A Zero-Augmented Multiplicative Error Model

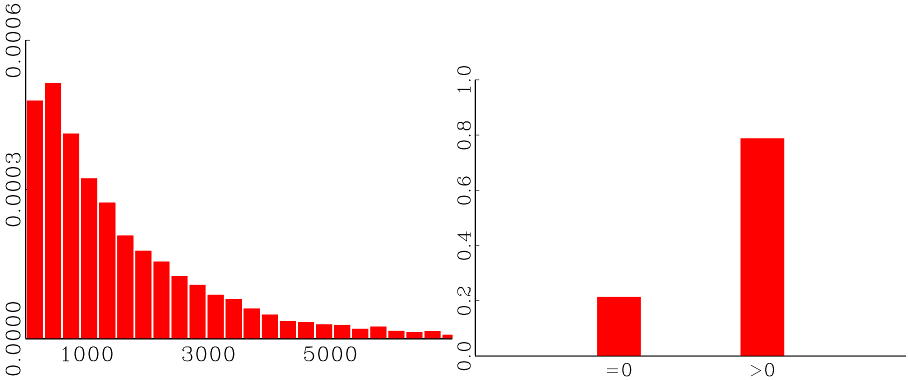
Nikolaus Hautsch

Peter Malec

Melanie Schienle

Chair of Econometrics
Institute for Statistics and Econometrics
Humboldt-Universität zu Berlin

Figure: Distribution of 15-Second-Trading Volumes of Boeing (NYSE)



(a) Histogram of nonzero volumes (b) Frequencies of zero/nonzero volumes

Sample Period: January 2006

The Multiplicative Error Model

- Define:
  - x_t: a positive valued process.
  - $\mathcal{F}_{t-1}$: information process up to time $t-1$.
- MEM for x_t (Engle, 2002):

$$x_t = \mu_t \, \varepsilon_t \,,$$

where:

$$\mu_t = E[x_t | \mathcal{F}_{t-1}] = \omega + \sum_{i=1}^{p} \alpha_i \, x_{t-i} + \sum_{i=1}^{q} \beta_i \, \mu_{t-i},$$

$$\varepsilon_t | \mathcal{F}_{t-1} \sim i.i.d. \; D\left(1, \sigma_\varepsilon^2\right) .$$

- In many cases: distribution of $\varepsilon_t | \mathcal{F}_{t-1}$ not defined for $\varepsilon_t = 0$.
  $\Rightarrow x_t = 0$ cannot be explained by the model!

Exact Zeros

- Common approach in the MEM-literature: removal of zeros.
  - Implies loss of information.
  - Absence of trades important in market microstructure theory: trading intensity has informational content (e.g. Easly & O'Hara, 1992).
- Models for continuous nonnegative data with excess zeros: Mainly applied in other fields of research (e.g. biometrics, environmetrics).
- Aim of this project:
  Develop a substantiated and parsimonious econometric approach for dealing with zero values in a MEM framework.

Outline

Generalized F-MEM

- ε_t follows Generalized F distribution:

$$f_\varepsilon(\varepsilon_t) = \frac{a\,\varepsilon_t^{a\,m-1}\,[\eta + (\varepsilon_t/\lambda)^a]^{(-\eta-m)}\,\eta^\eta}{\lambda^{a\,m}\,\mathcal{B}(m,\eta)}; \qquad \lambda = 1.$$

- Extended MEM structure:

$$x_t = \varkappa(\mu_t)\,\varepsilon_t,$$

where:

$$\varepsilon_t|\mathcal{F}_{t-1} \sim i.i.d.\ D\left(\xi, \sigma_\varepsilon^2\right) \qquad \varkappa(\mu_t) = \mu_t/\xi,$$

$$\xi = \eta^{1/a}\frac{\Gamma(m+1/a)\,\Gamma(\eta-1/a)}{\Gamma(m)\,\Gamma(\eta)}; \quad a\,\eta > 1.$$

- Log-likelihood function:

$$\mathcal{L} = \sum_{t=1}^{n} \log a + (am-1)\log x_t - (\eta + m)\log\{\eta + [x_t/\varkappa(\mu_t)]^a\}$$
$$+ \eta \log \eta - \log \mathcal{B}(m,\eta) - am \log \varkappa(\mu_t).$$

- Generalized F nests different error distributions:
  - $\eta \to \infty$: Generalized Gamma
  - $m = 1, \eta \to \infty$: Weibull
  - $m = 1, \eta = 1$: Log-Logistic
- Drawback: Distribution not defined for $\varepsilon_t = 0$ when $am < 1$. Log-Likelihood cannot be evaluated for $x_t = 0$ ($\log x_t$).

A Zero-Augmented Error-Distribution

- Idea: Place nonzero probability mass on the zero value,

$$P\{\varepsilon_t > 0\} = p, \quad P\{\varepsilon_t = 0\} = 1 - p.$$

- $\varepsilon_t | \varepsilon_t > 0$ follows a continuous distribution with p.d.f. $g_\varepsilon(\varepsilon_t)$:
  - continuous for $\varepsilon_t \in (0, \infty)$.
  - lower bound at zero.
- Distribution of ε_t is discontinuous with discontinuity at zero:

$$f_\varepsilon(\varepsilon_t) = (1 - p)\,\delta(\varepsilon_t) + p\; g_\varepsilon(\varepsilon_t)\, 1\!\!1_{(\varepsilon_t > 0)},$$

$\delta(\varepsilon_t)$: point probability mass at $\varepsilon_t = 0$,

$1\!\!1_{(\varepsilon_t > 0)}$: indicator function.

Application to the Generalized F-MEM

- $g_\varepsilon(\varepsilon_t)$ is given by the Generalized F distribution:

$$g_\varepsilon(\varepsilon_t) = \frac{a\,\varepsilon_t^{a\,m-1}\,[\eta + (\varepsilon_t/\lambda)^a]^{(-\eta-m)}\,\eta^\eta}{\lambda^{a\,m}\,\mathcal{B}(m,\eta)}; \qquad \varepsilon_t > 0,$$

- Hence:

$$E[\varepsilon_t|\varepsilon_t > 0] = \mu_+ = \lambda\,\xi.$$

- The condition $E[x_t|\mathcal{F}_{t-1}] = \mu_t$ requires that $E[\varepsilon_t] = \xi$:

$$E[\varepsilon_t] = p\,\mu_+ \overset{!}{=} \xi,$$
$$\lambda = \frac{1}{p}.$$

ZAGF-MEM

- Zero-augmented MEM structure:

$$x_t = \varkappa(\mu_t)\,\varepsilon_t,$$

where:

$$\varepsilon_t \sim \text{ZAGF}(\lambda, a, m, \eta, p)\,,$$
$$\varkappa(\mu_t) = \mu_t/\xi,$$
$$\xi = E[\varepsilon_t]\,.$$

- Conditional density of the endogenous variable x_t:

$$f_x(x_t|\mathcal{F}_{t-1}) = (1-p)\,\delta(x_t) + p\; g_\varepsilon[x_t/\varkappa(\mu_t)]\;\; [1/\varkappa(\mu_t)]\,\mathbb{1}_{(x_t>0)}.$$

- Log-likelihood:

$$\mathcal{L} = n_z \log(1-p) + n_{nz}\log p + \sum_{t,nz}\{\log g_\varepsilon[x_t/\varkappa(\mu_t)] - \log\varkappa(\mu_t)\}\,.$$

A Semiparametric Benchmark

- Estimation of θ in $\mu(\mathcal{F}_{t-1};\theta)$ without explicit assumptions regarding $f_\varepsilon(\varepsilon_t)$.
  $\Rightarrow$ Quasi-Likelihood function:
  $$\mathcal{Q} = -\sum_{t,n}\left[\frac{x_t}{\mu_t} + \log\mu_t\right].$$
  $\Rightarrow$ Drost & Werker (2004): Provides consistent estimate of θ if $\mu(\mathcal{F}_{t-1};\theta)$ is correctly specified.
- Semiparametric estimation of $f_\varepsilon(\varepsilon)$ based on $\widehat{\varepsilon}_t = \frac{x_t}{\widehat{\mu}_t}$:
  - $\widehat{g}_\varepsilon(\varepsilon)$: kernel density estimation using all $\widehat{\varepsilon}|\widehat{\varepsilon} > 0$,
  - $\widehat{p} = n^{-1}\sum_t \mathbb{1}_{(\widehat{\varepsilon}_t>0)}$ (rel. frequency of $\widehat{\varepsilon} > 0$),

$$\Rightarrow \qquad \widehat{f}_\varepsilon(\varepsilon) = (1-\widehat{p})\,\delta(\varepsilon) + \widehat{p}\,\widehat{g}_\varepsilon(\varepsilon)\,\mathbb{1}_{(\varepsilon_t>0)}.$$

KDE Approach

- Kernel density estimator of $g_\varepsilon(\varepsilon)$:

$$\widehat{g}_\varepsilon(\varepsilon) = \frac{1}{nb}\sum_{t=1}^{n} K\left(\frac{\varepsilon - \varepsilon_t}{b}\right),$$

  $K(z)$: symmetric pdf with $\int z\,K(z)dz = 0$ and $\int z^2 K(z)dz < \infty$,
  b: bandwidth satisfying $b \to 0$ and $nb \to \infty$ as $n \to \infty$.
- Support of $g_\varepsilon(\varepsilon)$ is bounded from below: boundary bias!
  $\Rightarrow$ Possible remedies:
  - Reflection method (Schuster, 1985).
  - Cut-and-Normalized kernels (Gasser & Müller, 1979).
  - Boundary kernel (Gasser & Müller, 1979).
  - Local linear method (Cheng et al., 1997).
  - ...

Asymmetric Kernels

- Gamma kernel estimator (Chen, 2000):

$$\widehat{g}_{\varepsilon}(\varepsilon) = \frac{1}{n} \sum_{t=1}^{n} K_{\rho_b(\varepsilon),b}(\varepsilon_t) \,,$$

where

$$K_{\rho_b(\varepsilon),b}\,(u) = \frac{u^{\rho_b(\varepsilon)-1} \exp(-u/b)}{b^{\rho_b(\varepsilon)}\,\Gamma(\rho_b)},$$

and

$$\rho_b\,(\varepsilon) = \begin{cases} \varepsilon/b & \text{if } \varepsilon \geq 2b \\ \frac{1}{4}\,(\varepsilon/b)^2 + 1 & \text{if } \varepsilon \in (0, 2b)\,. \end{cases}$$

- Potential bias problem in the interior of the density support.

Multiplicative Bias Correction (MBC)

- Semiparametric density structure (Hjort & Glad, 1995):

$$g_\varepsilon(\varepsilon) = g_\varepsilon(\varepsilon, \vartheta)\, r(\varepsilon)\,,$$

  $g_\varepsilon(\varepsilon, \vartheta)$: parametric "start",
  $r(\varepsilon)$: correction function.
- Idea: estimate $r(\varepsilon) = g_\varepsilon(\varepsilon)\,/g_\varepsilon(\varepsilon, \vartheta)$ via kernel smoothing, i.e.

$$\widehat{r}(\varepsilon) = \frac{1}{n}\sum_{t=1}^{n} K_{\rho_b(\varepsilon),b}(\varepsilon_t)\,/g_\varepsilon(\varepsilon_t, \widehat{\vartheta}).$$

$\Rightarrow$ Bias-corrected gamma KDE (Hagmann & Scaillet, 2007):

$$\widetilde{g}_\varepsilon(\varepsilon) = \frac{1}{n}\sum_{t=1}^{n} K_{\rho_b(\varepsilon),b}(\varepsilon_t)\;\frac{g_\varepsilon(\varepsilon, \widehat{\vartheta})}{g_\varepsilon(\varepsilon_t, \widehat{\vartheta})}.$$

Specification of the Conditional Mean

- Basic linear MEM(p, q):

$$\mu_t = \omega + \sum_{i=1}^{p} \alpha_i \, x_{t-i} + \sum_{i=1}^{q} \beta_i \, \mu_{t-i}; \qquad \omega, \alpha_i, \beta_i \geq 0.$$

- Log-MEM (Bauwens & Giot, 2000) makes parameter restrictions unnecessary:

$$\log \mu_t = \omega + \sum_{i=1}^{p} \alpha_i \, \log x_{t-i} \, \mathbb{1}_{(x_{t-i}>0)} + \sum_{i=1}^{p} \alpha_i^0 \, \mathbb{1}_{(x_{t-i}=0)} + \sum_{i=1}^{q} \beta_i \, \log \mu_{t-i}.$$

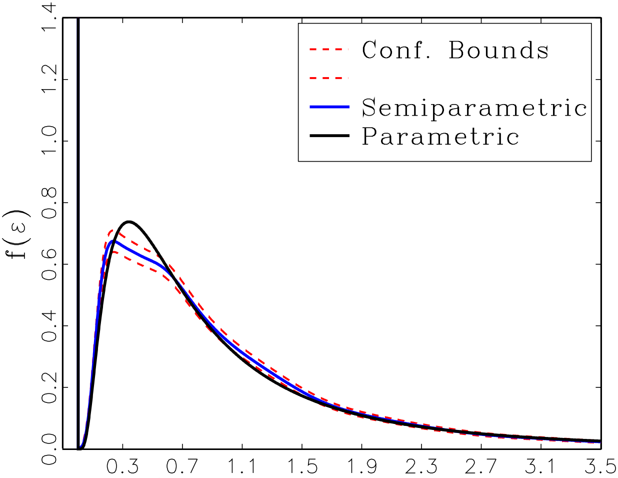
- Specification search provides p and q.

Data

- Dataset (original): tick data for AIG, Boeing, IBM, Merck (NYSE).
- Period: January 2-18, 2006.
- Frequency (aggregated): 15 seconds.
- Observations: 15829.
- Variable: average trading volume.

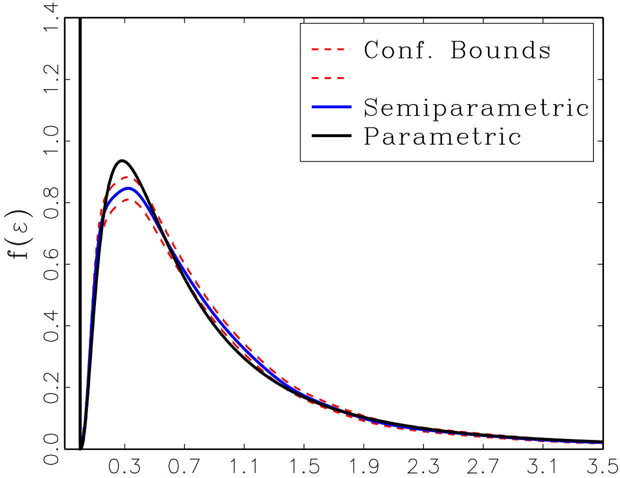
	AIG	**Boeing**	**IBM**	**Merck**
% Zeros	0.114	0.188	0.086	0.153

Figure: Estimated Density of ε_t (Boeing, Gamma Kernel)



$\widehat{p}_{ML} = 0.812 \qquad \widehat{p} = 0.811$

Figure: Estimated Density of ε_t (IBM, Gamma Kernel)



$\widehat{p}_{ML} = 0.913 \qquad \widehat{p} = 0.914$

Figure: Estimated Density of ε_t (Boeing, Gamma Kernel with MBC)

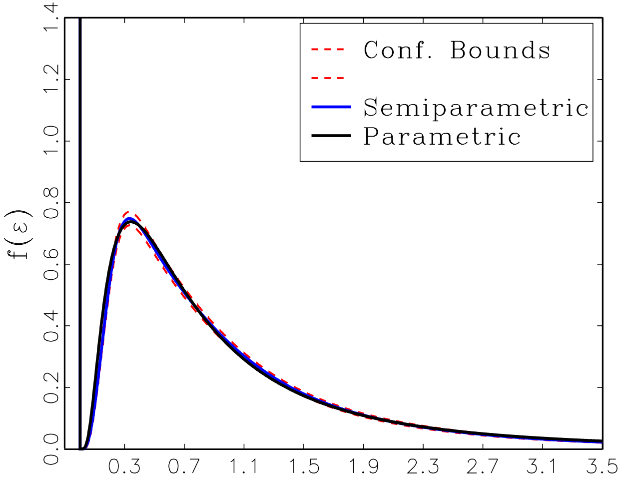
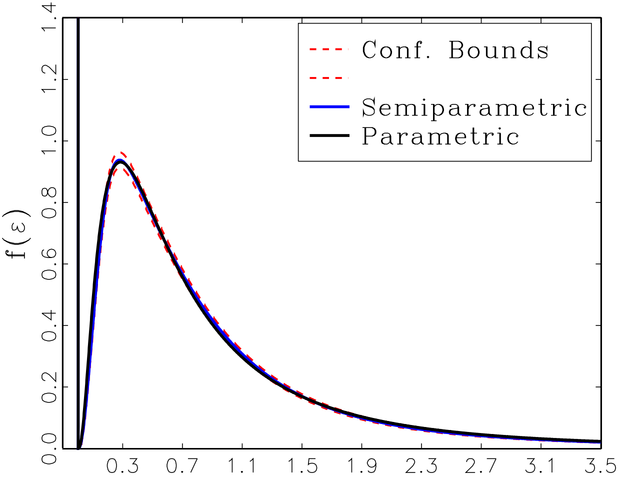


Figure: Estimated Density of ε_t (IBM, Gamma Kernel with MBC)



Semiparametric Specification Test

- We compare:

$$f_\varepsilon(\varepsilon) = (1-p)\,\delta(\varepsilon) + p\; g_\varepsilon(\varepsilon)\,\mathbb{1}_{(\varepsilon>0)} \quad \text{(semiparametric)}$$

$$f_\varepsilon(\varepsilon,\vartheta) = (1-p)\,\delta(\varepsilon) + p\; g_\varepsilon(\varepsilon,\vartheta)\,\mathbb{1}_{(\varepsilon>0)} \quad \text{(parametric).}$$

- H_0: $f_\varepsilon(\varepsilon) = f_\varepsilon(\varepsilon,\vartheta)$ for some $\vartheta \in \Theta$.
- Idea: If H_0 holds, then $r(\varepsilon) = g(\varepsilon)\,/g(\varepsilon,\vartheta) = 1$.
- Test statistic:

$$T_n = \widehat{p}\,n\,b^{1/2}\int_0^\infty \psi(\varepsilon)\,[\widehat{r}(\varepsilon)-1]^2\,d\varepsilon.$$

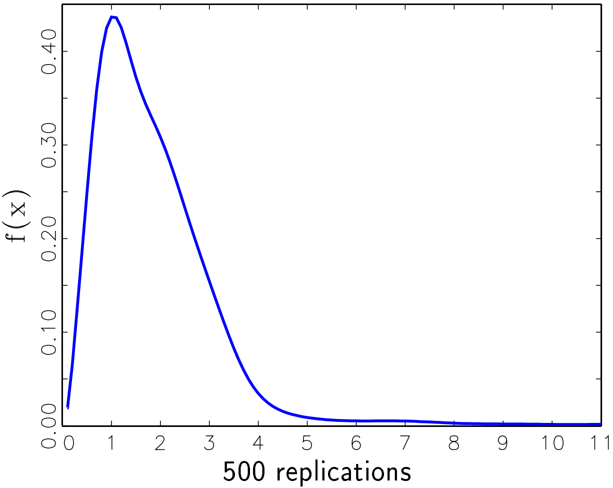
- Fernandes & Monteiro (2005): T_n is asymptotically normal under H_0.
- BUT: size distortions play an important role (Fan, 1995).
  $\Rightarrow$ Fan (1998): Parametric bootstrap approach.

Parametric Bootstrap

1. Draw random sample $\{\varepsilon_t^*\}_{t=1}^n$ from distribution with density $f_\varepsilon(\varepsilon, \widehat{\vartheta})$.
2. Evaluate $\{x_t^*\}_{t=1}^n$ as $x_t^* = \mu\left(\mathcal{F}_{t-1}^*; \widehat{\theta}\right) \varepsilon_t^* / \xi$.
3. Use $\{x_t^*\}_{t=1}^n$ to compute compute statistic T_n^* (involves re-estimation of parametric model and semiparametric benchmark).
4. Repeat steps 1,2 and 3 B times to obtain bootstrap distribution of $\left\{T_{n,r}^*\right\}_{r=1}^B$.
5. Calculate $(1-\alpha)$-quantile of bootstrap distribution (C_α).

   $\Rightarrow$ Reject H_0 at significance level α if $T_n > C_\alpha$.

Figure: Bootstrap Distribution of Semiparametric Specification Test (IBM)



Extension by a Binary Choice Model

- Probability p is time-varying:

$$p_t = P\left(\varepsilon_t > 0 | \mathcal{F}_{t-1}\right) = P\left(x_t > 0 | \mathcal{F}_{t-1}\right).$$

- Binary choice model for trade indicator:

$$p_t = P\left(\mathcal{I}_t = 1 | \mathcal{F}_{t-1}\right) = p(\mathcal{F}_{t-1}; \pi),$$

where:

$$\mathcal{I}_t = \begin{cases} 1 & \text{if } x_t > 0 \\ 0 & \text{if } x_t = 0. \end{cases}$$

- Implications for model properties:
  - $\varepsilon_t | \varepsilon_t > 0$ follows Generalized F distribution with time-varying scale parameter: $\lambda_t = 1/p_t$.
  - Error process $\{\varepsilon_t\}$ loses i.i.d. property:
    $\Rightarrow P\left(\varepsilon_t \leq \varepsilon | \mathcal{F}_{t-1}\right)$ depends on $\mathcal{H}_{t-1} \subseteq \mathcal{F}_{t-1}$
    $\Rightarrow \{\varepsilon_t | \mathcal{H}_{t-1}\}$ is i.n.i.d. (Drost & Werker, 2004).

Specification of p_t

- Two-State ACM model (Russel & Engle, 2004):

$$p_t = P\left(\mathcal{I}_t = 1 | \mathcal{H}_{t-1}\right) = \frac{exp(h_t)}{1 + exp(h_t)},$$

where

$$h_t = \gamma + \sum_{j=1}^{v} \eta_j \, \frac{\mathcal{I}_{t-j} - p_{t-j}}{\sqrt{p_{t-j}\left(1 - p_{t-j}\right)}} + \sum_{j=1}^{w} \zeta_j \, h_t.$$

- Note: $I_t | H_t \sim Be\left(p_t\right) \Rightarrow E\left[I_t | H_t\right] = p_t$.

Table: Estimation Results - DZAGF-MEM - ACM Part

	Boeing			**IBM**		
	Coef.	T-Stat.	P-Val.	Coef.	T-Stat.	P-Val.
γ	0.008	4.307	0.000	0.016	4.438	0.000
η_1	0.169	10.128	0.000	0.156	6.895	0.000
η_2	-0.114	-6.814	0.000	-0.083	-3.640	0.000
ζ_1	0.995	909.616	0.000	0.994	705.638	0.000

Density Forecast Evaluation

- Does not require assumptions regarding DGP.
- Idea: Check if one-step-ahead density forecasts are correct,

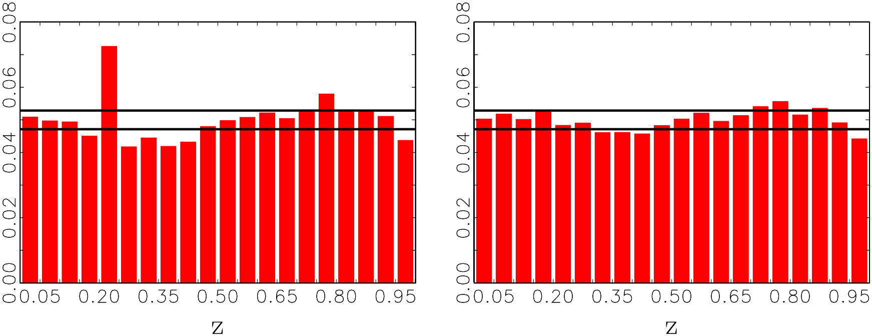
$$\{q_t(x_t|\mathcal{F}_{t-1})\}_{t=1}^n = \{f_t(x_t|\mathcal{F}_{t-1})\}_{t=1}^n .$$

- Method: Compute probability integral transforms (PIT)

$$z_t = \int_{-\infty}^{x_t} q_t(u)\, du = Q_t(x_t) .$$

- Diebold et al. (1998): If the model is correctly specified, $\{z_t\}_{t=1}^n$ is i.i.d. $U(0,1)$.

Figure: Histogram of Z (Boeing)



(a) ZAGF-MEM

(b) DZAGF-MEM

Table: χ^2-Test for Uniformity

	Boeing		IBM	
	χ^2	P-Val.	χ^2	P-Val
ZAGF-MEM	275.357	0.000	105.033	0.000
DZAGF-MEM	53.465	0.000	46.207	0.001

- Conclusions:
  - Approach allows MEM analysis of data with "many" zeros.
  - Extension by binary choice component (ACM model) improves performance.
- Agenda:
  - Test of parametric density against semiparametric benchmark: final implementation.
  - Simple economic application (market microstructure).

Figure: Estimated Density of ε_t (IBM, Epanechnikov Kernel with Reflection)

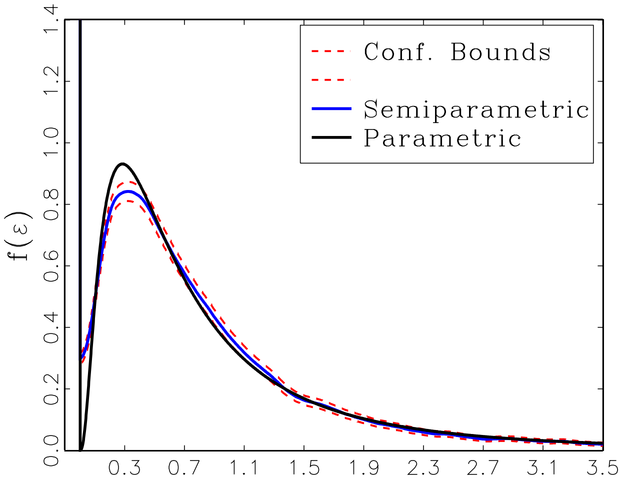


Figure: Estimated Density of ε_t (IBM, Boundary Kernel)

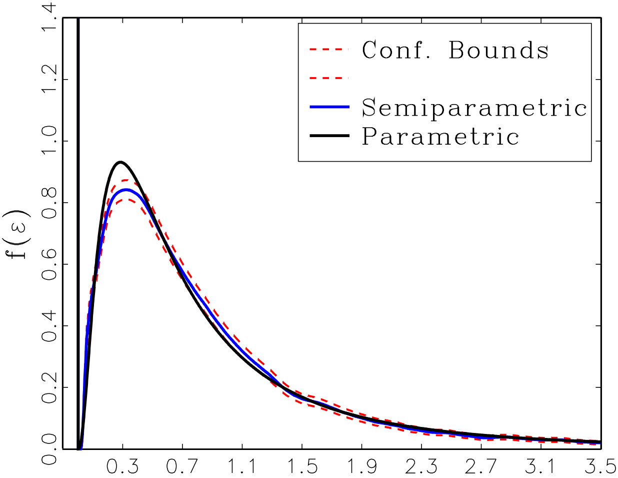
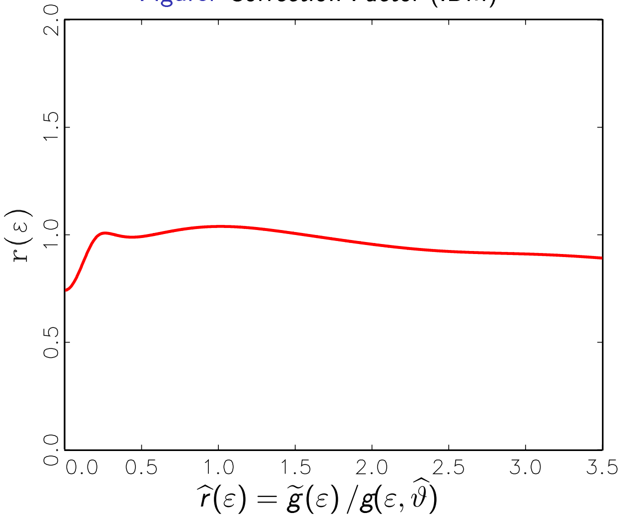


Figure: Correction Factor (IBM)



Wald-Wolfowitz-Runs-Test

- Define binary trade indicator:

$$\mathcal{I}_t = \begin{cases} 1 & \text{if } x_t > 0 \\ 0 & \text{if } x_t = 0. \end{cases}$$

- H_0: No serial dependence in $\{\mathcal{I}_t\}$.
- Runs: $\underbrace{1 \quad 1}_{run} \underbrace{0 \quad 0 \quad 0}_{run} \ldots \underbrace{1}_{run} \underbrace{0}_{run} \underbrace{1}_{run} \ldots \Rightarrow R = \#$ runs
- Under H_0: $Z = \frac{R - E[R]}{V[R]} \overset{a}{\sim} N(0,1)$.

Table: Runs-Test-Results

	Boeing	IBM
Z	-15.140	-8.915
p-val.	0.000	0.000

DZAGF-MEM: Log-Likelihood Function

- Log-likelihood function:

$$\mathcal{L} = \underbrace{\sum_t \{\mathcal{I}_t \log p_t + (1-\mathcal{I}_t)\log(1-p_t)\}}_{\text{ACM part}}$$

$$+ \underbrace{\sum_{t,nz} \{\log g_\varepsilon[x_t/\varkappa(\mu_t)] - \log \varkappa(\mu_t)\}}_{\text{MEM part}}$$

$$\mathcal{I}_t = \begin{cases} 1 & \text{if } x_t > 0 \\ 0 & \text{if } x_t = 0. \end{cases}$$

Unit mean condition: $\lambda_t = 1/p_t$.

Table: Ljung-Box-Statistics for ACM Residuals

	Boeing		IBM	
	Stat.	P-Val.	Stat.	P-Val
LB_{20}	17.095	0.646	17.182	0.641
LB_{50}	49.617	0.488	37.268	0.908
LB_{100}	92.502	0.690	68.166	0.993

Definition of residuals:

$$u_t = \frac{\mathcal{I}_t - p_t}{\sqrt{p_t\,(1 - p_t)}}.$$

Randomized PIT

- Add random noise to discrete random variables (e.g. Brockwell, 2007).
- Applied to zero-augmented MEM structure

$$z_t = \begin{cases} U_t \, Q_t(x_t) & \text{if } x_t = 0 \\ Q_t(x_t) & \text{if } x_t > 0, \end{cases}$$

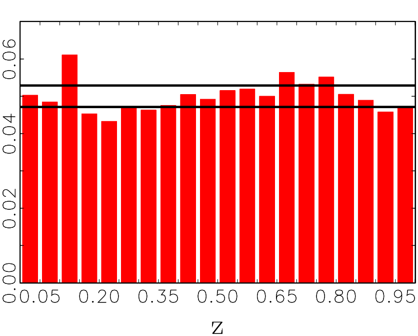
where $\{U_t\}_{t=1}^{n}$ is i.i.d. $U(0,1)$.

- Extensive notation

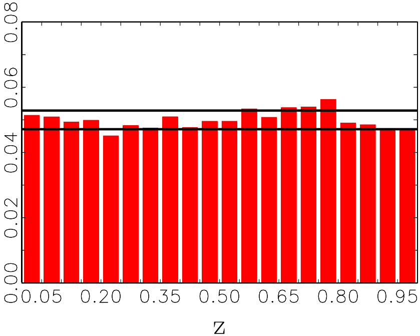
$$z_t = \begin{cases} U_t \, (1 - p_t) & \text{if } x_t = 0 \\ (1 - p_t) + p_t \, G_\varepsilon(x_t/\varkappa(\mu_t)) & \text{if } x_t > 0, \end{cases}$$

$G_\varepsilon(\cdot)$ denotes c.d.f. of the Generalized F distribution.

Figure: Histogram of Z (IBM)



(a) ZAGF-MEM



(b) DZAGF-MEM