# Using R for Hedge Fund of Funds Risk Management

## R/Finance 2009: Applied Finance with R
## University of Illinois Chicago, April 25, 2009

**Eric Zivot**
**Professor and Gary Waterman Distinguished Scholar, Department of Economics**
**Adjunct Professor, Departments of Finance and Statistics**
**University of Washington**

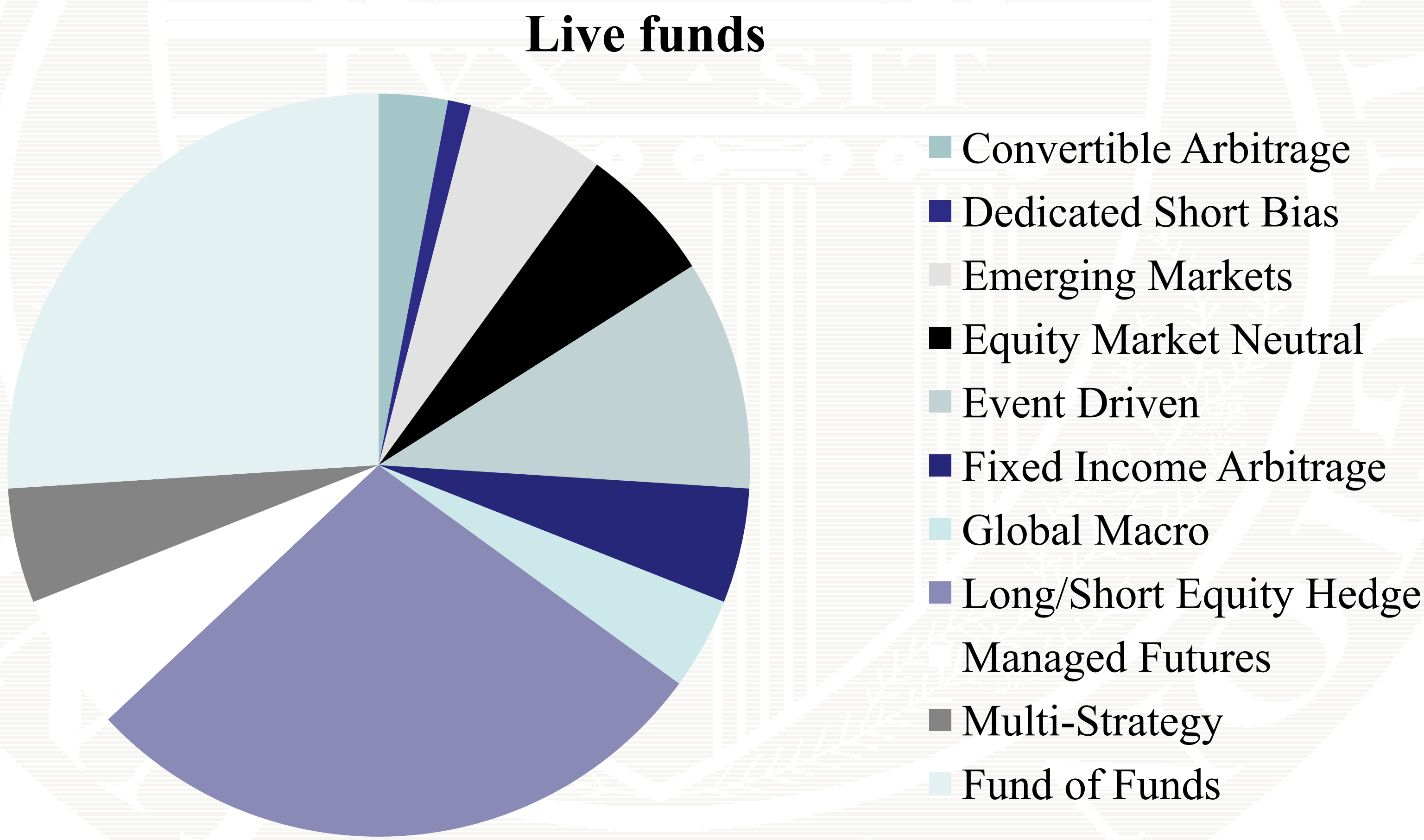
# Outline

- Hedge fund of funds environment
- Factor model risk measurement
- R implementation in corporate environment
- Dealing with unequal data histories
- Some thoughts on S-PLUS and S+FinMetrics vs. R in Finance

# Hedge Fund of Funds Environment

- HFoFs are hedge funds that invest in other hedge funds
  - 20 to 30 portfolios of hedge funds
  - Typical portfolio size is 30 funds
- Hedge fund universe is large: 5000 live funds
  - Segmented into 10-15 distinct strategy types
- Hedge funds voluntarily report monthly performance to commercial databases
  - Altvest, CISDM, HedgeFund.net, Lipper TASS, CS/Tremont, HFR
- HFoFs often have partial position level data on invested funds

# Hedge Fund Universe



# Characteristics of Monthly Returns

- Reporting biases
  - Survivorship, backfill
- Non-normal behavior
  - Asymmetry (skewness) and fat tails (excess kurtosis)
- Serial correlation
  - Performance smoothing, illiquid positions
- Unequal histories

# Characteristics of Hedge Fund Data

| | fund1 | fund2 | fund3 | fund4 | fund5 |
|---|---|---|---|---|---|
| Observations | 122.0000 | 107.0000 | 135.0000 | 135.0000 | 135.0000 |
| NAs | 13.0000 | 28.0000 | 0.0000 | 0.0000 | 0.0000 |
| Minimum | -0.0842 | -0.3649 | -0.0519 | -0.1556 | -0.2900 |
| Quartile 1 | -0.0016 | -0.0051 | 0.0020 | -0.0017 | -0.0021 |
| Median | 0.0058 | 0.0046 | 0.0060 | 0.0073 | 0.0049 |
| Arithmetic Mean | 0.0038 | -0.0017 | 0.0063 | 0.0059 | 0.0021 |
| Geometric Mean | 0.0037 | -0.0029 | 0.0062 | 0.0055 | 0.0014 |
| Quartile 3 | 0.0158 | 0.0129 | 0.0127 | 0.0157 | 0.0127 |
| Maximum | 0.0311 | 0.0861 | 0.0502 | 0.0762 | 0.0877 |
| Variance | 0.0003 | 0.0020 | 0.0002 | 0.0008 | 0.0013 |
| Stdev | 0.0176 | 0.0443 | 0.0152 | 0.0275 | 0.0357 |
| Skewness | -1.7753 | -5.6202 | -0.8810 | -2.4839 | -4.9948 |
| Kurtosis | 5.2887 | 40.9681 | 3.7960 | 13.8201 | 35.8623 |
| Rho1 | 0.6060 | 0.3820 | 0.3590 | 0.4400 | 0.383 |

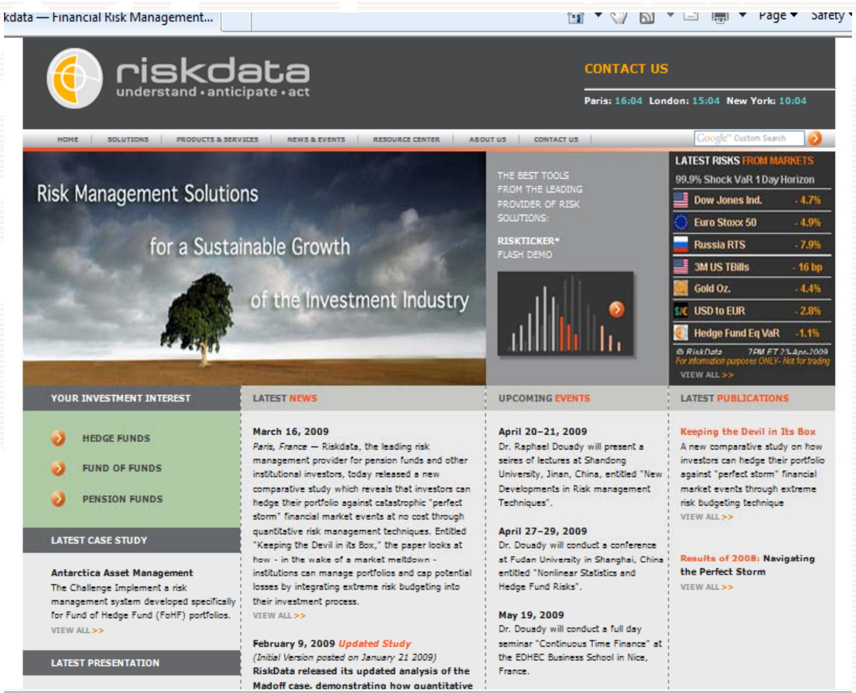
Sample: January 1998 – March 2009

# Factor Model Risk Measurement in HFoFs Portfolio

- Quantify factor risk exposures
  - Equity, rates, credit, volatility, currency, commodity, etc.
- Quantify tail risk
  - VaR, ETL
- Risk budgeting
  - Component, incremental, marginal
- Stress testing and scenario analysis

# Commercial Products

www.riskdata.com




www.finanalytica.com


Very expensive! R is not!

# Factor Model: Methodology

$$R_{it} = \alpha_i + \beta_{i1}F_{1t} + \cdots + \beta_{ik}F_{kt} + \varepsilon_{it},$$
$$= \alpha_i + \boldsymbol{\beta}_i'\mathbf{F_t} + \varepsilon_{it}$$
$$i = 1,\ldots,n;\ t = t_i,\ldots,T$$
$$\mathbf{F}_t \sim (\boldsymbol{\mu}_{\mathrm{F}}, \boldsymbol{\Sigma}_F)$$
$$\varepsilon_{it} \sim (0, \sigma^2_{\varepsilon,i})$$
$$\mathrm{cov}(f_{jt}, \varepsilon_{it}) = 0 \text{ for all } j,\ i \text{ and } t$$
$$\mathrm{cov}(\varepsilon_{it}, \varepsilon_{jt}) = 0 \text{ for } i \neq j$$

# Practical Considerations

- Many potential risk factors ( > 50)
- High collinearity among some factors
- Risk factors vary across discipline/strategy
- Nonlinear effects
- Dynamic effects
- Time varying coefficients
- Common histories for factors; unequal histories for fund performance

# Expected Return Decomposition

$$E[R_{it}] = \alpha_i + \beta_{i1}E[F_{1t}] + \cdots + \beta_{ik}E[F_{kt}]$$

Expected return due to "beta" exposure

$$\beta_{i1}E[F_{1t}] + \cdots + \beta_{ik}E[F_{kt}]$$

Expected return due to manager specific "alpha"

$$\alpha_i = E[R_{it}] - (\beta_{i1}E[F_{1t}] + \cdots + \beta_{ik}E[F_{kt}])$$

# Variance Decomposition

$$\mathrm{var}(R_{it}) = \underbrace{\boldsymbol{\beta}_i' \,\mathrm{var}(\mathbf{F}_t)\boldsymbol{\beta}_i}_{\text{systematic}} + \underbrace{\mathrm{var}(\varepsilon_{it})}_{\text{specific}} = \boldsymbol{\beta}_i'\boldsymbol{\Sigma}_{\mathbf{F}}\boldsymbol{\beta}_i + \sigma^2_{\varepsilon,i}$$

Variance contribution due to factor exposures

$$\beta_1^2 \,\mathrm{var}(F_{1t}) + \beta_2^2 \,\mathrm{var}(F_{2t}) + \cdots + \beta_k^2 \,\mathrm{var}(F_{kt})$$

Variance contribution due to covariances between factors

$$2\beta_1\beta_2 \,\mathrm{cov}(F_{1t}, F_{2t}) + \cdots + 2\beta_{k-1}\beta_k \,\mathrm{cov}(F_{k-1t}, F_{kt})$$

# Covariance

$$\underset{n\times 1}{\boldsymbol{R}_t} = \underset{n\times 1}{\boldsymbol{\alpha}} + \underset{n\times k}{\boldsymbol{B}}\,\underset{k\times 1}{\boldsymbol{F}_t} + \underset{n\times 1}{\boldsymbol{\varepsilon}_t}$$

$$\mathrm{var}(R_t) = \Sigma_{FM} = \mathbf{B}\boldsymbol{\Sigma}_{\mathbf{F}}\mathbf{B}' + \mathbf{D}_{\varepsilon}$$

$$\mathbf{D}_{\varepsilon} = diag(\sigma^2_{\varepsilon,1},\ldots,\sigma^2_{\varepsilon,n})$$

Note: $\mathrm{cov}(R_{it}, R_{jt}) = \boldsymbol{\beta}_i' \,\mathrm{var}(\mathbf{F}_t)\boldsymbol{\beta}_j = \boldsymbol{\beta}_i'\boldsymbol{\Sigma}_{\mathbf{F}}\boldsymbol{\beta}_j$

# Portfolio Analysis

$\mathbf{w} = (w_1, \ldots, w_n)' =$ portfolio weights

$$\begin{aligned} R_{pt} &= \mathbf{w}'\mathbf{R}_t = \mathbf{w}'\boldsymbol{\alpha} + \mathbf{w}'\mathbf{B}\mathbf{F}_t + \mathbf{w}'\boldsymbol{\varepsilon}_t \\ &= \sum_{i=1}^{n} w_i R_{it} = \sum_{i=1}^{n} w_i \alpha_i + \sum_{i=1}^{n} w_i \boldsymbol{\beta}_i' \mathbf{F}_t + \sum_{i=1}^{n} w_i \varepsilon_{it} \\ &= \alpha_p + \boldsymbol{\beta}_p' \mathbf{F}_t + \varepsilon_{pt} \end{aligned}$$

# Portfolio Variance Decomposition

$$\sigma_p^2 = \mathrm{var}(R_{pt}) = \mathbf{w}'\,\mathrm{var}(\mathbf{R}_t)\mathbf{w} = \mathbf{w}'\mathbf{B}\mathbf{\Sigma}_F\mathbf{B}'\mathbf{w} + \mathbf{w}'\mathbf{D}\mathbf{w}$$

$$\sigma^2_{p,systematic} = \mathbf{w}'\mathbf{B}\mathbf{\Sigma}_F\mathbf{B}'\mathbf{w}$$

$$\sigma^2_{p,specific} = \mathbf{w}'\mathbf{D}\mathbf{w} = \sum_{i=1}^{n} w_i \sigma^2_{\varepsilon,i} \qquad R_p^2 = \frac{\sigma^2_{p,systematic}}{\sigma_p^2}$$

$$\sigma^2_{p,systematic} = \beta_p' \Sigma_F \beta_p = \sum_{j=1}^{k} \beta^2_{p,j}\sigma^2_{jj} + \text{covariance terms}$$

$$\text{covariance terms} = \sigma^2_{p,systematic} - \sum_{j=1}^{k} \beta^2_{p,j}\sigma^2_{jj}$$

# Risk Budgeting: Volatility

$$\mathbf{MCR} = \frac{\partial \sigma_p}{\partial \mathbf{w}} = \frac{\mathbf{B}\Sigma_{\mathbf{F}}\mathbf{B}'\mathbf{w} + \mathbf{D}\mathbf{w}}{\sigma_p}$$

Marginal contributions to risk

$$\mathbf{MCR}_{systematic} = \frac{\mathbf{B}\Sigma_{\mathbf{F}}\mathbf{B}'\mathbf{w}}{\sigma_p}$$

$$\mathbf{MCR}_{specific} = \frac{\mathbf{D}\mathbf{w}}{\sigma_p}$$

$$\mathbf{CR} = \mathbf{w} \odot \frac{\partial \sigma_p}{\partial \mathbf{w}} = \frac{\mathbf{w} \odot (\mathbf{B}\Sigma_{\mathbf{F}}\mathbf{B}'\mathbf{w} + \mathbf{D}\mathbf{w})}{\sigma_p}$$

Components to risk

$$\mathbf{1}'\mathbf{CR} = \sum_{i=1}^{n} CR_i = \sigma_p$$

# Tail Risk Measures

Value-at-Risk (VaR)

$$VaR_{\alpha} = -q_{\alpha} = -F^{-1}(\alpha)$$

$$F = CDF \text{ of returns } R$$

Expected Shortfall (ES)

$$ES_{\alpha} = -E[R \mid R \leq VaR_{\alpha}]$$

# Tail Risk Measures: Normal Distribution

$$R_p \sim N(\mu_p, \sigma_p^2),\ \sigma_p^2 = w'\Sigma_{FM} w$$

$$VaR_\alpha^N = -\mu_p - \sigma_p \times z_\alpha,\ z_\alpha = \Phi^{-1}(\alpha)$$

$$ES_\alpha^N = \mu_p - \sigma_p \frac{1}{\alpha} \phi(z_\alpha)$$

See functions in PerformanceAnalytics

# Tail Risk Measures: Non-Normal Distributions

Use Cornish-Fisher expansion to account for asymmetry and fat tails

$$VaR_{\alpha}^{CF} = -\mu_i - \sigma_i \times z_{\alpha}$$
$$+\sigma_i \left[ -\frac{1}{6}\left(z_{\alpha}^2 - 1\right) skew_i - \frac{1}{24}\left(z_{\alpha}^3 - 3z_{\alpha}\right) ekurt_i + \frac{1}{36}\left(2z_{\alpha}^3 - 5z_{\alpha}\right) skew_i^2 \right]$$

$ES_{\alpha}^{CF}$ : Formula given in Boudt, Peterson and Croux (2008) "Estimation and Decomposition of Downside Risk for Portfolios with Non-Normal Returns," *Journal of Risk* and implementation in PerformanceAnalytics

# Risk Budgeting: Tail Risk

Value-at-Risk (VaR)

$$VaR_{\alpha} = \sum_{i=1}^{n} w_i \frac{\partial VaR_{\alpha}}{\partial w_i} = \sum_{i=1}^{n} \overbrace{w_i \times mVaR_{\alpha,i}}^{cVaR_{\alpha,i}},$$

$$mVaR_{\alpha,i} = \frac{\partial VaR_{\alpha}}{\partial w_i} = -E[R_i \mid R_p = VaR_{\alpha}]$$

Expected Shortfall (ES)

$$ES_{\alpha} = \sum_{i=1}^{n} w_i \frac{\partial ES_{\alpha}}{\partial w_i} = \sum_{i=1}^{n} \overbrace{w_i \times mES_{\alpha,i}}^{cES_{\alpha,i}},$$

$$mES_{\alpha,i} = \frac{\partial ES_{\alpha}}{\partial w_i} = -E[R_i \mid R_p \leq VaR_{\alpha}]$$

# Risk Budgeting: Explicit Formulas

- Normal distribution
  - See Jorian (2007) or Dowd (2002)
- Non-normal distribution using Cornish-Fisher expansion
  - See Boudt, Peterson and Croux (2008) "Estimation and Decomposition of Downside Risk for Portfolios with Non-Normal Returns," *Journal of Risk* and implementation in PerformanceAnalytics

# Risk Budgeting: Simulation

$\{R_{it}\}_{t=1}^{M} = M$ simulated returns

Method 1: Brute Force

$$mVaR_{\alpha,i} \approx \frac{\Delta VaR_{\alpha}}{\Delta w_i},\ mES_{\alpha,i} = \frac{\Delta ES_{\alpha}}{\Delta w_i}$$

Method 2: Average $R_{it}$ around values for which $R_{pt} = VaR_{\alpha}$

$$mVaR_{\alpha,i} \approx -\sum_{t:R_{pt}=VaR_{\alpha}\pm\varepsilon} R_{it},\ mES_{\alpha,i} \approx -\sum_{t:R_{pt}\le VaR_{\alpha}} R_{it}$$

# R Functions for Factor Model Risk Analysis

| Function | Function |
| --- | --- |
| factorModelCovariance | normalES |
| factorModelRiskDecomposition | normalPortfolioES |
| normalVaR | normalMarginalES |
| normalPortfolioVaR | normalComponentES |
| normalMarginalVaR | modifiedES |
| normalComponentVaR | modifiedPortfolioES |
| normalVaRreport | modifiedESreport |
| modifiedVaR | simulatedMarginalVaR |
| modifiedPortfolioVaR | simulatedComponentVaR |
| modifiedMarginalVaR | simulatedMarginalES |
| modifiedComponentVaR | simulatedComponentES |

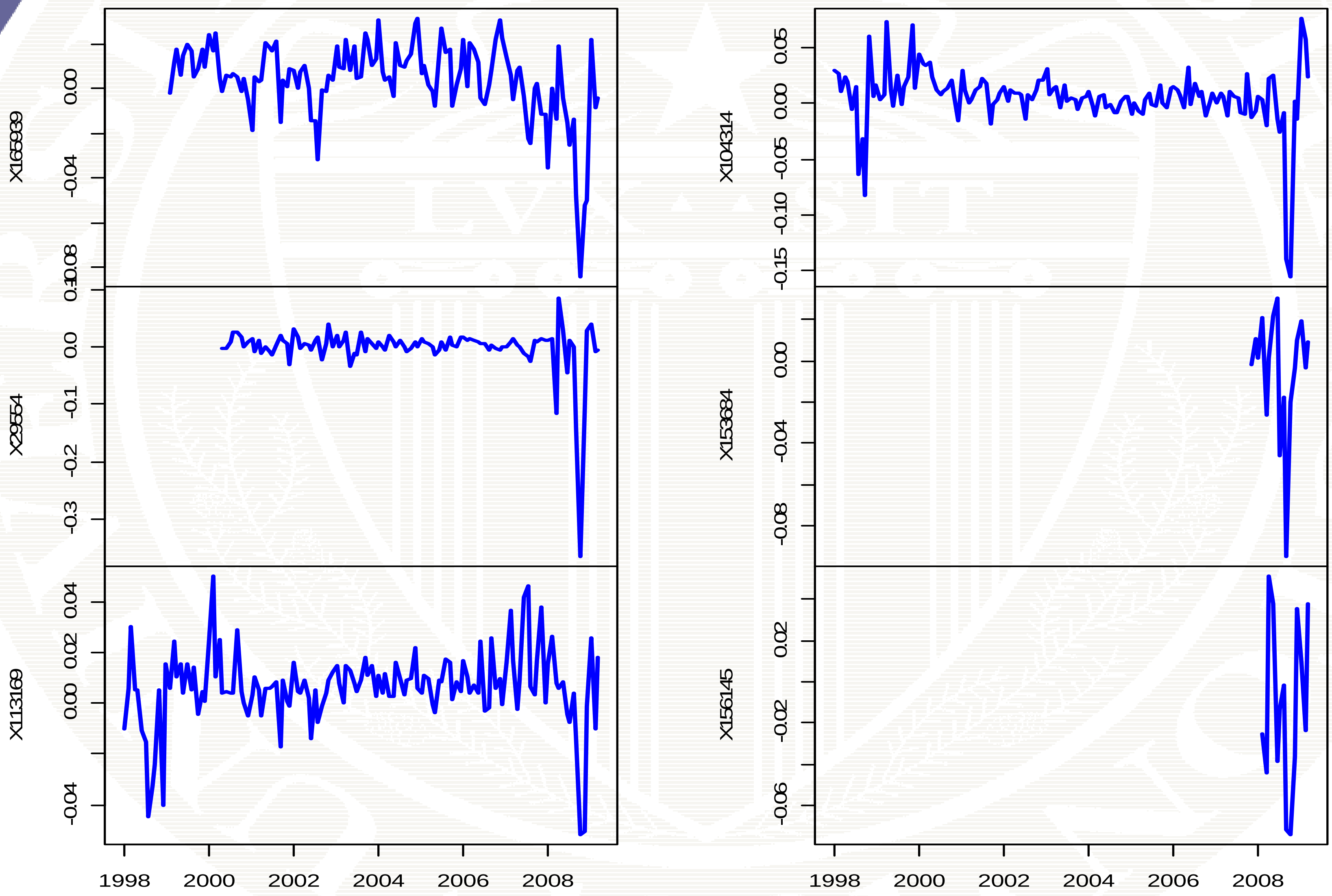
# Unequal Histories

| Risk factors | Fund performance | |
|---|---|---|
| $F_{1,T}, \ldots, F_{k,T}$ | $R_{1,T}$ | $R_{n,T}$ |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $F_{1,T-T_i}, \ldots, F_{k,T-T_i}$ | $R_{1,T-T_1}$ | $\vdots$ |
| $\vdots$ | | $R_{n,T-T_n}$ |
| $F_{1,1}, \ldots, F_{k,1}$ | | |

Observe full history

Observe partial histories

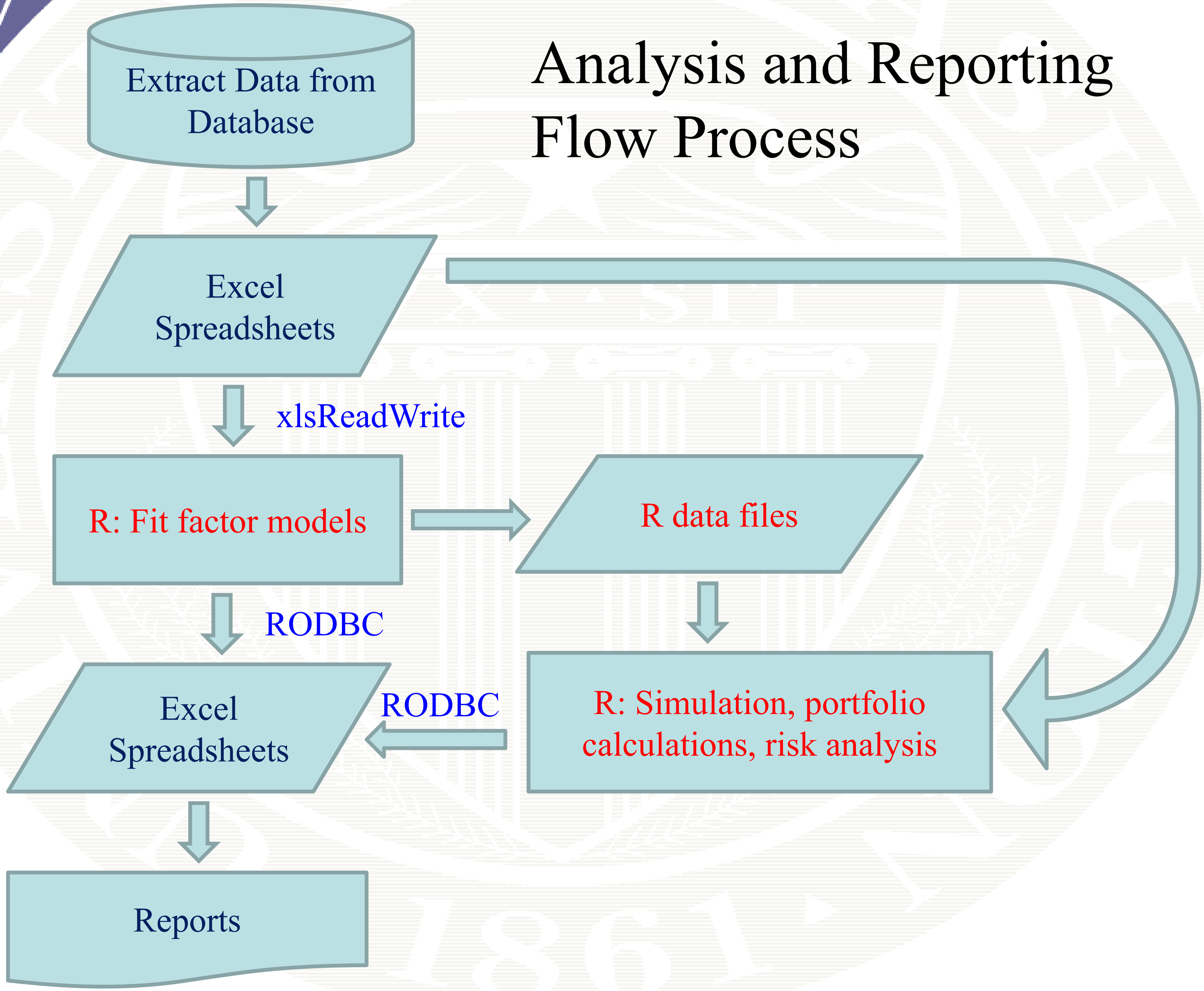
# Example Portfolio: Unequal Histories of Individual Funds



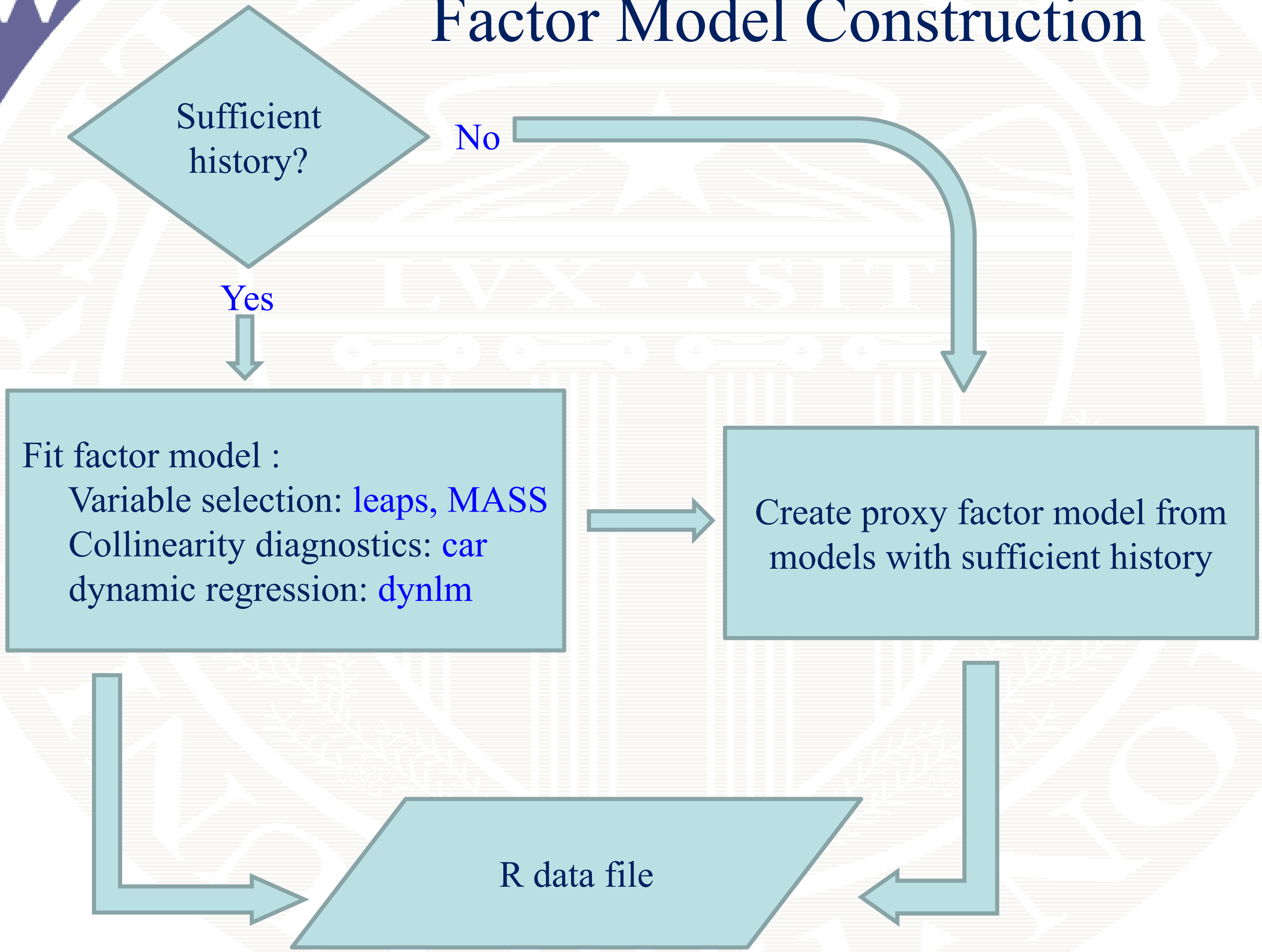
# Implication of Unequal Histories

- Can’t fit factor models to some funds
  - Need to create proxy factor model
- Statistics on common histories (truncated data) may be unreliable
- Difficult to compute non-normal tail risk measures

# Analysis and Reporting Flow Process



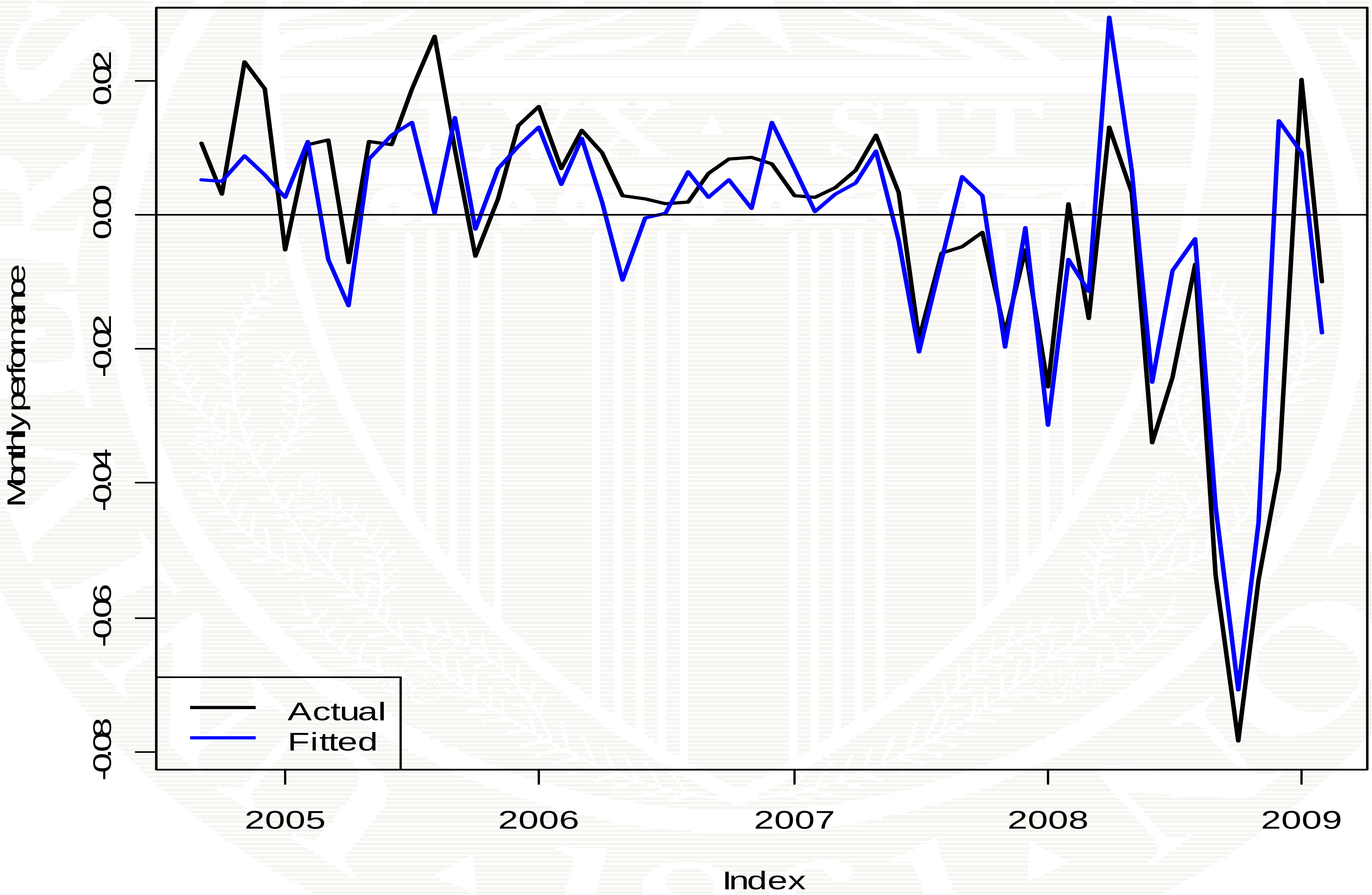
# Factor Model Construction



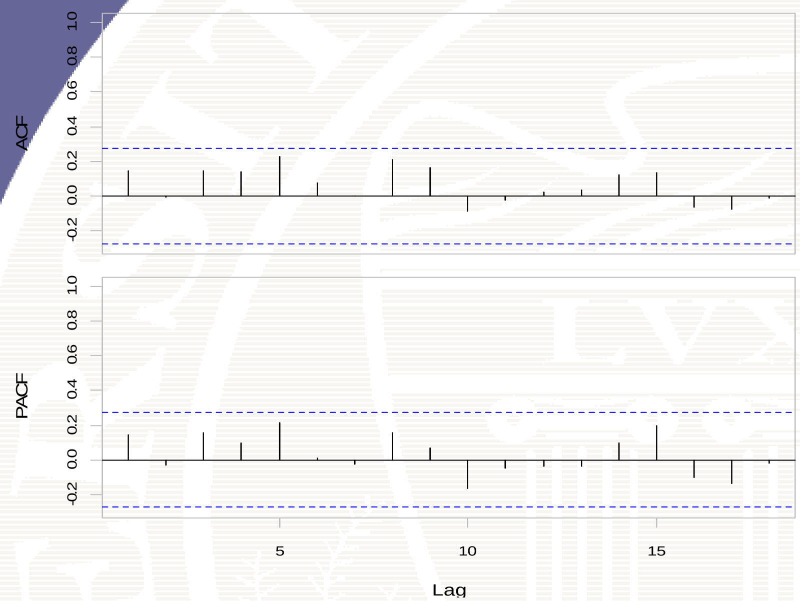
# Evaluation of Fitted Factor Models

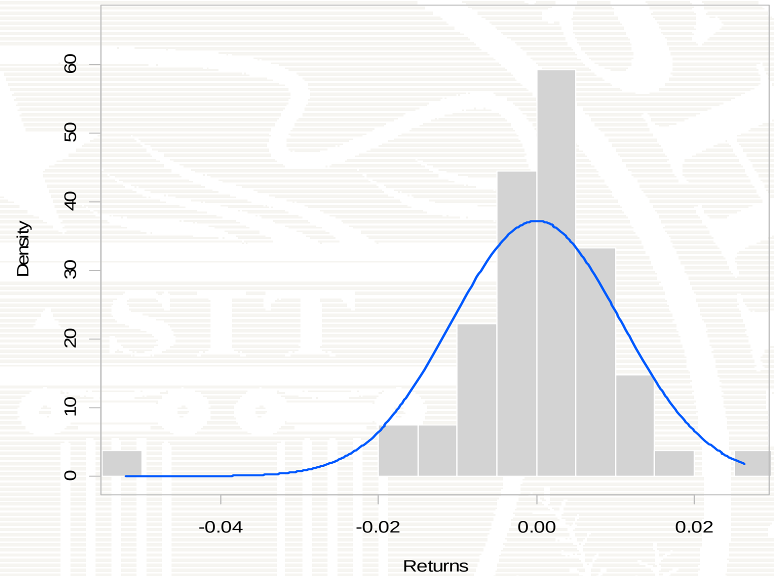
- Graphical diagnostics
  - Created plot method appropriate for time series regression.
- Stability analysis
  - CUSUM etc: strucchange
  - Rolling analysis: rollapply (zoo)
  - Time varying parameters: dlm
- Dynamic effects
  - dynlm, lmtest

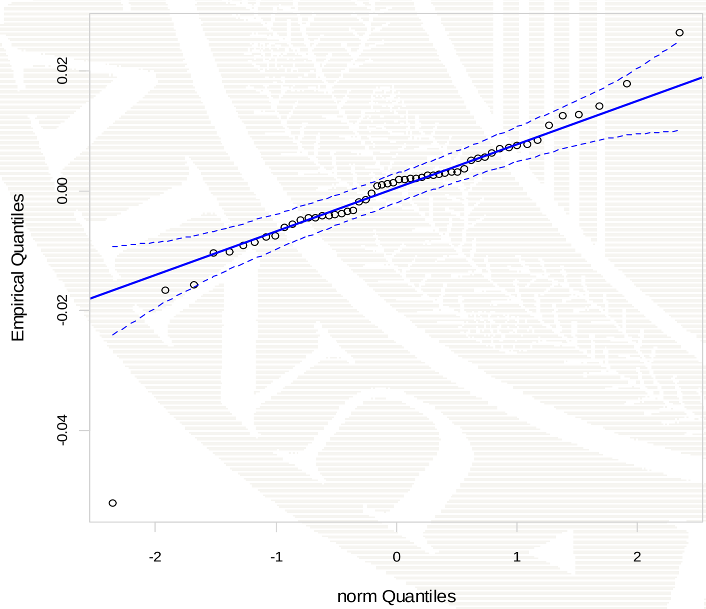
# Diagnostic Plots: Example Fund

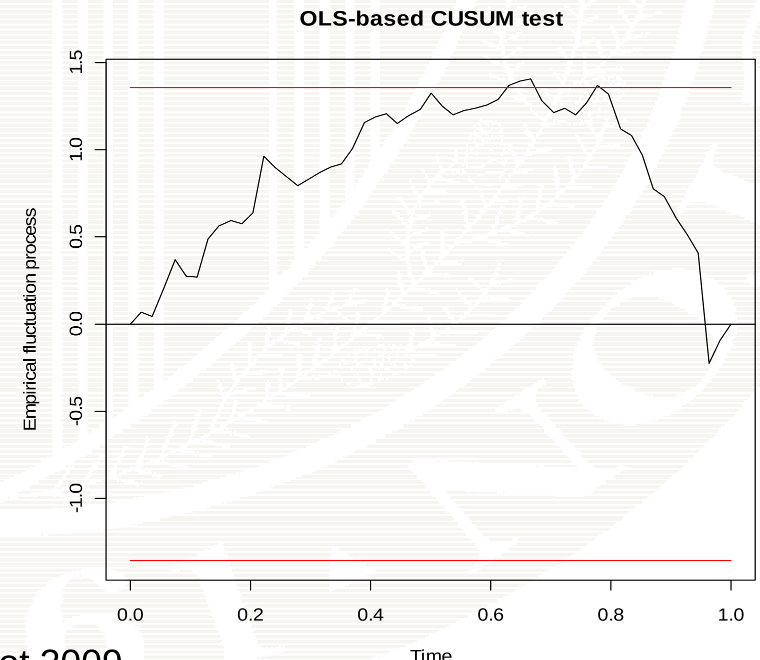


# Time Series Regression Residual diagnostics

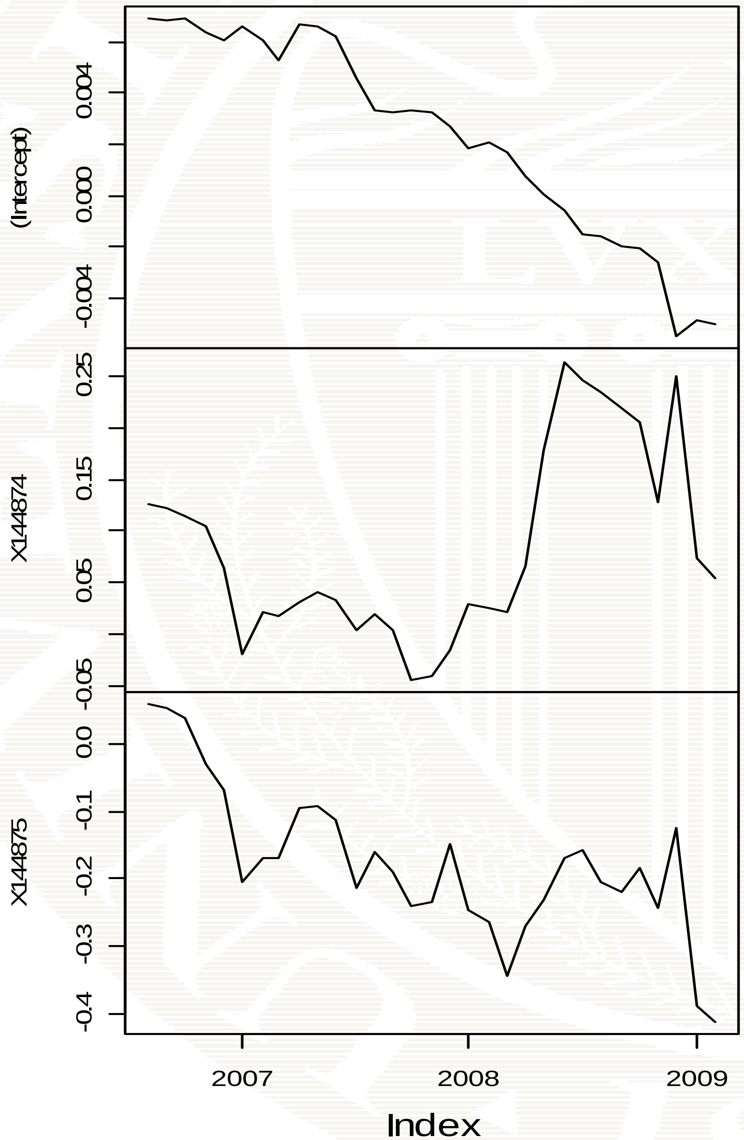


# 24-month rolling estimates




# Dealing with Unequal Histories

- Estimate conditional distribution of $R_i$ given $\boldsymbol{F}$
  - Fitted factor model or proxy factor model
- Estimate marginal distribution of $\boldsymbol{F}$
  - Empirical distribution, multivariate normal, copula
- Derive marginal distribution of $R_i$ from $p(R_i|\boldsymbol{F})$ and $p(\boldsymbol{F})$
- Simulate $R_i$ and Calculate functional of interest
  - Unobserved performance, Sharpe ratio, ETL etc

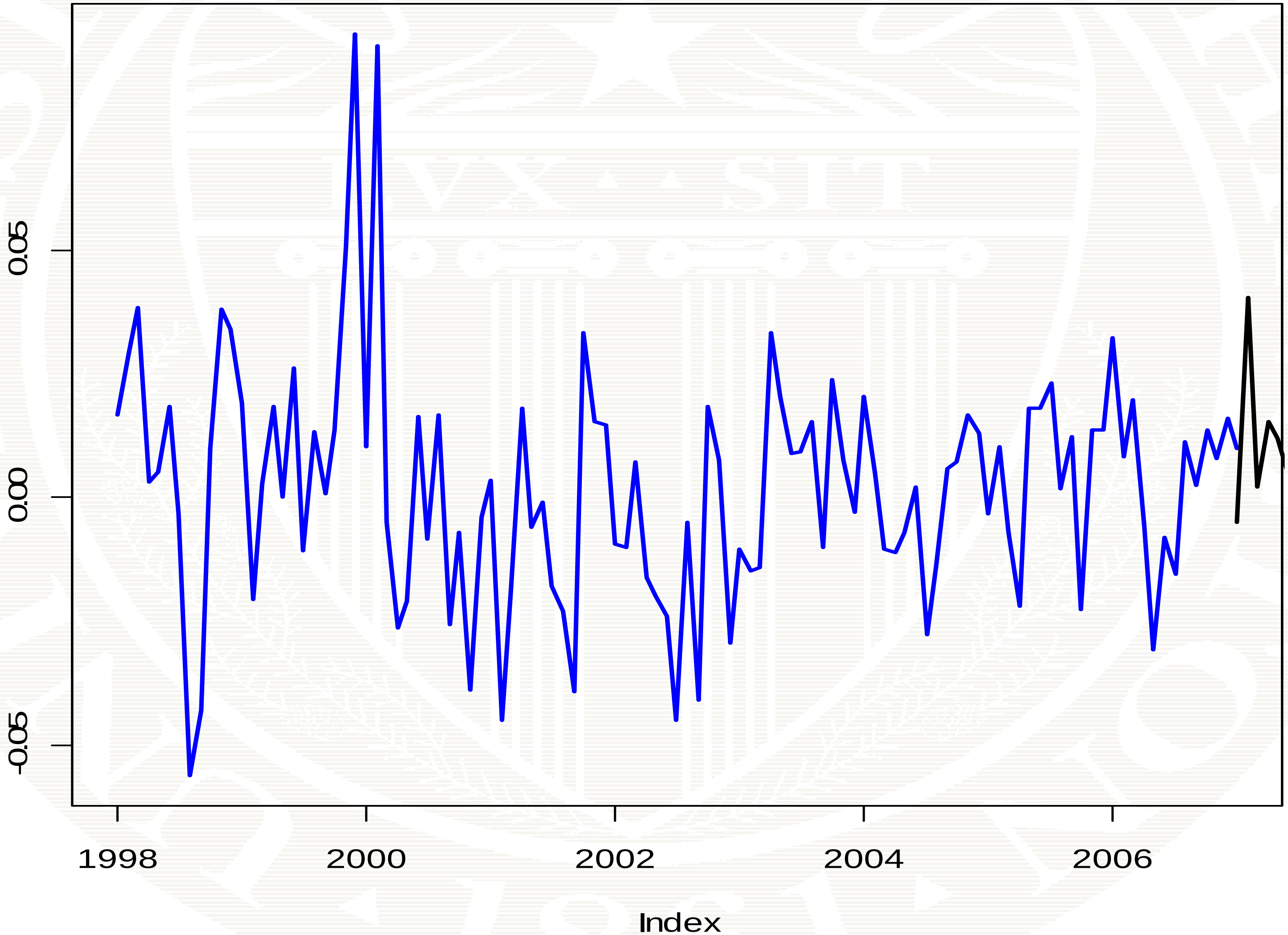
# Simulation Algorithm

- Draw $\left\{ \tilde{\boldsymbol{F}}_1, \ldots, \tilde{\boldsymbol{F}}_M \right\}$ by resampling from the empirical distribution of $\boldsymbol{F}$.
- For each $\widetilde{\boldsymbol{F}_u}$ ($u = 1, \ldots, M$), draw a value $\tilde{R}_{i,u}$ from the estimated conditional distribution of $R_i$ given $\boldsymbol{F}=\tilde{\boldsymbol{F}}_u$ (e.g., from fitted factor model assuming normal errors)
- $\left\{ \tilde{R}_{i,u} \right\}_{u=1}^{M}$ is the desired sample for $R_i$
- M ≈ 5000

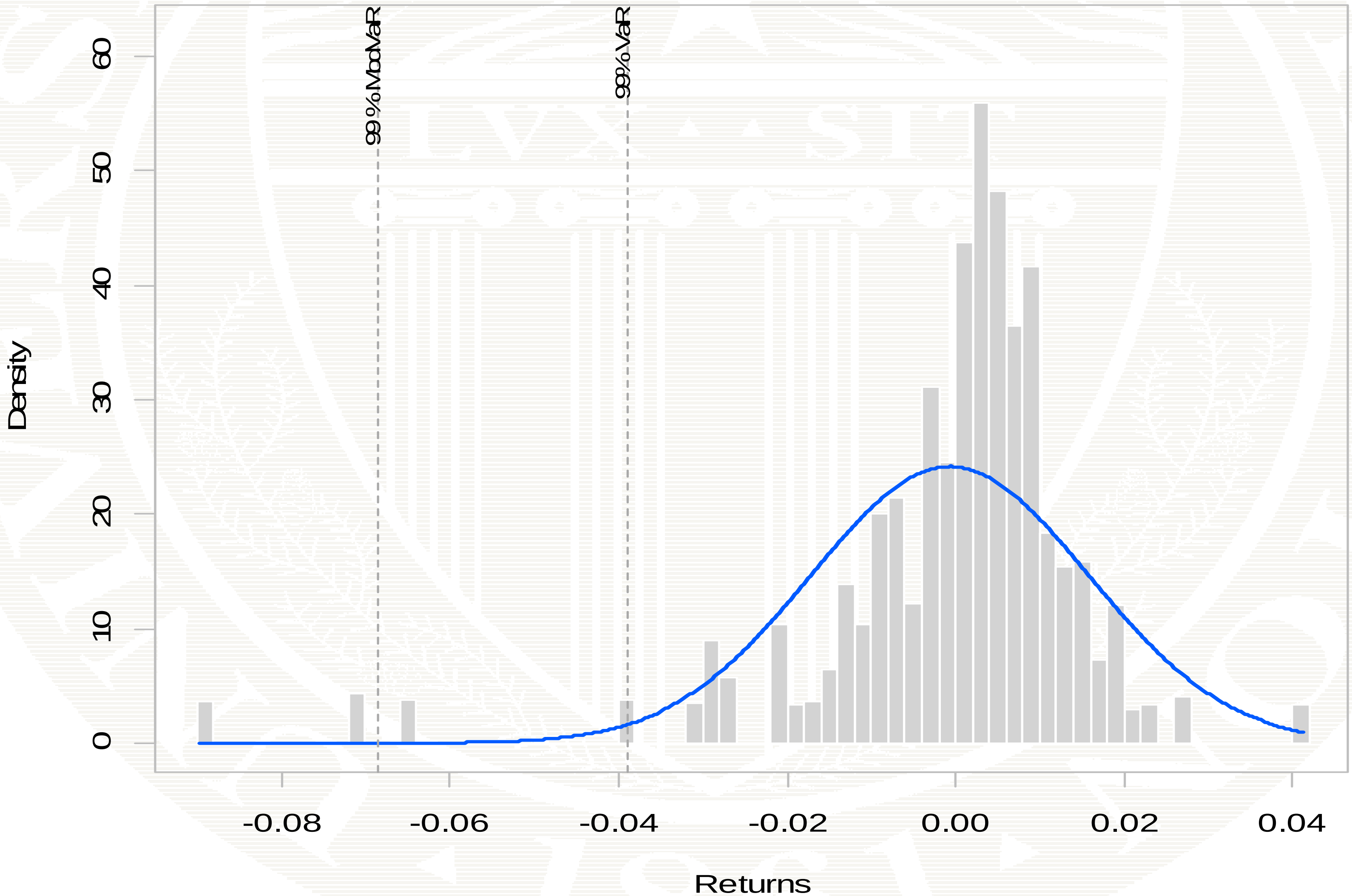
# What to do with $\left\{\tilde{R}_{i,u}\right\}_{u=1}^{M}$?

- Backfill missing fund performance
- Compute fund and portfolio performance measures
- Estimate non-parametric fund and portfolio tail risk measures
- Compute non-parametric risk budgeting measures
- Standard errors can be computed using a bootstrap procedure

# Example: Backfilled Fund Performance



# Example: Simulated portfolio distribution



# S-PLUS and S+FinMetrics vs R

- Dealing with time series objects in R can be difficult and confusing
  - timeSeries, zoo, xts
- Time series regression in R is incompletely implemented
  - Diagnostic plots, prediction
- R packages give about 80% functional coverage to S+FinMetrics

# Some Thoughts About Using R in a Corporate Environment

- IT doesn’t want to support it
- Firewalls block R downloads
- The world runs from an Excel spreadsheet
- Analysts with some programming experience learn R quickly
- Not good for the casual user

# References


- Boudt, K., B. Peterson and C. Croux (2008) "Estimation and Decomposition of Downside Risk for Portfolios with Non-Normal Returns," *Journal of Risk*
- Dowd, K. (2002), *Measuring Market Risk*, John Wiley and Sons
- Goodworth, T. and C. Jones (2007). “Factor-based, Non-parametric Risk Measurement Framework for Hedge Funds and Fund-of-Funds,” *The European Journal of Finance*.
- Hallerback, J. (2003). “Decomposing Portfolio Value-at-Risk: A General Analysis”, *The Journal of Risk* 5/2.
- Jorian, P. (2007), *Value at Risk*, Third Edition, McGraw Hill

# References


- Jiang, Y. (2009). *Overcoming Data Challenges in Fund-of-Funds Portfolio Management*, PhD Thesis, Department of Statistics, University of Washington.
- Lo, A. (2007). *Hedge Funds: An Analytic Perspective*, Princeton.
- Yamai, Y. and T. Yoshiba (2002). "Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization, *Institute for Monetary and Economic Studies*, Bank of Japan.