p^{-1} \int_{0}^{p} VaR^{\gamma}_{T+1|T} \, d\gamma$$

- *ES* assesses the entire distributional tail
- *ES* captures the beneficial effects of portfolio diversification

Penn
University of Pennsylvania

Exponential Smoothing and RiskMetrics

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1-\lambda) r_{t-1}^2$$

$$\sigma_t^2 = \sum_{j=0}^{\infty} \varphi_j r_{t-1-j}^2$$

$$\varphi_j = (1-\lambda) \lambda^j$$

– Many initializations possible (r_1^2, sample variance, etc.)

$$\text{RM-VaR}_{T+1|T}^{p} \equiv \sigma_{T+1} \Phi_p^{-1}$$

(Assumes a particular volatility dynamic and conditional normality)

RiskMetrics Term Structure of Volatility

Conditional variance for the k-day return in RM is

$$Var(r_{t+k} + r_{t+k-1} + ... + r_{t+1} \,|\, F_t) \;\equiv\; \sigma^2_{t:t+k|t} \;=\; k\,\sigma^2_{t+1}$$

- Random walk for variance
- Random walk plus noise model for squared returns
- Optimal volatility forecast at any horizon is simply the current smoothed value
- Flat volatility term structure is not realistic!

Rigorous Modeling I

Conditional Portfolio-Level Volatility Dynamics from "Daily" Data

Conditional Return Distributions

$$f(r_t) \ vs. \ f(r_t|\Omega_{t-1})$$

Key 1: $E(r_t|\Omega_{t-1})$

Are returns conditional mean independent? Arguably yes.

Returns are (arguably) approximately serially uncorrelated, and (arguably) approximately free of additional non-linear conditional mean dependence.

Penn
University of Pennsylvania

Conditional Return Distributions, Continued

Key 2: $var(r_t|\Omega_{t-1}) \; = \; E((r_t - \mu)^2|\Omega_{t-1})$

Are returns conditional variance independent? No way!

Squared returns serially correlated, often with very slow decay.

The Standard Model
(Linearly Indeterministic Process with iid Innovations)

$$y_t = \sum_{i=0}^{\infty} b_i \, \varepsilon_{t-i}$$

$$\varepsilon \sim iid \, (0, \, \sigma_\varepsilon^2) \quad \sum_{i=0}^{\infty} b_i^2 < \infty \quad b_0 = 1$$

Uncond. mean: $E(y_t) = 0$ (constant)
Uncond. variance: $E(y_t - E(y_t))^2 = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} b_i^2$ (constant)
Cond. mean: $E(y_t \mid \Omega_{t-1}) = \sum_{i=1}^{\infty} b_i \, \varepsilon_{t-i}$ (varies)
Cond. variance: $E([y_t - E(y_t \mid \Omega_{t-1})]^2 \mid \Omega_{t-1}) = \sigma_\varepsilon^2$ (constant)

The Standard Model, Continued
k-Step-Ahead Least Squares Forecasting

$$E(y_{t+k} \mid \Omega_t) \; = \; \sum_{i=0}^{\infty} b_{k+i} \, \varepsilon_{t-i}$$

Associated prediction error:

$$y_{t+k} - E(y_{t+k} \mid \Omega_t) \; = \; \sum_{i=0}^{k-1} b_i \, \varepsilon_{t+k-i}$$

Conditional prediction error variance:

$$E([y_{t+k} - E(y_{t+k} \mid \Omega_t)]^2 \mid \Omega_t) \; = \; \sigma_\varepsilon^2 \sum_{i=0}^{k-1} b_i^2$$

Key: Depends only on k, not on Ω_t

ARCH(1) Process

$$r_t \mid \Omega_{t-1} \;\sim\; N(0,\; h_t)$$

$$h_t \;=\; \omega \;+\; \alpha \; r_{t-1}^2$$

$$E(r_t) \;=\; 0$$

$$E(r_t - E(r_t))^2 \;=\; \omega/(1-\alpha)$$

$$E(r_t \;\mid\; \Omega_{t-1}) \;=\; 0$$

$$E([r_t - E(r_t \mid \Omega_{t-1})]^2 \mid \Omega_{t-1}) \;=\; \omega \;+\; \alpha r_{t-1}^2$$

GARCH(1,1) Process

"Generalized ARCH"

$$r_t \mid \Omega_{t-1} \sim N(0,\ h_t)$$

$$h_t = \omega + \alpha\, r_{t-1}^2 + \beta\, h_{t-1}$$

$$E(r_t) = 0$$

$$E(r_t - E(r_t))^2 = \omega/(1-\alpha-\beta)$$

$$E(r_t \mid \Omega_{t-1}) = 0$$

$$E([r_t - E(r_t \mid \Omega_{t-1})]^2 \mid \Omega_{t-1}) = \omega + \alpha r_{t-1}^2 + \beta h_{t-1}$$

GARCH-Based 1-Day VaR

$$\text{GARCH-VaR}^{p}_{T+1|T} \equiv \sigma_{T+1}\, \Phi_p^{-1}$$

– Assumes Gaussian conditional distribution
– Consistent with fat tails of unconditional return distributions
– Can be extended to allow for fat-tailed conditional distributions

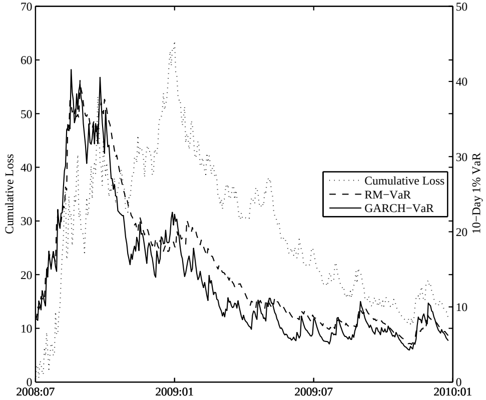


Figure: Cumulative S&P500 Loss (dots, left scale) and 1% 10-day RM-*VaR* and GARCH-*VaR* (solid and dashed, right scale), July 1, 2008 - December 31, 2009.

Unified Theoretical Framework

- Volatility dynamics (of course, by construction)
- Conditional symmetry translates into unconditional symmetry
- Volatility clustering produces unconditional leptokurtosis

Penn
University of Pennsylvania

Tractable Empirical Framework

$$L(\theta; r_1, \ldots, r_T) \approx f(r_T|\Omega_{T-1};\theta)\, f(r_{T-1}|\Omega_{T-2};\theta) \ldots f(r_{p+1}|\Omega_p;\theta)$$

If the conditional densities are Gaussian,

$$f(r_t|\Omega_{t-1};\theta) = \frac{1}{\sqrt{2\pi}} h_t(\theta)^{-1/2} \exp\left(-\frac{1}{2}\frac{r_t^2}{h_t(\theta)}\right)$$

$$\ln L(\theta; r_{p+1}, \ldots, r_T) \approx -\frac{T-p}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=p+1}^{T} \ln h_t(\theta) - \frac{1}{2}\sum_{t=p+1}^{T} \frac{r_t^2}{h_t(\theta)}$$

Penn
University of Pennsylvania

The Squared Return as a Noisy Volatility Proxy

Note that we can write:

$$r_t^2 = h_t + \nu_t$$

Thus r_t^2 is a *noisy* indicator of h_t

Various approaches handle the noise in various ways.

GARCH(1,1) and Exponential Smoothing

Exponential smoothing recursion:

$$\bar{r}_t^2 \;=\; \gamma\, r_t^2 \;+\; (1-\gamma)\,\bar{r}_{t-1}^2$$

Back substitution yields:

$$\bar{r}_t^2 \;=\; \sum w_j \; r_{t-j}^2$$

where $w_j \;=\; \gamma\,(1-\gamma)^j$

But in GARCH(1,1) we have:

$$h_t \;=\; \omega \;+\; \alpha r_{t-1}^2 \;+\; \beta h_{t-1}$$

$$h_t \;=\; \frac{\omega}{1-\beta} \;+\; \alpha \sum \beta^{j-1} r_{t-j}^2$$

Fat-Tailed Conditional Densities: t-GARCH

If r is conditionally Gaussian, then $\frac{r_t}{\sqrt{h_t}} \sim N(0,1)$

But often with high-frequency data, $\frac{r_t}{\sqrt{h_t}} \sim$ *fat tailed*

So take:

$$r_t = h_t^{1/2} z_t$$

$$z_t \overset{iid}{\sim} \frac{t_d}{std(t_d)}$$

Univariate GARCH(1,1) Variance Targeting

Sample unconditional variance:

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} r_t^2$$

Implied unconditional GARCH(1,1) variance:

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta}$$

We can constrain $\sigma^2 = \hat{\sigma}^2$ by constraining:

$$\omega = (1 - \alpha - \beta)\hat{\sigma}^2$$

– Saves a degree of freedom and ensures reasonableness

Additional GARCH material starts here...

Time Variation in Volatility and Prediction Error Variance

Conditional variance is a serially correlated random variable

$$h_t = \omega + \alpha \, r_{t-1}^2 + \beta \, h_{t-1}$$

Prediction error variance depends on Ω_{t-1}
e.g., 1-step-ahead prediction error variance is now

$$h_t = \omega + \alpha \, r_{t-1}^2 + \beta \, h_{t-1}$$

ARMA Representation in Squares

r_t^2 has the ARMA(1,1) representation:

$$r_t^2 = \omega + (\alpha + \beta) r_{t-1}^2 - \beta \nu_{t-1} + \nu_t,$$

where $\nu_t = r_t^2 - h_t$.

Variations on the GARCH Theme

- Regression with GARCH Disturbances
- Incorporating Exogenous Variables
- Asymmetric Response and the Leverage Effect:
- Fat-Tailed Conditional Densities
- Time-Varying Risk Premia

Regression with GARCH Disturbances

$$y_t = x_t' \beta + \varepsilon_t$$

$$\varepsilon_t | \Omega_{t-1} \;\sim\; N(0, h_t)$$

Incorporating Exogenous Variables

$$h_t = \omega + \alpha\, r_{t-1}^2 + \beta\, h_{t-1} + \gamma' z_t$$

γ is a parameter vector

z is a set of positive exogenous variables.

Asymmetric Response and the Leverage Effect I: TARCH

Standard GARCH: $h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1}$

TARCH: $h_t = \omega + \alpha r_{t-1}^2 + \gamma r_{t-1}^2 D_{t-1} + \beta h_{t-1}$

$$D_t = \begin{cases} 1 \ \textit{if} \ \ r_t < 0 \\ 0 \ \textit{otherwise} \end{cases}$$

positive return (good news): α effect on volatility

negative return (bad news): $\alpha + \gamma$ effect on volatility

$\gamma \neq 0$: Asymetric news response
$\gamma > 0$: "Leverage effect"

Asymmetric Response II: E-GARCH

$$\ln(h_t) \;=\; \omega \;+\; \alpha \left| \frac{r_{t-1}}{h_{t-1}^{1/2}} \right| \;+\; \gamma \frac{r_{t-1}}{h_{t-1}^{1/2}} \;+\; \beta \;\ln(h_{t-1})$$

- Log specification ensures that the conditional variance is positive.
- Volatility driven by both size and sign of shocks
- Leverage effect when $\gamma < 0$

Time-Varying Risk Premia: GARCH-M

Standard GARCH regression model:

$$y_t = x_t'\beta + \varepsilon_t$$

$$\varepsilon_t | \Omega_{t-1} \sim N(0, h_t)$$

GARCH-M model is a special case:

$$y_t = x_t'\beta + \gamma h_t + \varepsilon_t$$

$$\varepsilon_t | \Omega_{t-1} \sim N(0, h_t)$$

A GARCH(1,1) Example

Dependent Variable: R
Method: ML - ARCH (Marquardt)

Sample: 1 3461
Included observations: 3461
Convergence achieved after 19 iterations
Variance backcast: ON

Coefficient	Std. Error	z-Statistic	Prob.	
C	0.000640	0.000127	5.036942	0.0000
Variance Equation				
C	1.06E-06	1.49E-07	7.136840	0.0000
ARCH(1)	0.067410	0.004955	13.60315	0.0000
GARCH(1)	0.919714	0.006122	150.2195	0.0000

Figure: GARCH(1,1) Estimation, Daily NYSE Returns.

A GARCH(1,1) Example

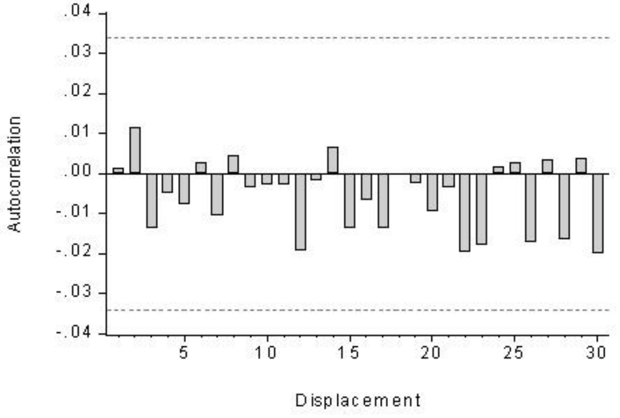


Figure: Correlogram of Squared Standardized GARCH(1,1) Residuals, Daily NYSE Returns.

A GARCH(1,1) Example

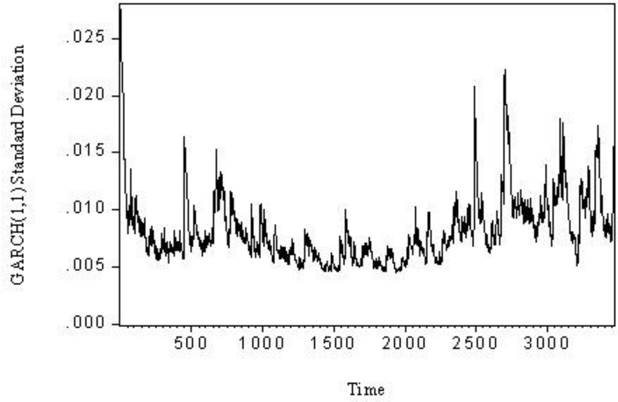


Figure: Estimated Conditional Standard Deviation, Daily NYSE Returns.

A GARCH(1,1) Example

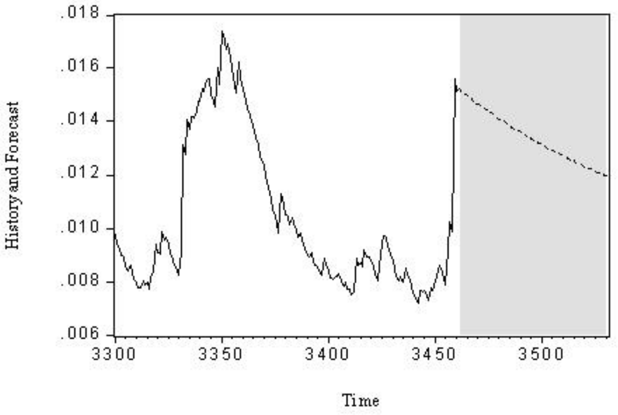


Figure: Conditional Standard Deviation, History and Forecast, Daily NYSE Returns.

After Exploring Lots of Possible Extensions...

Dependent Variable: R
Method: ML - ARCH (Marquardt) - Student's t distribution
Date: 04/10/12 Time: 13:48
Sample (adjusted): 2 3461
Included observations: 3460 after adjustments
Convergence achieved after 19 iterations
Presample variance: backcast (parameter = 0.7)
GARCH = C(4) + C(5)*RESID(-1)^2 + C(6)*RESID(-1)^2*(RESID(-1)<0)
+ C(7)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
@SQRT(GARCH)	0.083360	0.053138	1.568753	0.1167
C	1.28E-05	0.000372	0.034443	0.9725
R(-1)	0.073763	0.017611	4.188535	0.0000
Variance Equation				
C	1.03E-06	2.23E-07	4.628790	0.0000
RESID(-1)^2	0.014945	0.009765	1.530473	0.1259
RESID(-1)^2*(RESID(-1)<0)	0.094014	0.014945	6.290700	0.0000
GARCH(-1)	0.922745	0.009129	101.0741	0.0000
T-DIST. DOF	5.531579	0.478432	11.56188	0.0000

Additional GARCH material ends here...

Rigorous Modeling II

Conditional Portfolio-Level Volatility Dynamics from High-Frequency Data

Intraday Data and Realized Volatility

$$dp(t) = \mu(t)dt + \sigma(t)dW(t)$$

$$RV_t\left(\Delta\right) \equiv \sum_{j=1}^{N(\Delta)} \left(p_{t-1+j\Delta} - p_{t-1+(j-1)\Delta}\right)^2$$

$$RV_t\left(\Delta\right) \;\rightarrow\; IV_t = \int_{t-1}^{t} \sigma^2\left(\tau\right) d\tau$$

Microstructure Noise

– State space signal extraction

– AvgRV

– Realized kernel

– Many others

– Interestingly, MSN is rapidly becoming less important as spreads narrow (but they may widen again in crises)

Penn
UNIVERSITY of PENNSYLVANIA

RV is Persistent

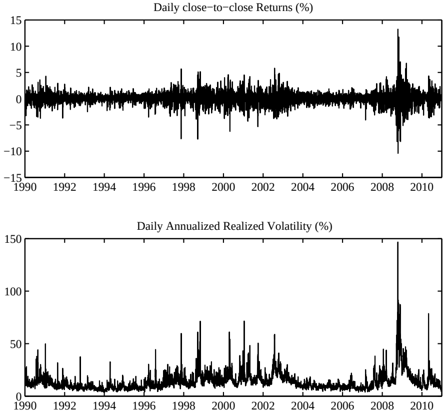


Figure: S&P500 Daily Returns and Volatilities (Percent). The top panel shows daily S&P500 returns, and the bottom panel shows daily S&P500 realized volatility. We compute realized volatility as the square root of *AvgRV*, where *AvgRV* is the average of five daily RVs each computed from 5-minute squared returns on a 1-minute grid of S&P500 futures prices.

RV is Log-Normal

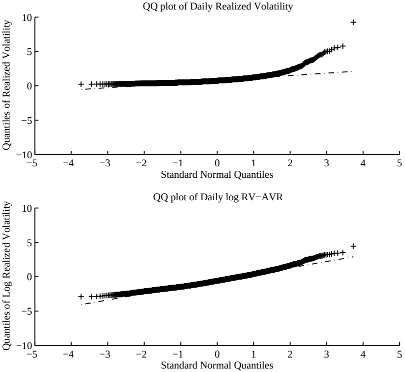


Figure: S&P500: QQ Plots for Realized Volatility and Log Realized Volatility. The top panel plots the quantiles of daily realized volatility against the corresponding normal quantiles. The bottom panel plots the quantiles of the natural logarithm of daily realized volatility against the corresponding normal quantiles. We compute realized volatility as the square root of *AvgRV*, where *AvgRV* is the average of five daily RVs each computed from 5-minute squared returns on a 1-minute grid of S&P500 futures prices.

Crucially: RV is Long-Memory

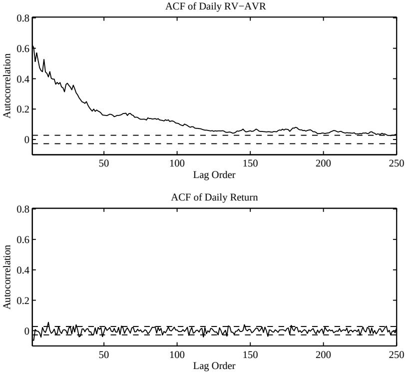


Figure: S&P500: Sample Autocorrelations of Daily Realized Variance and Daily Return. The top panel shows realized variance autocorrelations, and the bottom panel shows return autocorrelations, for displacements from 1 through 250 days. Horizontal lines denote 95% Bartlett bands. Realized variance is *AvgRV*, the average of five daily RVs each computed from 5-minute squared returns on a 1-minute grid of S&P500 futures prices.

Penn
UNIVERSITY of PENNSYLVANIA

Exact and Approximate Long Memory

Exact long memory:

$$(1-L)^d RV_t = \beta_0 + \nu_t$$

Corsi (2009) approximation (HAR):

$$RV_t \;=\; \beta_0 \,+\, \beta_1\, RV_{t-1} \,+\, \beta_2\, RV_{t-5:t-1} \,+\, \beta_3 RV_{t-21:t-1} \,+\, \nu_t$$

Even better:

$$\log RV_t \;=\; \beta_0 + \beta_1 \log RV_{t-1} + \beta_2 \log RV_{t-5:t-1} + \beta_3 \log RV_{t-21:t-1} + \nu_t$$

– Ensures positivity and promotes normality

RV-VaR

$$RV - VaR^{p}_{T+1|T} \; = \; \widehat{RV}_{T+1|T}\,\Phi^{-1}_{p},$$

where Φ^{-1}_{p} is the relevant standard normal quantile

Easily adapted for multi-period variance forecasts:

$$RV_{t:t+9} \; = \; \beta_0 + \beta_1\,RV_{t-1} + \beta_2\,RV_{t-5:t-1} + \beta_3\,RV_{t-21:t-1} + \nu_{t:t+9}$$

$$RV - VaR^{p}_{T+10|T} \; = \; \widehat{RV}_{T+1:T+10|T}\,\Phi^{-1}_{p}$$

e.g., HAR-VaR vs. GARCH-VaR

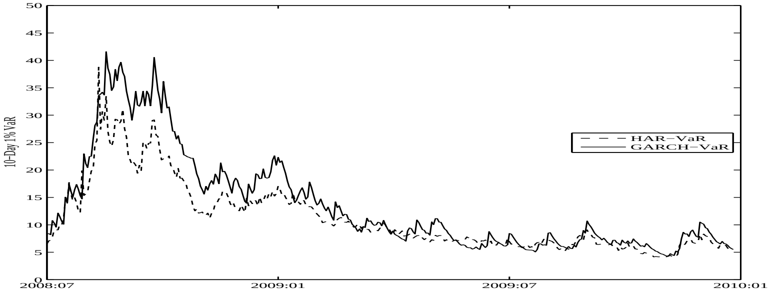


Figure: 10-day 1% HAR-*VaR* and GARCH-*VaR*, July 1, 2008 - December 31, 2009. The dashed line shows 10-day 1% HAR-*VaR* based on the HAR forecasting model for 10-day realized volatility. The solid line shows 10-day 1% GARCH-*VaR*. When computing *VaR* the 10-day returns divided by the expected volatility are assumed to be normally distributed.

GARCH-RV

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \gamma RV_{t-1}$$

- Fine for 1-step
- Multi-step requires "closing the system" with an RV equation
  - Hansen, Huang and Shek (2010), "Realized GARCH"
  - Shephard and Sheppard (2010), "HEAVY"

Separating Jumps

$$QV_t \;=\; IV_t \,+\, JV_t$$

where

$$JV_t \;=\; \sum_{j=1}^{\mathcal{J}_t} J_{t,j}^2$$

e.g., we might want to explore:

$$\begin{aligned} RV_t \;&= \beta_0 + \beta_1\, IV_{t-1} + \beta_2\, IV_{t-5:t-1} + \beta_3\, IV_{t-21:t-1} \\ &\quad + \alpha_1\, JV_{t-1} + \alpha_2\, JV_{t-5:t-1} + \alpha_3\, JV_{t-21:t-1} + \nu_t \end{aligned}$$

Penn
University of Pennsylvania

But How to Separate Jumps?

- Mancini (2001) Truncation:

$$TV_t(\Delta) = \sum_{j=1}^{N(\Delta)} \Delta p_{t-1+j\Delta}^2 \, I\left(\Delta p_{t-1+j\Delta} < \mathcal{T}\right)$$

- Barndorff-Nielsen and Shephard (2004) Bi-Power:

$$BPV_t(\Delta) = \frac{\pi}{2} \frac{N(\Delta)}{N(\Delta)-1} \sum_{j=1}^{N(\Delta)-1} \left|\Delta p_{t-1+j\Delta}\right| \left|\Delta p_{t-1+(j+1)\Delta}\right|$$

- Andersen, Dobrev and Schaumburg (2010) Minimum:

$$MinRV_t(\Delta) = \frac{\pi}{\pi-2} \left(\frac{N(\Delta)}{N(\Delta)-1}\right) \sum_{j=1}^{N(\Delta)-1} \min\left\{\left|\Delta p_{t-1+j\Delta}\right|, \left|\Delta p_{t-1+(j+1)\Delta}\right|\right\}^2$$

Modeling Entire Return Distributions: Returns are not Unconditionally Gaussian

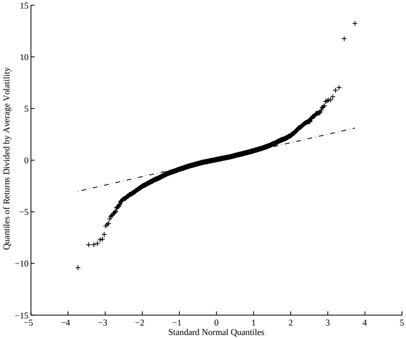


Figure: QQ Plot of S&P500 Returns. We show quantiles of daily S&P500 returns from January 2, 1990 to December 31, 2010, against the corresponding quantiles from a standard normal distribution.

Modeling Entire Return Distributions: Returns are Often not Conditionally Gaussian

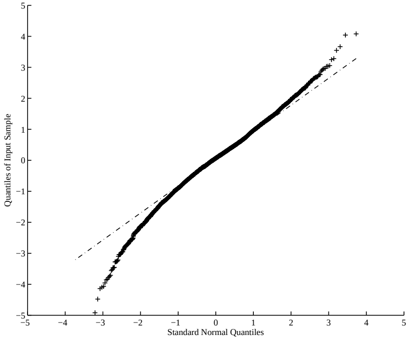


Figure: QQ Plot of S&P500 Returns Standardized by NGARCH Volatilities. We show quantiles of daily S&P500 returns standardized by the dynamic volatility from a NGARCH model against the corresponding quantiles of a standard normal distribution. The sample period is January 2, 1990 through December 31, 2010. The units on each axis are standard deviations.

Modeling Entire Return Distributions: Issues

- Gaussian QQ plots effectively show calibration of Gaussian VaR at different levels
- Gaussian unconditional VaR is terrible
- Gaussian conditional VaR is somewhat better but left tail remains bad
- Gaussian conditional expected shortfall, which integrates over the left tail, would be terrible
- So we want more accurate assessment of things like $VaR^{p}_{T+1|T}$ than those obtained under Gaussian assumptions
  –Doing so for all values of $p \in [0, 1]$ requires estimating the entire conditional return distribution
  – More generally, best-practice risk measurement is about tracking the entire conditional return distribution

Penn
University of Pennsylvania

Observation-Driven Density Forecasting Using $r = \sigma\,\varepsilon$ and GARCH

Assume:

$$r_{T+1} = \sigma_{T+1/T}\varepsilon_{T+1}$$

$$\varepsilon_{T+1} \sim iid(0,1)$$

Multiply ε_{T+1} draws by $\sigma_{T+1/T}$ (fixed across draws, from a GARCH model) to build up the conditional density of r_{T+1}.

- ε_{T+1} simulated from standard normal
- ε_{T+1} simulated from standard t
- ε_{T+1} simulated from kernel density fit to $\frac{r_{T+1}}{\sigma_{T+1/T}}$
- ε_{T+1} simulated from any density that can be simulated

Penn
UNIVERSITY of PENNSYLVANIA

Parameter-Driven Density Forecasting Using $r = \sigma \varepsilon$ and SV

Assume:

$$r_{T+1} = \sigma_{T+1} \varepsilon_{T+1}$$

$$\varepsilon_{T+1} \sim iid(0, 1)$$

Multiply ε_{T+1} draws by σ_{T+1} draws (from a simulated SV model) to build up the conditional density of r_{T+1}.

– Again, ε_{T+1} simulated from any density deemed relevant

Modeling Entire Return Distributions: Returns Standardized by RV *are* Approximately Gaussian

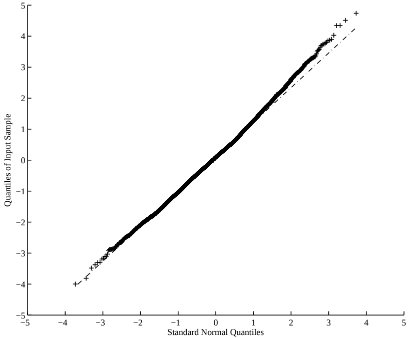


Figure: QQ Plot of S&P500 Returns Standardized by Realized Volatilities. We show quantiles of daily S&P500 returns standardized by *AvgRV* against the corresponding quantiles of a standard normal distribution. The sample period is January 2, 1990 through December 31, 2010. The units on each axis are standard deviations.

A Special Parameter-Driven Density Forecasting Approach Using $r = \sigma\,\varepsilon$ and RV (Log-Normal / Normal Mixture)

Assume:

$$r_{T+1} = \sigma_{T+1}\varepsilon_{T+1}$$

$$\varepsilon_{T+1} \sim iid(0,1)$$

Multiply ε_{T+1} draws from $N(0,1)$ by σ_{T+1} draws (from a simulated RV model fit to log realized standard deviation) to build up the conditional density of r_{T+1}.

Pitfalls of the "$r = \sigma\,\varepsilon$" Approach

In the conditionally *Gaussian* case we can write with no loss of generality:

$$r_{T+1} = \sigma_{T+1/T}\varepsilon_{T+1}$$
$$\varepsilon_{T+1} \sim iidN(0,1)$$

But in the conditionally non-Gaussian case there *is* potential loss of generality in writing:

$$r_{T+1} = \sigma_{T+1/T}\varepsilon_{T+1}$$
$$\varepsilon_{T+1} \sim iid(0,1),$$

because there may be time variation in conditional moments other than $\sigma_{T+1/T}$, and using $\varepsilon_{T+1} \sim iid(0,1)$ assumes that away

Rigorous Modeling III

Conditional Asset-Level (Multivariate) Volatility Dynamics from "Daily" Data

Multivariate

Univariate volatility models useful for portfolio-level risk measurement (VaR, ES, etc.)

But what about risk *management* questions:

- Portfolio risk change under a certain scenario involving price movements of set of assets or asset classes?
- Portfolio risk change if certain correlations increase suddenly
- Portfolio risk change if I double my holdings of Intel?
- How do optimal portfolio shares change if the covariance matrix moves in a certain way?

Similarly, what about almost any other question in asset pricing, hedging, trading? Almost all involve correlation.

Penn
University of Pennsylvania

Basic Framework and Issues I

$$N \times 1 \text{ return vector } R_t$$

$$N \times N \text{ covariance matrix } \Omega_t$$

- $\frac{N(N+1)}{2}$ distinct elements
- Structure needed for pd or even psd
- Huge number of parameters even for moderate N
- And N may be *not* be moderate!

Basic Framework and Issues II

Univariate:

$$r_t = \sigma_t z_t$$

$$z_t \sim i.i.d.(0, 1)$$

Multivariate:

$$R_t = \Omega_t^{1/2} Z_t$$

$$Z_t \sim i.i.d.(\underline{0}, \mathcal{I})$$

where $\Omega_t^{1/2}$ is a "square-root" (e.g., Cholesky factor) of Ω_t

Ad Hoc Exponential Smoothing (RM)

$$\Omega_t \;=\; \lambda\,\Omega_{t-1} \,+\, (1-\lambda)\,R_{t-1}R_{t-1}'$$

- Assumes that the dynamics of all the variances and covariances are driven by a single scalar parameter λ (identical smoothness)
- Guarantees that the smoothed covariance matrices are pd so long as Ω_0 is pd
- Common strategy is to set Ω_0 equal to the sample covariance matrix $\frac{1}{T}\sum_{t=1}^{T} R_t R_t'$ (which is pd if $T > N$)
- But covariance matrix forecasts inherit the implausible scaling properties of the univariate RM forecasts and will in general be suboptimal

Multivariate GARCH(1,1)

$$vech\,(\Omega_t) \;=\; vech\,(C) + B\, vech\,(\Omega_{t-1}) + A\, vech\,(R_{t-1}R'_{t-1})$$

- *vech* operator converts the upper triangle of a symmetric matrix into a $\frac{1}{2}N(N+1) \times 1$ column vector
- A and B matrices are both of dimension $\frac{1}{2}N(N+1) \times \frac{1}{2}N(N+1)$
- Even in this "parsimonious" GARCH(1,1) there are $O(N^4)$ parameters
  – More than 50 *million* parameters for $N = 100$!

Encouraging Parsimony: Diagonal GARCH(1,1)

Diagonal GARCH constrains A and B matrices to be diagonal.

$$vech(\Omega_t) = vech(C) + (I\underline{\beta})\, vech(\Omega_{t-1}) + (I\underline{\alpha})\, vech(R_{t-1}R'_{t-1})$$

– Still $O(N^2)$ parameters.

Encouraging Parsimony: Scalar GARCH(1,1)

Scalar GARCH constrains A and B matrices to be scalar:

$$vech(\Omega_t) \;=\; vech(C) + (I\beta)\, vech(\Omega_{t-1}) + (I\alpha)\, vech(R_{t-1}R'_{t-1})$$

– Mirrors RM, but with the important difference that the Ω_t forecasts now revert to $\Omega = (1 - \alpha - \beta)^{-1} C$

– Fewer parameters than diagonal, but still $O(N)^2$ (because of C)

Encouraging Parsimony: Covariance Targeting

Recall variance targeting:

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} r_t^2, \quad \sigma^2 = \frac{\omega}{1-\alpha-\beta} \implies \text{take } \omega = (1-\alpha-\beta)\hat{\sigma}^2$$

Covariance targeting is the obvious multivariate generalization:

$$vech(C) \;=\; (\mathcal{I} - A - B)\; vech\left(\frac{1}{T}\sum_{t=1}^{T} R_t R_t'\right)$$

– Encourages both parsimony and reasonableness

Constant Conditional Correlation (CCC) Model

[Key is to recognize that correlation matrix
is the covariance matrix of standardized returns]

Two-step estimation:

- Estimate N appropriate univariate GARCH models
- Calculate standardized return vector, $\hat{e}_t = R_t \hat{D}_t^{-1}$
- Estimate correlation matrix Γ (assumed constant) as $\frac{1}{T}\sum_{t=1}^{T} \hat{e}_t \hat{e}_t'$
  - Quite flexible as the N models can differ across returns

Dynamic Conditional Correlation (DCC) Model

Two-step estimation:

- Estimate N appropriate univariate GARCH models
- Calculate standardized return vector, $\hat{e}_t = R_t \hat{D}_t^{-1}$
- Estimate correlation matrix Γ_t (assumed to have scalar GARCH(1,1)-style dynamics) as following

$$vech(\Gamma_t) = vech(C) + (I\beta)vech(\Gamma_{t-1}) + (I\alpha)vech(e_{t-1}e'_{t-1})$$

– “Correlation targeting” is helpful

DECO

- Time-varying correlations assumed identical across all pairs of assets, which implies:

$$\Gamma_t \; = \; (1 - \rho_t)\mathcal{I} \; + \; \rho_t \, \mathcal{J},$$

  where $\mathcal{J}$ is an $N \times N$ matrix of ones
- Analytical inverse facilitates estimation:

$$\Gamma_t^{-1} \; = \; \frac{1}{(1 - \rho_t)} \left[\mathcal{I} \; - \; \frac{\rho_t}{1 + (N-1)\rho_t} \, \mathcal{J} \right]$$

- Assume GARCH(1,1)-style conditional correlation structure:

$$\rho_t \; = \; \omega_\rho \; + \; \alpha_\rho \, u_t \; + \; \beta_\rho \, \rho_{t-1}$$

- Updating rule is naturally given by the average conditional correlation of the standardized returns,

$$u_t \; = \; \frac{2 \sum_{i=1}^{N} \sum_{j>i}^{N} \; e_{i,t} \, e_{j,t}}{N \sum_{i=1}^{N} \; e_{i,t}^2}$$

- Three parameters, ω_ρ, α_ρ and β_ρ, to be estimated.

DECO Example

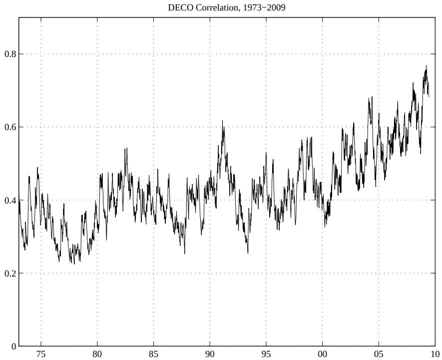


Figure: Time-Varying International Equity Correlations. The figure shows the estimated equicorrelations from a DECO model for the aggregate equity index returns for 16 different developed markets from 1973 through 2009.

Factor Structure

$$R_t = \lambda F_t + \nu_t$$

where

$$F_t = \Omega_{Ft}^{1/2} Z_t$$

$$Z_t \sim i.i.d.(0,\ \mathcal{I})$$

$$\nu_t \sim i.i.d.(0,\ \Omega_\nu)$$

$$\Longrightarrow\ \Omega_t\ =\ \lambda\,\Omega_{Ft}\,\lambda'\ +\ \Omega_{\nu t}$$

Penn
UNIVERSITY of PENNSYLVANIA

One-Factor Case with Everything Orthogonal

$$R_t = \lambda f_t + \nu_t$$

where

$$f_t = \sigma_{ft} z_t$$

$$z_t \sim i.i.d.(0,\ 1)$$

$$\nu_t \sim i.i.d.(0,\ \sigma_\nu^2)$$

$$\implies \Omega_t \ = \ \sigma_{ft}^2 \lambda \lambda' \ + \ \sigma_\nu^2$$

$$\sigma_{it}^2 = \sigma_{ft}^2 \lambda_i^2 + \sigma_\nu^2$$

$$\sigma_{ijt}^2 = \sigma_{ft}^2 \lambda_i \lambda_j$$

Rigorous Modeling IV

Conditional Asset-Level (Multivariate) Volatility Dynamics from High-Frequency Data

Realized Covariance

$$dP(t) \; = \; M(t)\, dt \, + \, \Omega(t)^{1/2}\, dW(t)$$

$$RCov_t(\Delta) \; \equiv \; \sum_{j=1}^{N(\Delta)} R_{t-1+j\Delta,\Delta}\, R'_{t-1+j\Delta,\Delta}$$

$$RCov_t(\Delta) \; \rightarrow \; ICov_t \; = \; \int_{t-1}^{t} \Omega\,(\tau)\, d\tau$$

– p.d. so long as $N(\Delta) > N$; else use regularization methods

Asynchronous Trading and the Epps Effect

– Epps effect biases covariance estimates downward

– Can overcome Epps by lowering sampling frequency to accommodate least-frequently-traded asset, but that wastes data

– Opposite extreme: Calculate each pairwise realized covariance matrix using appropriate sampling; then assemble and regularize

Regularization(Shrinkage)

$$\hat{\Omega}_t^S \;=\; \kappa\, RCov_t(\Delta) \;+\; (1-\kappa)\,\Upsilon_t$$

– Υ_t is p.d. and $0 < \kappa < 1$

– $\Upsilon_t = I$ (naive benchmark)
– $\Upsilon_t = \Omega$ (unconditional covariance matrix)
– $\Upsilon_t \;=\; \sigma_M^2\, b\, b' \;+\; D_\nu$ (one-factor market model)

Models for Realized Covariance Matrices

– Mimic univariate exponential smoothing:

$$\hat{\Omega}_{t+1|t} = \lambda \hat{\Omega}_{t|t-1} + (1-\lambda) \hat{\Omega}_t$$

– Mimic scalar diagonal GARCH with multivariate regression:

$$vech(\hat{\Omega}_{t+1}) = vech(C) + \beta \, vech(\hat{\Omega}_t) + \xi_{t+1}$$

– Mimic DCC:

$$vech(Q_t) = vech(C) + \beta \, vech(Q_{t-1}) + \xi_t$$

– Maintain p.d. using Cholesky or matrix log

Multivariate GARCH-RV

$$vech\,(\Omega_t) \;=\; vech\,(C) \,+\, B\,vech\,(\Omega_{t-1}) \,+\, A\,vech(\hat{\Omega}_{t-1})$$

- Fine for 1-step
- Multi-step requires “closing the system” with an RV equation

  – Noureldin et al. (2011), multivariate HEAVY

Multivariate Return Distributions

– If reliable realized covariances are available, one could do a multivariate analog of the earlier lognormal/normal mixture model. But the literature thus far has focused primarily on conditional distributions for "daily" data.

Return version:

$$Z_t \;=\; \Omega_t^{-1/2}\, R_t, \qquad Z_t \sim i.i.d., \quad E_{t-1}(Z_t) = 0 \quad Var_{t-1}(Z_t) = \mathcal{I}$$

Standardized return version (as in DCC):

$$e_t \;=\; D_t^{-1}\, R_t, \quad E_{t-1}(e_t) \,=\, 0, \qquad Var_{t-1}(e_t) \,=\, \Gamma_t$$

where D_t denotes the diagonal matrix of conditional standard deviations for each of the assets, and Γ_t refers to the potentially time-varying conditional correlation matrix.

Leading Examples

Multivariate normal:

$$f(e_t) \;=\; C\left(\Gamma_t\right) \exp\left(-\tfrac{1}{2}\, e_t'\, \Gamma_t^{-1}\, e_t\right)$$

Multivariate t:

$$f(e_t) \;=\; C\left(d, \Gamma_t\right) \left(1 + \frac{e_t' \Gamma_t^{-1} e_t}{(d-2)}\right)^{-(d+N)/2}$$

Multivariate asymmetric t:

$$f\left(e_t\right) = \frac{C\left(d, \dot{\Gamma}_t\right) K_{\frac{d+N}{2}}\left(\sqrt{\left(d + (e_t - \dot{\mu})'\, \dot{\Gamma}_t^{-1}\,(e_t - \dot{\mu})\right) \xi' \dot{\Gamma}_t^{-1} \xi}\right) \exp\left((e_t - \dot{\mu})'\, \dot{\Gamma}_t^{-1} \xi\right)}{\left(1 + \frac{(e_t - \dot{\mu})' \dot{\Gamma}_t^{-1} (e_t - \dot{\mu})}{d}\right)^{\frac{(d+N)}{2}} \left(\sqrt{\left(d + (e_t - \dot{\mu})'\, \dot{\Gamma}_t^{-1}\,(e_t - \dot{\mu})\right) \xi'\, \dot{\Gamma}_t^{-1} \xi}\right)^{-\frac{(d+N)}{2}}}$$

– More flexible than symmetric t but requires estimation of N asymmetry parameters simultaneously with the other parameters, which is challenging in high dimensions.
Copula methods sometimes provide a simpler two-step approach.

Copula Methods

Sklar's Theorem:

$$F(e) \; = \; G(\, F_1(e_1), ..., F_N(e_N)\,) \; \equiv \; G(\, u_1, ..., u_N \,) \; \equiv \; G(u)$$

$$f(e) \; = \; \frac{\partial^N G(F_1(e_1), ..., F_N(e_N))}{\partial e_1 ... \partial e_N} \; = \; g\,(u) \times \prod_{i=1}^{N} f_i(e_i)$$

$$\Longrightarrow \quad \log L \; = \; \sum_{t=1}^{T} \log g(u_t) \; + \; \sum_{t=1}^{T} \sum_{i=1}^{N} \log f_i(e_{i,t})$$

Penn
University of Pennsylvania

Standard Copulas

Normal:

$$g(u_t; \Gamma_t^*) = |\Gamma_t^*|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\Phi^{-1}(u_t)'(\Gamma_t^{*-1} - I)\Phi^{-1}(u_t)\right\}$$

where $\Phi^{-1}(u_t)$ refers to the $N \times 1$ vector of standard inverse univariate normals, and the correlation matrix Γ_t^* pertains to the $N \times 1$ vector e_t^* with typical element,

$$e_{i,t}^* \;=\; \Phi^{-1}(u_{i,t}) \;=\; \Phi^{-1}(F_i\,(e_{i,t})).$$

– Often does not allow for sufficient dependence between tail events.

– t copula

– Asymmetric t copula

Multivariate Distribution Simulation (General Case)

Simulate using:

$$R_t = \hat{\Omega}_t^{1/2} Z_t$$

$$Z_t \sim i.i.d.(\underline{0}, \mathcal{I})$$

– Z_t may be drawn from parametrically-(Gaussian, t, ...) or nonparametrically-fitted distributions, or with replacement from the empirical distribution.

Multivariate Distribution Simulation (Factor Case)

Simulate using:

$$R_{F,t} = \hat{\Omega}_{F,t}^{1/2} Z_{F,t}$$

$$R_t \ = \ \hat{B}\, R_{F,t} + \nu_t$$

– $Z_{F,t}$ and ν_t may be drawn from parametrically- or nonparametrically-fitted distributions, or with replacement from the empirical distribution.

Asymmetric Tail Correlations

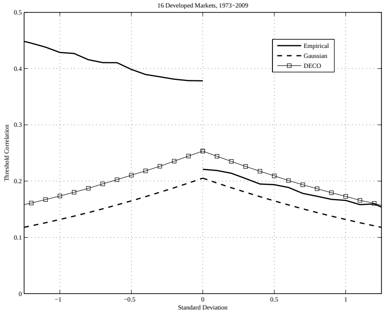


Figure: Average Threshold Correlations for Sixteen Developed Equity Markets. The solid line shows the average empirical threshold correlation for GARCH residuals across sixteen developed equity markets. The dashed line shows the threshold correlations implied by a multivariate standard normal distribution with constant correlation. The line with square markers shows the threshold correlations from a DECO model estimated on the GARCH residuals from the 16 equity markets. The figure is based on weekly returns from 1973 to 2009.

Rigorous Modeling V

Measuring Connectedness
in the Small and in the Large

Financial and Economic Connectedness (Diebold-Yilmaz)

- Market Risk, Portfolio Concentration Risk
  (return connectedness)
- Credit Risk
  (default connectedness)
- Counterparty Risk, Gridlock Risk
  (bilateral and multilateral contractual connectedness)
- Systemic Risk
  (system-wide connectedness)
- Business Cycle Risk
  (local or global real output connectedness)

Covariance and Correlation

- So pairwise...
- So linear...
- So Gaussian...

We Will Take a Different Approach...

– Natural framework with direct motivation

– Firmly grounded in network theory

– Track transmissions and receipts, from highly granular to highly aggregated

Two Natural Questions

A natural modeling question:
What fraction of the H-step-ahead prediction-error variance of variable i is due to shocks in variable $j, \forall i, j$?

Variance decomposition: $d_{ij}^H, \forall i, j$

A natural financial/economic connectedness question:
What fraction of the H-step-ahead prediction-error variance of variable i is due to shocks in variable $j, \forall j \neq i$?

Non-own elements of the variance decomposition: $d_{ij}^H, \forall j \neq i$

Variance Decompositions and the Connectedness Table

N-Variable Connectedness Table

	x_1	x_2	...	x_N	From Others to i
x_1	d^H_{11}	d^H_{12}	$\cdots$	d^H_{1N}	$\Sigma^N_{j=1} d^H_{1j}, j \neq 1$
x_2	d^H_{21}	d^H_{22}	$\cdots$	d^H_{2N}	$\Sigma^N_{j=1} d^H_{2j}, j \neq 2$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_N	d^H_{N1}	d^H_{N2}	$\cdots$	d^H_{NN}	$\Sigma^N_{j=1} d^H_{Nj}, j \neq N$
To Others From j	$\Sigma^N_{i=1} d^H_{i1}$ $i \neq 1$	$\Sigma^N_{i=1} d^H_{i2}$ $i \neq 2$	$\cdots$	$\Sigma^N_{i=1} d^H_{iN}$ $i \neq N$	$\Sigma^N_{i,j=1} d^H_{ij}$ $i \neq j$

Upper-left block is variance decomposition matrix, D

Connectedness involves the **non-diagonal** elements of D

Penn
UNIVERSITY of PENNSYLVANIA

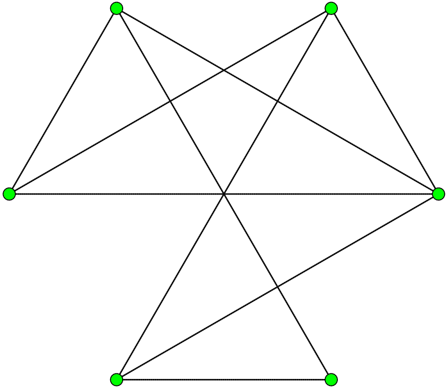
Connectedness Measures

- Pairwise Directional: $C^H_{i\leftarrow j} = d^H_{ij}$ ("i's imports from j")
- Net: $C^H_{ij} = C^H_{j\leftarrow i} - C^H_{i\leftarrow j}$ ("ij bilateral trade balance")

- Total Directional:
  - From others to i: $C^H_{i\leftarrow \bullet} = \sum_{\substack{j=1 \\ j\neq i}}^{N} d^H_{ij}$ ("i's total imports")
  - To others from j: $C^H_{\bullet\leftarrow j} = \sum_{\substack{i=1 \\ i\neq j}}^{N} d^H_{ij}$ ("j's total exports")
- Net: $C^H_{i} = C^H_{\bullet\leftarrow i} - C^H_{i\leftarrow \bullet}$ ("i's multilateral trade balance")

- Total: $C^H = \frac{1}{N}\sum_{\substack{i,j=1 \\ i\neq j}}^{N} d^H_{ij}$ ("total world exports")

Networks I: Representation



$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

Adjacency Matrix (Symmetric)
$A_{ij} = 1$ if nodes i, j linked
$A_{ij} = 0$ otherwise

Networks I: Degree

Degree of node i, d_i:

$$d_i = \sum_{j=1}^{N} A_{ij}$$

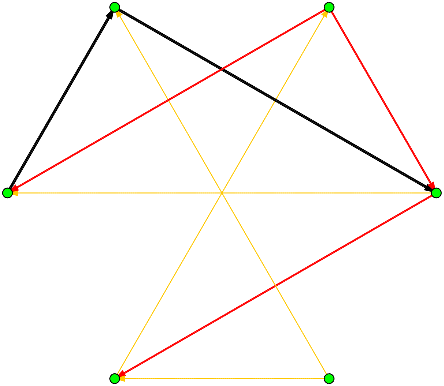
Discrete *degree distribution*, $P(d)$, on 0, ..., $N-1$

Mean degree, $E(d)$, is the key connectedness measure

Beautiful results (e.g., "small world") involve the mean degree:

$$\textit{diameter} \approx \frac{\ln N}{\ln E(d)}$$

Networks II : Representation (Weighted, Directed)



$$A = \begin{pmatrix} 0 & .5 & .7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .3 & 0 \\ 0 & 0 & 0 & .7 & 0 & .3 \\ .3 & .5 & 0 & 0 & 0 & 0 \\ .5 & 0 & 0 & 0 & 0 & .3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

"to i, from j"

Networks II: Degree (Weighted, Directed)

$A_{ij} \in [0, 1]$ depending on connection strength

Two degrees:

$$d_i^{from} = \sum_{j=1}^{N} A_{ij}$$

$$d_j^{to} = \sum_{i=1}^{N} A_{ij}$$

Continuous "from" and "to" degree distributions on $[0, N-1]$

Mean degree $E(d)$ remains the key connectedness measure

Central Observation: D is a Weighted, Directed Network

Connectedness Table

	x_1	x_2	...	x_N	From Others
x_1	d^H_{11}	d^H_{12}	$\cdots$	d^H_{1N}	$\sum_{j\neq 1} d^H_{1j}$
x_2	d^H_{21}	d^H_{22}	$\cdots$	d^H_{2N}	$\sum_{j\neq 2} d^H_{2j}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_N	d^H_{N1}	d^H_{N2}	$\cdots$	d^H_{NN}	$\sum_{j\neq N} d^H_{Nj}$
To Others	$\sum_{i\neq 1} d^H_{i1}$	$\sum_{i\neq 2} d^H_{i2}$	$\cdots$	$\sum_{i\neq N} d^H_{iN}$	$\sum_{i\neq j} d^H_{ij}$

$$C^H_{i\leftarrow\bullet} = \sum_{\substack{j=1\\ j\neq i}}^{N} d^H_{ij}, \text{ are the “from degrees”}$$

$$C^H_{\bullet\leftarrow j} = \sum_{\substack{i=1\\ i\neq j}}^{N} d^H_{ij}, \text{ are the “to degrees”}$$

$$C^H = \frac{1}{N}\sum_{\substack{i,j=1\\ i\neq j}}^{N} d^H_{ij}, \text{ is the mean degree (to or from)}$$

Relationships to other market-based measures?

- Marginal expected shortfall
- Expected capital shortfall
- CoVaR
- ΔCoVaR

MES and ECS

$$MES^{j|mkt}_{T+1|T} = E_T \left[r_{j,T+1} | \mathbb{C} \left(r_{mkt,T+1} \right) \right]$$

- Sensitivity of firm j's return to extreme market event $\mathbb{C}$
- Market-based "stress test" of firm j's fragility
- Like "total directional connectedness *from*" (from degree)

$$ECS^{j|mkt}_{T+1|T} = a_{0j} + a_{1j} MES^{j|mkt}_{T+1|T}$$

– a_{0j} depends on firm j's "prudential ratio" of asset value to equity as well as its debt composition
– a_{1j} depends on firm j's prudential ratio and initial capital

CoVaR and ΔCoVaR

$$\text{VaR}: \; p = \Pr_T\left(r_{j,T+1} < -VaR^{p,j}_{T+1|T}\right)$$

$$\text{CoVaR}: \; p = \Pr_T\left(r_{j,T+1} < -CoVaR^{j|i}_{T+1|T} \mid \mathbb{C}\left(r_{i,T+1}\right)\right)$$

$$\text{Marketwide}: \; p = \Pr_T\left(r_{mkt,T+1} < -CoVaR^{mkt|i}_{T+1|T} \mid \mathbb{C}\left(r_{i,T+1}\right)\right)$$

- Measures tail-event linkages
- Leading choice of $\mathbb{C}\left(r_{i,T+1}\right)$ is that firm i breaches its VaR
- Like "total directional connectedness *to*" (to degree)

$$\text{Delta CoVaR}: \; \Delta CoVaR^{j|i}_{T+1|T} = CoVaR^{j|VaR(i)}_{T+1|T} - CoVaR^{j|Med(i)}_{T+1|T}$$

$$\text{Marketwide}: \; \Delta CoVaR^{mkt|i}_{T+1|T} = CoVaR^{mkt|VaR(i)}_{T+1|T} - CoVaR^{mkt|Med(i)}_{T+1|T}$$

Penn
University of Pennsylvania

Estimating Connectedness

Thus far we've worked under correct specification, in population:

$$C(x, H, B(L))$$

Now we want:

$$\widehat{C}\left(x, H, B(L), M(L; \hat{\theta})\right),$$

and similarly for other variants of connectedness

Many Interesting Issues

- x objects: Returns? **Return volatilities**? Real activities?
- x universe: How many and which ones?
  ($\approx$ **15 major financial institutions**)
- x frequency: **Daily?** Monthly? Quarterly?
- H: **Match VaR horizon**? Holding period?
- M: **VAR**? Structural?
- Identification of variance decompositions:
  Cholesky? **Generalized**? Structural?
- Estimation: **Classical**? Bayesian?

Connectedness of Major U.S. Financial Institutions

$$\widehat{C}\left(x, H, B(L), M(L;\hat{\theta})\right)$$

- x: Thirteen daily realized stock return volatilities

  Commercial banks: JP Morgan Chase (JPM), Bank of America (BAC), CitiGroup (C), Wells Fargo (WFC), Bank of New York Mellon (BK), U.S. BankCorp (USB), PNC Bank (PNC)

  Investment Banks: Goldman Sachs (GS), Morgan Stanley (MS)

  GSEs: Fannie Mae (FNM), Freddie Mac (FRE)

  Insurance: AIG (AIG)

  Specialized: American Express (AXP)

- H: 12 days

- $M(L;\theta)$: logarithmic $VAR(3)$, generalized identification, 5/4/1999 - 4/30/2010

Penn
University of Pennsylvania

Full-Sample Connectedness Table

	AXP	BAC	BK	C	GS	JPM	MS	PNC	USB	WFC	AIG	FNM	FRE	**FROM**
AXP	20.0	8.5	7.1	10.3	5.8	9.8	8.8	5.1	8.0	7.8	3.2	2.6	3.0	80.0
BAC	8.3	19.1	6.0	10.6	5.8	8.0	7.4	6.1	7.1	9.2	4.2	3.5	4.6	80.9
BK	8.4	8.3	18.8	8.4	6.2	9.3	8.5	5.7	8.4	8.3	4.2	2.4	3.0	81.2
C	9.5	9.6	5.4	20.4	4.9	8.7	7.8	5.2	7.0	8.0	5.4	3.5	4.7	79.6
GS	8.2	8.6	6.8	7.6	22.1	8.8	13.3	4.0	6.0	7.6	2.4	1.9	2.6	77.9
JPM	10.2	8.6	7.1	10.6	6.2	18.8	9.5	5.2	7.8	7.3	3.6	2.5	2.6	81.2
MS	9.2	8.3	7.1	8.9	9.8	9.7	20.5	4.2	5.5	7.1	3.4	2.8	3.6	79.5
PNC	7.7	8.8	7.4	8.5	4.6	7.6	6.6	18.1	7.6	8.8	5.2	4.2	4.9	81.9
USB	9.3	9.9	7.6	9.9	5.7	8.7	6.4	5.4	20.1	8.5	4.3	1.6	2.7	79.9
WFC	8.3	10.2	6.5	9.8	6.2	7.6	7.1	5.9	7.3	18.0	3.8	3.8	5.3	82.0
AIG	5.3	7.3	4.9	8.8	2.6	5.2	4.9	6.2	6.0	5.6	27.5	6.6	9.0	72.5
FNM	4.2	5.4	2.5	6.0	2.3	3.5	3.8	5.5	1.9	6.8	6.5	29.6	22.0	70.4
FRE	4.3	6.3	2.9	6.5	2.6	3.3	4.1	5.2	2.9	7.3	7.4	17.6	29.6	70.4
TO	92.9	99.7	71.3	106.1	62.7	90.2	88.2	63.7	75.5	92.2	53.8	53.1	68.1	**78.3**

Estimating Time-Varying Connectedness

Before:

$$C\left(x, H, B(L), M(L;\theta)\right)$$
$$\widehat{C}\left(x, H, B(L), M(L;\hat{\theta})\right)$$

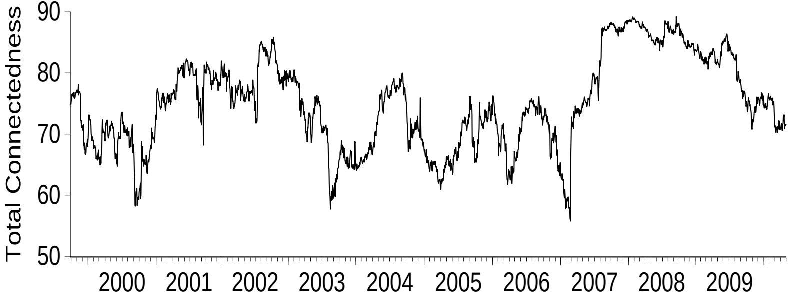
Now:

$$C_t\left(x, H, B_t(L), M(L;\theta_t)\right)$$
$$\widehat{C}_t\left(x, H, B_t(L), M(L;\hat{\theta}_t)\right)$$

- Time-varying parameters: **Rolling estimation**? Smooth TVP model? Regime-switching?

  (100-day estimation window)

Rolling Total Connectedness



Net Pairwise Directional Connectedness, The Lehman Bankruptcy

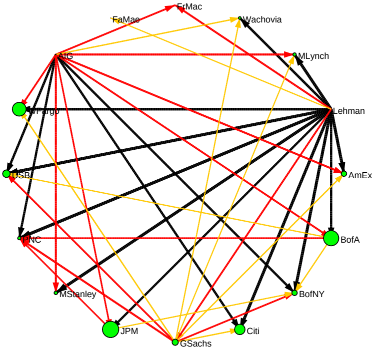


Figure: Sept. 16, 2008

In Conclusion...

Epilogue

- Reliable risk measurement requires *conditional* models that allow for time-varying volatility.
- Risk measurement may be done using univariate density models directly for portfolio returns. Many important recent developments.
- Certain aspects of risk *management* require fully-specified multivariate models. Many important recent developments, especially for N large.
- Volatility measures based on high-frequency return data hold great promise. A base-asset factor approach is often useful.
- The business cycle emerges as a key macroeconomic fundamental driving risk. Hence conditioning also applies at longer horizons.