Efficient Method of Moments

by

A. Ronald Gallant
Department of Economics
University of North Carolina
Chapel Hill NC 27599-3305 USA

George Tauchen
Department of Economics
Duke University
Durham NC 27708-0097 USA

WEB: www.unc.edu/~arg

FTP: ftp.econ.duke.edu, pub/arg/papers/ee.ps

Dynamic Models with Unobserved States

- Standard statistical methods, both classical and Bayesian, are usually not applicable either because it is not practicable to obtain the density of the state vector or because the integration required to eliminate unobserved states from the likelihood is infeasible.

- On a case-by-case basis, statistical methods are sometimes available. However, our purpose here is to describe methods that are generally applicable.

- Simulating the evolution of the state vector is often practicable. Our methods rely on this.

- We describe simulated method of moments methods in general and then focus the discussion on efficient method of moments (EMM).

References

Theory

Gallant, A. Ronald, and George Tauchen (1996), "Which Moments to Match?," *Econometric Theory* 12, 657–681.

Gallant, A. R., and Long, J. R. (1997), "Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Squared," *Biometrika*, 84, 125–141.

Application

Gallant, A. Ronald, and George Tauchen (2001), "Efficient Method of Moments," Manuscript, Deparment of Economics, University of North Carolina.

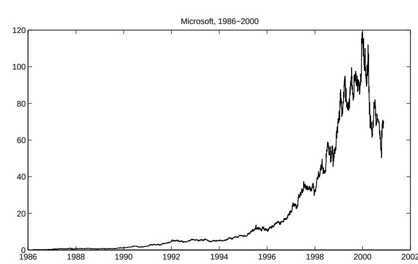
Efficiency

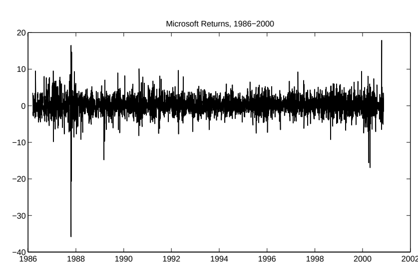
Gallant, A. Ronald, and George Tauchen (1999), "The Relative Efficiency of Method of Moments Estimators," *Journal of Econometrics* 92, 149–172.

Cross Validation

Coppejans, Mark, and A. Ronald Gallant (2001), "Cross Validated SNP Density Estimates," Manuscript, Department of Economics, Duke University.

Microsoft, 1986–2000





- 3,711 daily observations

A Stochastic Volatility Model of Equity Prices

$$\begin{pmatrix} dU_1 \\ dU_2 \\ dU_3 \end{pmatrix} = \begin{pmatrix} \alpha_{10} \\ \alpha_{22}U_2 \\ \alpha_{33}U_3 \end{pmatrix} dt + \begin{pmatrix} e^{(\beta_{10}+\beta_{12}U_2+\beta_{13}U_3)} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} dW_1 \\ dW_2 \\ dW_3 \end{pmatrix}$$

$$y_t = 100(U_{1,t} - U_{1,t-1}), \quad t = 1, 2, \cdots$$

- U_{1t} is the continuous record of the log stock price
- U_{2t} and U_{3t} are unobserved volatility factors
- y_t is the discretely sampled daily return.

Parameters of the System

$$\rho = (\alpha_{10}, \alpha_{22}, \alpha_{33}, \beta_{10}, \beta_{12}, \beta_{13}),$$

Matrix Formulation

$$dU_t = A(U_t)dt + B(U_t)dW_t$$

Simulation Methods

SDE

$$dU_t = A(U_t)dt + B(U_t)dW_t$$

Iterate for small delta

$$U_\Delta - U_0 = A(U_0)\Delta + B(U_0)(W_\Delta - W_0)$$

Sum

$$\sum_{i=1}^{t/\Delta}[U_{i\Delta} - U_{(i-1)\Delta}] = \sum_{i=1}^{t/\Delta} A[U_{(i-1)\Delta}]\Delta + \sum_{i=1}^{t/\Delta} B[U_{(i-1)\Delta}][W_{i\Delta} - W_{(i-1)\Delta}]$$

Limit as delta decreases

$$U_t - U_0 = \int_0^t A(U_s)ds + \int_0^t B(U_s)dW_s$$

Outline

- Dynamical Systems with Unobserved States
  - Motivation
  - Example: A diffusion model of equity prices
  - Simulation methods for diffusions
- Simulated Method of Moments
  - Minimum chi-squared estimators
  - Efficiency considerations
  - Data determined bandwidth
- Methodology
  - Projection
  - Estimation
  - Reprojection

Assumption Underlying the Methodology

For ρ in the parameter space, the system is presumed to be stationary, ergodic, and convenient to simulate.

Consequence

For any lag length L and parameter setting ρ there exists a stationary density for observables

$$(y_{t-L}, \ldots, y_t) \sim p(y_{-L}, \ldots, y_0, \rho).$$

And, an unconditional expectation

$$\mathcal{E}_\rho\Big\{\psi(y_{-L}, \ldots, y_0)\Big\}$$

can be computed by generating a long simulation $\{\hat{y}_t\}_{t=-L}^N$ and averaging

$$\mathcal{E}_\rho\Big\{\psi(y_{-L}, \ldots, y_0)\Big\} = \frac{1}{N}\sum_{t=0}^{N} \psi(\hat{y}_{t-L}, \ldots, \hat{y}_t).$$

Notation

Random Variables: $y_t \quad t = -\infty, \ldots, \infty$
Data: $\tilde{y}_i \quad i = -L, \ldots, n$
Simulation: $\hat{y}_i \quad i = -L, \ldots, N$

Dummy arguments of summation:

$$(y_{t-L}, \ldots, y_{t-1}, y_t)$$

$$y_t, \quad x_{t-1} = (y_{t-L}, \ldots, y_{t-1})$$

Dummy arguments of integration (put $t = 0$):

$$(y_{-L}, \ldots, y_{-1}, y_0)$$

$$y = y_0, \quad x = (y_{-L}, \ldots, y_{-1})$$

EMM Estimator

For a QMLE estimator

$$\tilde{\theta}_n = \underset{\theta}{\text{argmax}} \frac{1}{n} \sum_{t=1}^{n} \log f(\tilde{y}_t | \tilde{x}_{t-1}, \theta),$$

a sample average satisfies

$$0 = \frac{1}{n} \sum_{t=1}^{n} \frac{\partial}{\partial \theta} \log f(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n).$$

because these are the first order conditions of the optimization problem.

Therefore a large simulation from a the model with parameters ρ should satisfy the estimating equations

$$0 = m(\rho, \tilde{\theta}_n) = \frac{1}{N} \sum_{t=1}^{N} \frac{\partial}{\partial \theta} \log f(\hat{y}_t | \hat{x}_{t-1}, \tilde{\theta}_n),$$

except for sampling variation in $\tilde{\theta}_n$. These equations hold exactly in the limit as n and N tend to infinity.

The EMM estimator attempts to find ρ that solves these estimating equations in a sense made precise in the next transparency.

Minimum Chi-squared

If the equations $m(\rho, \tilde{\theta}_n) = 0$ cannot be solved because the dimension of θ is larger than the dimension of ρ, then

- Use a nonlinear optimizer to minimize:
$$\hat{\rho}_n = \underset{\rho}{\text{argmin}}\ m_n'(\rho, \tilde{\theta}_n) \left(\tilde{\mathcal{I}}_n\right)^{-1} m_n(\rho, \tilde{\theta}_n);$$
- $\tilde{\mathcal{I}}_n$ estimates the variance of $\sqrt{n}\, m_n'(\rho^o, \tilde{\theta}_n)$. If $f(y|x, \theta)$ is a good approximation to the true data generating process $p(y|x, \rho^o)$, then an adequate estimator is

$$\tilde{\mathcal{I}}_n = \frac{1}{n} \sum_{t=1}^{n} \left[\frac{\partial}{\partial \theta} \log f_K(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n)\right]\left[\frac{\partial}{\partial \theta} \log f_K(\tilde{y}_t | \tilde{x}_{t-1}, \tilde{\theta}_n)\right]'$$

Otherwise a HAC estimator of the variance must be used.

Q: Which density for QMLE?

A: SNP (modified Hermite expansion)

Location-scale transform of an innovation z

$$y = R_x z + \mu_x$$

Innovation density

$$h_K(z|x) = \frac{[\mathcal{P}(z, x)]^2 \phi(z)}{\int [\mathcal{P}(u, x)]^2 \phi(u)\, du}$$

$\mathcal{P}(z, x)$ is a polynomial in z of degree K_z whose coefficients are polynomials of degree K_x in x.

Location function

$$\mu_x = b_0 + B x_{t-1},$$

ARCH scale function

$$\text{vech}(R_{x_{t-1}}) = \rho_0 + \sum_{i=1}^{L_r} P_{(i)} |y_{t-1-L_r+i} - \mu_{x_{t-2-L_r+i}}|$$

GARCH scale function

$$\text{vech}(R_{x_{t-1}}) = \rho_0 + \sum_{i=1}^{L_r} P_{(i)} |y_{t-1-L_r+i} - \mu_{x_{t-2-L_r+i}}| + \sum_{i=1}^{L_g} \text{diag}(G_{(i)}) R_{x_{t-2-L_g+i}}$$

EMM Estimating Equations Under SNP

$$m(\rho,\tilde{\theta}_n) = \mathcal{E}_\rho\Big[\frac{\partial}{\partial\theta}\log f_K(y_0|y_{-L},\ldots,y_{-1},\tilde{\theta}_n)\Big]$$

where ρ denotes the parameters of the putative model $p(y|x,\rho)$.

The expectation is computed by (1) fixing ρ, (2) simulating the model $p(y|x,\rho)$ model, and (3) averaging

$$\mathcal{E}_\rho\Big[\frac{\partial}{\partial\theta}\log f_K(y_t|y_{t-L},\ldots,y_{t-1},\theta)\Big] = \frac{1}{N}\sum_{t=1}^{N}\frac{\partial}{\partial\theta}\log f_K(\hat{y}_t|\hat{y}_{t-L},\ldots,\hat{y}_{t-1},\theta)$$

The EMM estimator $\hat{\rho}$ is the solution of the estimating equations

$$m(\rho,\tilde{\theta}_n) = 0$$

in the following sense

EMM Estimator

$$\hat{\rho}_n = \underset{\rho}{\text{argmin}}\; m'(\rho,\tilde{\theta}_n)(\tilde{\mathcal{I}}_n)^{-1}m(\rho,\tilde{\theta}_n)$$

Recall

$$m(\rho,\theta) = \mathcal{E}_\rho\left\{\frac{\partial}{\partial\theta}\log[f_K(y_0|y_{-L},\ldots,y_{-1},\theta)]\right\}$$

$$\tilde{\theta}_n = \underset{\theta}{\text{argmax}}\,\frac{1}{n}\sum_{t=0}^{n}\log[f_K(\tilde{y}_t|\tilde{y}_{t-L},\ldots,\tilde{y}_{t-1},\theta)]$$

$$\tilde{\mathcal{I}}_n = \frac{1}{n}\sum_{t=1}^{n}\Big[\frac{\partial}{\partial\theta}\log f_K(\tilde{y}_t|\tilde{x}_{t-1},\tilde{\theta}_n)\Big]\Big[\frac{\partial}{\partial\theta}\log f_K(\tilde{y}_t|\tilde{x}_{t-1},\tilde{\theta}_n)\Big]'$$

Why not do this instead? (method of moments)

- Set forth a moment function, such as

$$\tilde{\psi}_c(y_{-L},\ldots,y_{-1},y_0) = \begin{pmatrix} y_0 & - & \tilde{\mu}_1 \\ (y_0)^2 & - & \tilde{\mu}_2 \\ & \vdots & \\ (y_0)^K & - & \tilde{\mu}_k \\ y_{-1}\,y_0 & - & \tilde{\gamma}(1) \\ y_{-2}\,y_0 & - & \tilde{\gamma}(2) \\ & \vdots & \\ y_{-L}\,y_0 & - & \tilde{\gamma}(L) \end{pmatrix}$$

  where

$$\tilde{\mu}_k = \frac{1}{n}\sum_{t=1}^{n}(\tilde{y}_t)^k, \quad \tilde{\gamma}(\ell) = \frac{1}{n}\sum_{t=1}^{n}\tilde{y}_{t-\ell}\tilde{y}_t$$

- Compute the moment equations

$$m_n(\rho) = \mathcal{E}_\rho(\tilde{\psi}_c)$$

- Attempt to solve the estimating equations

$$m_n(\rho) = 0$$

  for the system parameters ρ.

Minimum Chi-squared

If the equations $m_n(\rho) = 0$ cannot be solved because the dimension of ψ is larger than the dimension of ρ, then

- Use a nonlinear optimizer to minimize:

$$\hat{\rho}_n = \underset{\rho}{\text{argmin}}\; m_n'(\rho)\left(\tilde{\mathcal{I}}_n\right)^{-1}m_n(\rho);$$

- $\tilde{\mathcal{I}}_n$ is an estimate of the variance of $\sqrt{n}\,m_n'(\rho)$

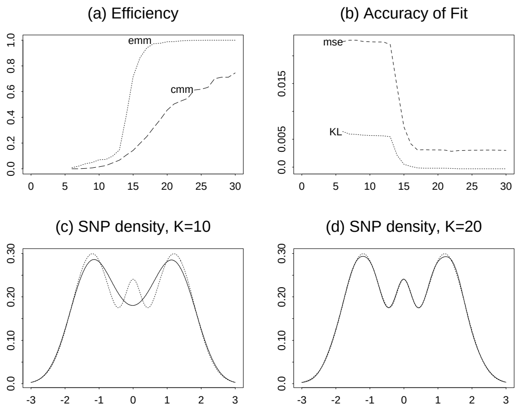
$$\tilde{\mathcal{I}}_n = \sum_{\tau=-[n^{1/5}]}^{[n^{1/5}]} w\left(\frac{\tau}{[n^{1/5}]}\right)\tilde{\mathcal{I}}_{n\tau}$$

  where

$$w(u) = \begin{cases} 1-6|u|^2+6|u|^3 & \text{if } 0 < u < \frac{1}{2} \\ 2(1-|u|)^3 & \text{if } \frac{1}{2} \le u < 1 \end{cases}$$

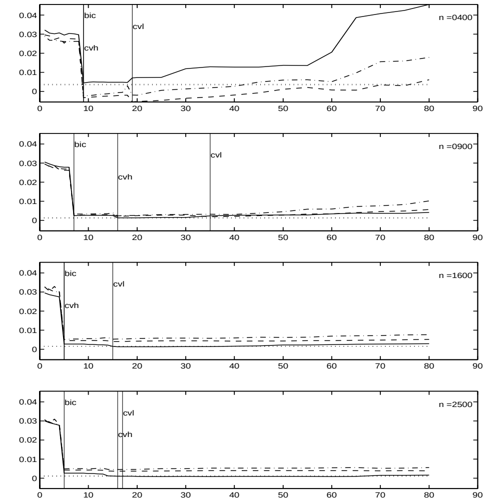
$$\tilde{\mathcal{I}}_{n\tau} = \begin{cases} \frac{1}{n}\sum_{t=1+\tau}^{n}\tilde{\psi}_c(\tilde{y}_{t-L},\ldots,\tilde{y}_t)\tilde{\psi}_c'(\tilde{y}_{t-L-\tau},\ldots,\tilde{y}_{t-\tau}) & \tau \ge 0 \\ \tilde{\mathcal{I}}_{n,-\tau} & \tau < 0 \end{cases}$$

Relative Efficiency for the Trimodal Density



- (a) $10 \le K \le 20$: EMM efficiency increases rapidly
- (b) $10 \le K \le 20$: SNP approximation error decreases rapidly
- (c) $K = 10$: SNP approximates a trimodal density by a bimodal density
- (d) $K = 20$: SNP correctly determines the number of modes.

Cross-Validation for the Trimodal Density



Mean squared error (MSE) and cross validated estimate (CV) for a realization of size n from a trimodal density. Solid line is MSE, dashed line is leave-one-out CV estimate (CVL), and dashed-dotted line is average of ten 10% hold-out-sample CV estimates (CVH). Upper horizontal line is MSE achieved by a cross-validated kernel estimate and lower is best kernel MSE. Vertical lines are BIC, CVL, and CVH choices of K.

Outline

- Dynamical Systems with Unobserved States
  - Motivation
  - Example: A diffusion model of equity returns
  - Simulation methods for diffusions
- Simulated Method of Moments
  - Minimum chi-squared estimators
  - Efficiency considerations
  - Data determined bandwidth
- Methodology
  - Projection
  - Estimation
  - Reprojection

Projection (data summary)

Use SNP density for observables

$$y_t \sim f_K(y|x,\theta)$$

which is described by

Location function: μ_x
Scale function: R_x
Hermite function: $P(z,x)\phi(z)$

with tuning parameters

$$K = (L_u, L_g, L_r, L_p, K_z, K_x)$$

to summarize the data

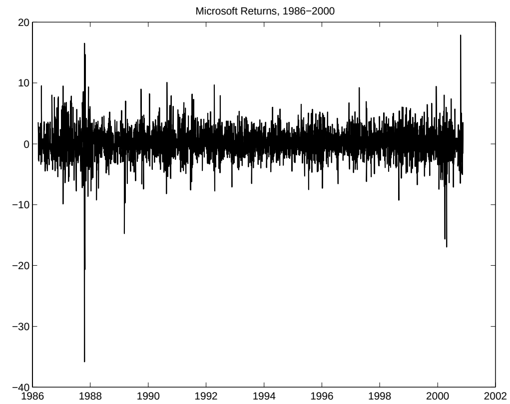
$$\tilde{\theta}_n = \underset{\theta}{\text{argmax}} \frac{1}{n} \sum_{t=0}^{n} \log[f_K(\tilde{y}_t|\tilde{x}_{t-1},\theta)]$$

$$\tilde{\mathcal{I}}_n = \frac{1}{n}\sum_{t=1}^{n}[\frac{\partial}{\partial\theta}\log f_K(y_t|x_{t-1},\tilde{\theta}_n)][\frac{\partial}{\partial\theta}\log f_K(y_t|x_{t-1},\tilde{\theta}_n)]'$$

with BIC used to determine

$$K = (L_u, L_g, L_r, L_p, K_z, K_x)$$

Microsoft Returns, 1985–2000



Data characteristics
- 3,711 daily observations.
- Mean 0.158, Std. Dev. 2.502
- Min -35.828, Median 0.00, Max 17.869, IQ range 2.77
- Annualized Mean 39.868, Std. Dev. 39.720

Projection (data summary)
- $K = (L_u, L_g, L_r, L_z, K_z, K_x) = (1, 1, 1, 1, 6, 0)$.
- $\dim(\theta) = 11$

Microsoft Data

The BIC preferred SNP model for Microsoft returns has tuning parameters

$$K = (L_u, L_g, L_r, L_p, K_z, K_x) = (1, 1, 1, 1, 6, 0)$$

which is a AR(1) with GARCH(1,1) volatility and a Hermite innovation density of degree 6.

Quite similar to a GARCH(1,1) with innovations that follow Student's t-distribution with small degrees freedom.

This total number of parameters is 11 which gives a saturation ratio of 3711/11=377 observations per parameter.

Estimation (parameter determination)

EMM Estimator

$$\hat{\rho}_n = \underset{\rho}{\operatorname{argmin}}\; m_n'(\rho, \tilde{\theta}_n)(\tilde{\mathcal{I}}_n)^{-1} m_n(\rho, \tilde{\theta}_n)$$

Recall

$$\begin{aligned} m_n(\rho, \tilde{\theta}_n) &= \mathcal{E}_\rho \frac{\partial}{\partial\theta} \log[f_K(y_0|y_{-L}, \ldots, y_{-1}, \tilde{\theta}_n)] \\ \tilde{\theta}_n &= \underset{\theta}{\operatorname{argmax}} \frac{1}{n}\sum_{t=0}^{n} \log[f_K(\tilde{y}_t|\tilde{y}_{t-L}, \ldots, \tilde{y}_{t-1}, \theta)] \\ \tilde{\mathcal{I}}_n &= \frac{1}{n}\sum_{t=1}^{n} \left[\frac{\partial}{\partial\theta}\log f_K(\tilde{y}_t|\tilde{x}_{t-1}, \tilde{\theta}_n)\right]\left[\frac{\partial}{\partial\theta}\log f_K(\tilde{y}_t|\tilde{x}_{t-1}, \tilde{\theta}_n)\right]' \end{aligned}$$

An unconditional expectation can be computed by generating a long simulation $\{\hat{y}_t\}_{t=-L}^{N}$ and averaging

$$m_n(\rho, \tilde{\theta}_n) \doteq \frac{1}{N}\sum_{t=0}^{N} \log[f_K(\tilde{y}_t|\hat{y}_{t-L}, \ldots, \hat{y}_{t-1}, \theta)]$$

For instance, if the Microsoft data is presumed to satisfy the the stochastic volatility model

$$\begin{aligned} dU_{1t} &= \alpha_{10}\, dt + \exp(\beta_{10} + \beta_{12}U_{2t} + \beta_{13}U_{3t})\, dW_{1t} \\ dU_{2t} &= \alpha_{22}U_{2t}dt + dW_{2t} \\ dU_{3t} &= \alpha_{33}U_{3t}dt + dW_{3t}, \end{aligned}$$

then, given values for the 6 parameters

$$\rho = (\alpha_{10}, \alpha_{22}, \alpha_{33}, \beta_{10}, \beta_{12}, \beta_{13}),$$

one can simulate the model and retain

$$\hat{y}_t = 100(U_{1,t} - U_{1,t-1})$$

$$\hat{x}_{t-1} = (y_{1,t-L}, \ldots, y_{1,t-1})$$

for integer values of time, $t = 1, \ldots, N$, to get daily and lagged daily returns for computing

$$m_n(\rho, \tilde{\theta}_n) \doteq \frac{1}{N}\sum_{t=0}^{N} \log[f_K(\hat{y}_t|\hat{x}_{t-1}, \theta)]$$

Asymptotics

If ρ^o denotes the true value of ρ and θ^o is an isolated solution of

$$\mathcal{E}_{\rho^o}\left[\frac{\partial}{\partial\theta}\log f_K(y|x,\theta)\right]=0,$$

then

$$\lim_{n\to\infty}\hat{\rho}_n=\rho^o \quad \text{a.s.}$$

$$\sqrt{n}(\hat{\rho}_n-\rho^o)\xrightarrow{\mathcal{L}} N\{0,[(M^o)'(\mathcal{I}^o)^{-1}(M^o)]^{-1}\}$$

$$\lim_{n\to\infty}\hat{M}_n=M^o \quad \text{a.s.}$$

$$\lim_{n\to\infty}\tilde{\mathcal{I}}_n=\mathcal{I}^o \quad \text{a.s.}$$

where

$$\hat{M}_n=M(\hat{\rho}_n,\tilde{\theta}_n)$$

$$M^o=M(\rho^o,\theta^o)$$

$$M(\rho,\theta)=(\partial/\partial\rho')\,\mathcal{E}_\rho\left[\frac{\partial}{\partial\theta}\log f_K(y|x,\theta)\right]$$

$$\mathcal{I}^o=\mathcal{E}_{\rho^o}\left[\frac{\partial}{\partial\theta}\log f_K(y|x,\theta^o)\right]\left[\frac{\partial}{\partial\theta}\log f_K(y|x,\theta^o)\right]'.$$

Diagnostics

Under the null hypothesis that $p(y_{-L},\ldots,y_0|\rho)$ is the correct model,

$$L_0=\;n\,m'(\hat{\rho}_n)(\tilde{\mathcal{I}}_n)^{-1}m(\hat{\rho}_n)$$

is asymptotically chi-squared on dim(θ)−dim(ρ) degrees of freedom.

Because

$$\sqrt{n}\,m(\hat{\rho}_n)\xrightarrow{\mathcal{L}} N\{0,\mathcal{I}^o-(M^o)[(M^o)'(\mathcal{I}^o)^{-1}(M^o)]^{-1}(M^o)'\},$$

inspection of the t-ratios

$$T_n=S_n^{-1}\sqrt{n}\,m(\hat{\rho}_n),$$

where

$$S_n=(\text{diag}\{\tilde{\mathcal{I}}_n-(\hat{M}_n)[(\hat{M}_n)'(\tilde{\mathcal{I}}_n)^{-1}(\hat{M}_n)]^{-1}(\hat{M}_n)'\})^{1/2}$$

can suggest reasons for model failure. Different elements of the score correspond to different characteristics of the data. Large t-ratios reveal those characteristics that are not well approximated.

EMM Objective Function Surface **MSFT**

Model	N	$\chi^2(\hat{\rho})$	df	p-value
LLV1	50k	18.348	7	0.010
LLV1	75k	17.888	7	0.012
LLV1	100k	18.189	7	0.011
LLV2	50k	9.978	5	0.076
LLV2	75k	9.254	5	0.099
LLV2	100k	8.642	5	0.124

Notes:
The LLV1 model has one volatility factor; the LLV2 model has two volatility factors. 100k denotes a simulation of length $N=100000$ simulated at $1/\Delta=24*252=6048$ steps per year.

Parameter Estimates **MSFT**

LLV1: One volatility factor

	Est	SE
α_{10}	0.4215	0.0638
α_{22}	-12.3711	2.2803
β_{10}	-1.1441	0.0384
β_{12}	1.4656	0.0566

LLV2: Two volatility factors

	Est	SE
α_{10}	0.4247	0.0874
α_{22}	-0.000861	0.001633
α_{33}	-102.9206	21.2190
β_{10}	-0.6759	0.1489
β_{13}	0.0371	0.0027
β_{14}	5.0979	0.2444

Diagnostic t-Statistics.

Coefficient		LLV1	LLV2
Location Function:			
b_0	psi(1)	0.424	-1.120
b_1	psi(2)	1.244	1.420
Scale Function:			
ρ_0	tau(1)	2.564	0.503
ρ_1	tau(2)	1.956	1.024
g_1	tau(3)	2.318	0.722
Hermite Polynomial:			
$a_{0,1}$	A(2)	0.381	0.563
$a_{0,2}$	A(3)	3.390	-1.013
$a_{0,3}$	A(4)	0.456	0.660
$a_{0,4}$	A(5)	3.570	-0.734
$a_{0,5}$	A(6)	0.203	0.639
$a_{0,6}$	A(7)	3.099	-0.830

Reprojection (post estimation analysis)

Once model parameter estimates are in hand the model can be simulated at these values.

For instance, for the Microsoft data one can generate a large simulation of the values of

$$(y_t, x_{t-1}, U_{2t}, U_{3t})$$

or even integrated volatility

$$\int_t^{t+T} \exp(\beta_{10} + \beta_{12} U_{2t} + \beta_{13} U_{3t})\, dt)$$

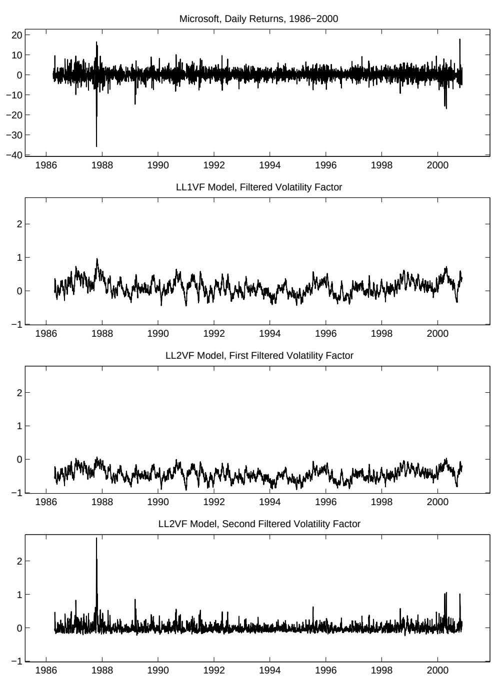
at any desired sampling frequency.

Using statistical methods, e.g., regression or density estimation, functional relationships such as

$$U_2 = g(y, x)$$

can be determined. Such a relationship can then be applied to the data to accomplish, e.g., filtering

$$U_2 = g(\tilde{y}_t, \tilde{x}_{t-1})$$



Outline

- Dynamical Systems with Unobserved States
  - Motivation
  - Example: A diffusion model of equity returns
  - Simulation methods for diffusions
- Simulated Method of Moments
  - Minimum chi-squared estimators
  - Efficiency considerations
  - Data determined bandwidth
- Methodology
  - Projection
  - Estimation
  - Reprojection