# A Zero-Augmented Multiplicative Error Model

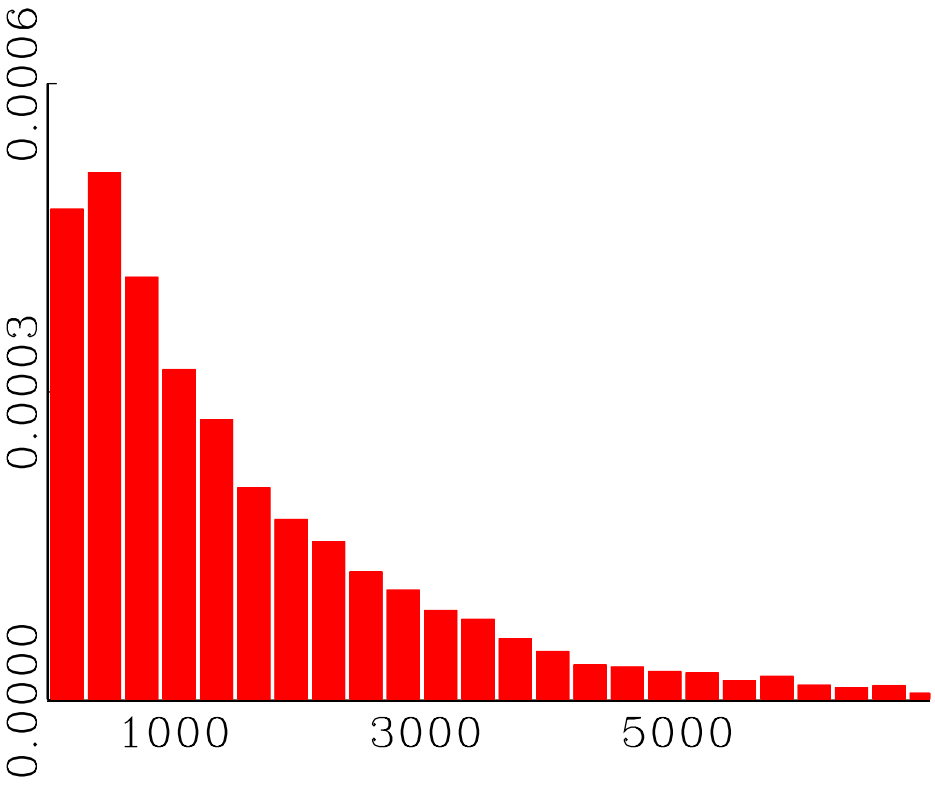
Nikolaus Hautsch

Peter Malec

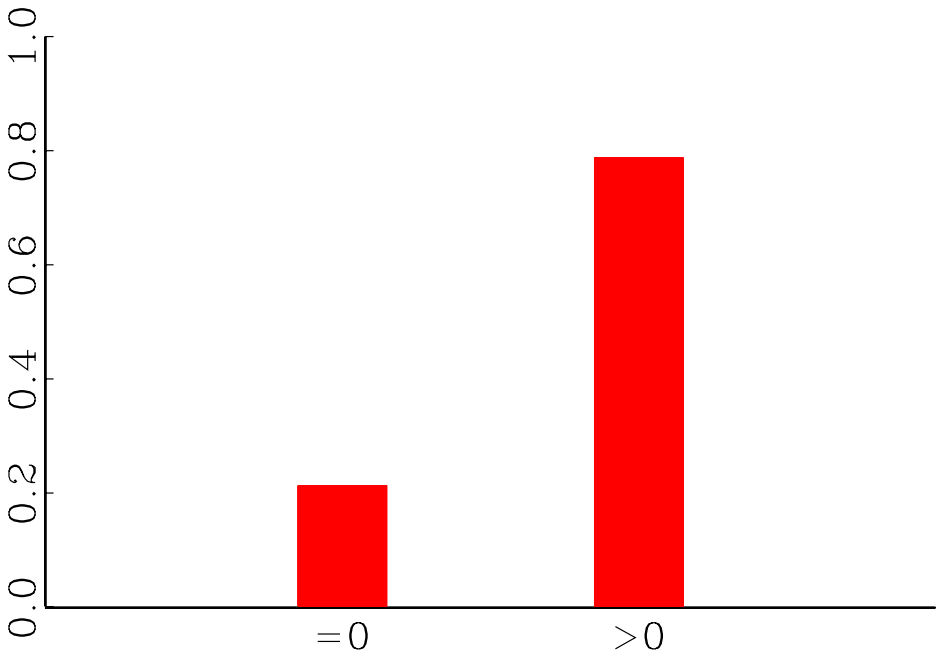
Melanie Schienle

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 Generalized F MEM

- $x_t$: a positive valued process.
- Extended MEM structure:

$$x_t = \varkappa(\mu_t)\,\varepsilon_t; \quad \varkappa(\mu_t) = \mu_t/\xi,$$

where:

$$\begin{aligned} \mu_t &= E(x_t|\mathcal{F}_{t-1})\,, \\ \xi &= E(\varepsilon_t)\,, \\ \varepsilon_t|\mathcal{F}_{t-1} &\sim i.i.d. \end{aligned}$$

- Assumption: $\varepsilon_t$ follows Generalized F distribution.
  $\Rightarrow$ Model nests different error distributions: Generalized Gamma, Weibull, Log-Logistic...

- pdf and mean of $\varepsilon_t$:

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

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

- Log-likelihood function:

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

- Zero values:
  - Distribution not defined for $\varepsilon_t = 0$ when $am < 1$.
  - Log-Likelihood cannot be evaluated for $x_t = 0$ ($\log x_t$).

# MEMs and Zero Values

- 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. Admati & Pfleiderer, 1988).
- Models for continuous nonnegative data with excess zeros: Mainly applied in other fields of research (e.g. biometrics, environmetrics).
- Main objective of the project: Develop a substantiated econometric approach to dealing with zero values in a MEM framework.

# Outline

1. Introduction ✓
2. Modeling Approach
3. Data
4. Empirical Results I
5. Model Extension
6. Empirical Results II
7. Outlook
8. Appendix

# 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 $\varepsilon_t$ is discontinuous with discontinuity at zero:

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

where $\quad \delta(\varepsilon_t) = \begin{cases} \infty & \text{for } \varepsilon_t = 0 \\ 0 & \text{for } \varepsilon_t \neq 0. \end{cases}$ (heuristic definition!)

- Density modeling approach follows Weglarczyk et al. (2005).

# Application to the Generalized F MEM

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

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

- Hence:

$$E(\varepsilon_t|\varepsilon_t > 0) = \mu_{\text{org}} = \lambda\,\xi.$$

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

$$E(\varepsilon_t) = p\,\mu_{\text{org}} \stackrel{!}{=} \xi,$$
$$\lambda = \frac{1}{p}.$$

# DGF-MEM

- Zero-augmented MEM structure:

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

where:

$$\begin{aligned} \varepsilon_t &\sim \mathrm{DGF}(\lambda, a, m, \eta, p)\,, \\ \varkappa(\mu_t) &= \mu_t/\xi, \\ \xi &= E(\varepsilon_t)\,. \end{aligned}$$

- Conditional density of the endogenous variable $x_t$:

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

- In the following: Log-MEM$(1,1)$ specification for $\mu_t$

$$\log\mu_t = \omega + \alpha\,\log x_{t-1}\,\mathbb{1}_{(x_{t-1}>0)} + \alpha^0\,\mathbb{1}_{(x_{t-1}=0)} + \beta\,\log\mu_{t-1}.$$

# A Semiparametric Benchmark

- Estimation of $\theta$ in $\mu\left(\mathcal{F}_{t-1};\theta\right)$ without explicit assumptions regarding $f_\varepsilon(\varepsilon_t)$.
  $\Rightarrow$ Quasi-Likelihood function:

$$\mathcal{Q} = -\sum_t \left[\frac{x_t}{\mu_t} + \log \mu_t\right].$$

- 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 $\widehat{\varepsilon}|\widehat{\varepsilon} > 0$ ,
  - $\widehat{p}$: frequency $N\left(\widehat{\varepsilon} > 0\right)/N$,

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

# Data

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

| | AIG | Boeing | IBM | Merck |
|---|---|---|---|---|
| % Zeros | 0.114 | 0.188 | 0.086 | 0.153 |

Figure: PDF of $\varepsilon_t$ implied by ML-estimates of DGF-MEM and semiparametric benchmark (Boeing)

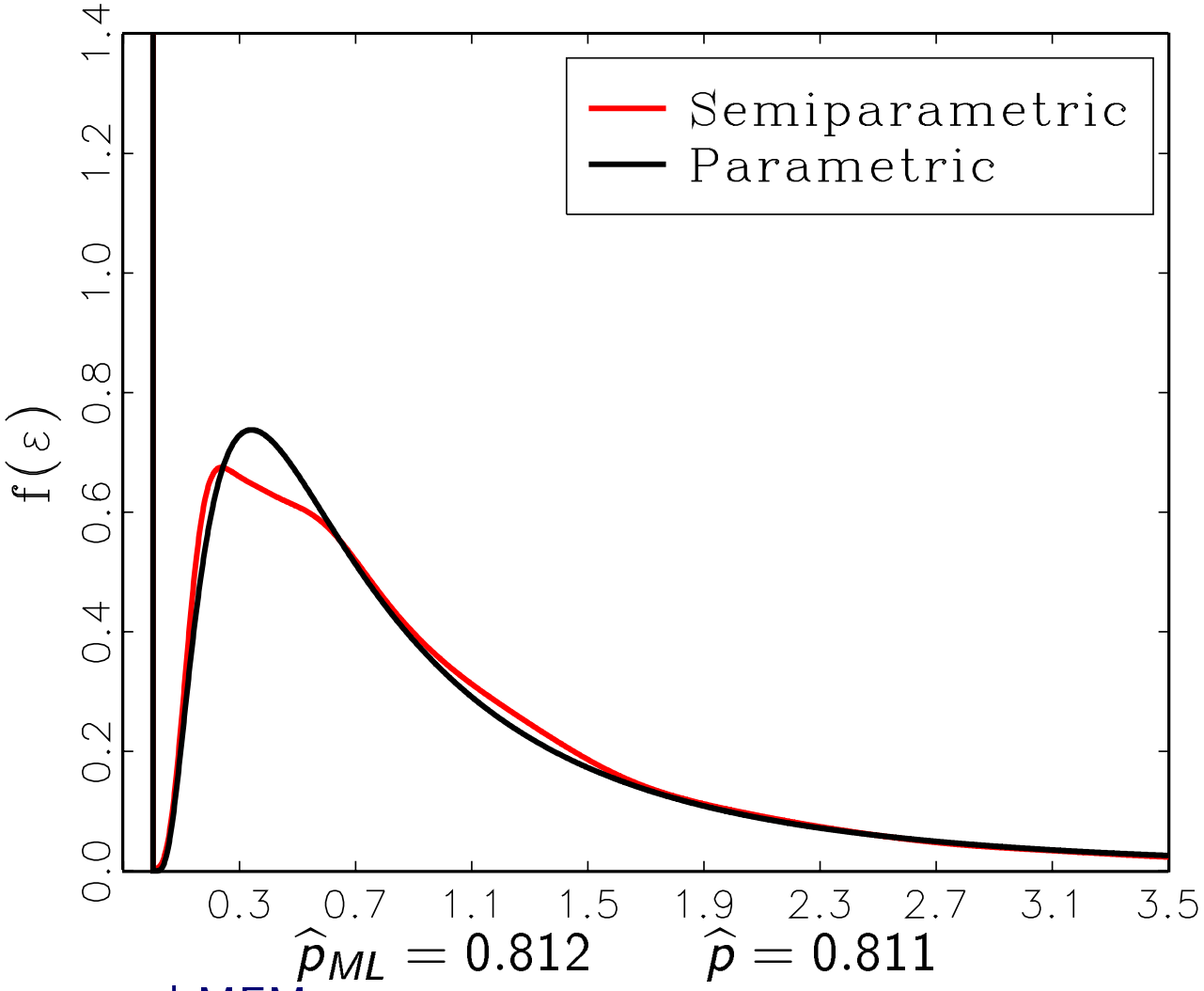
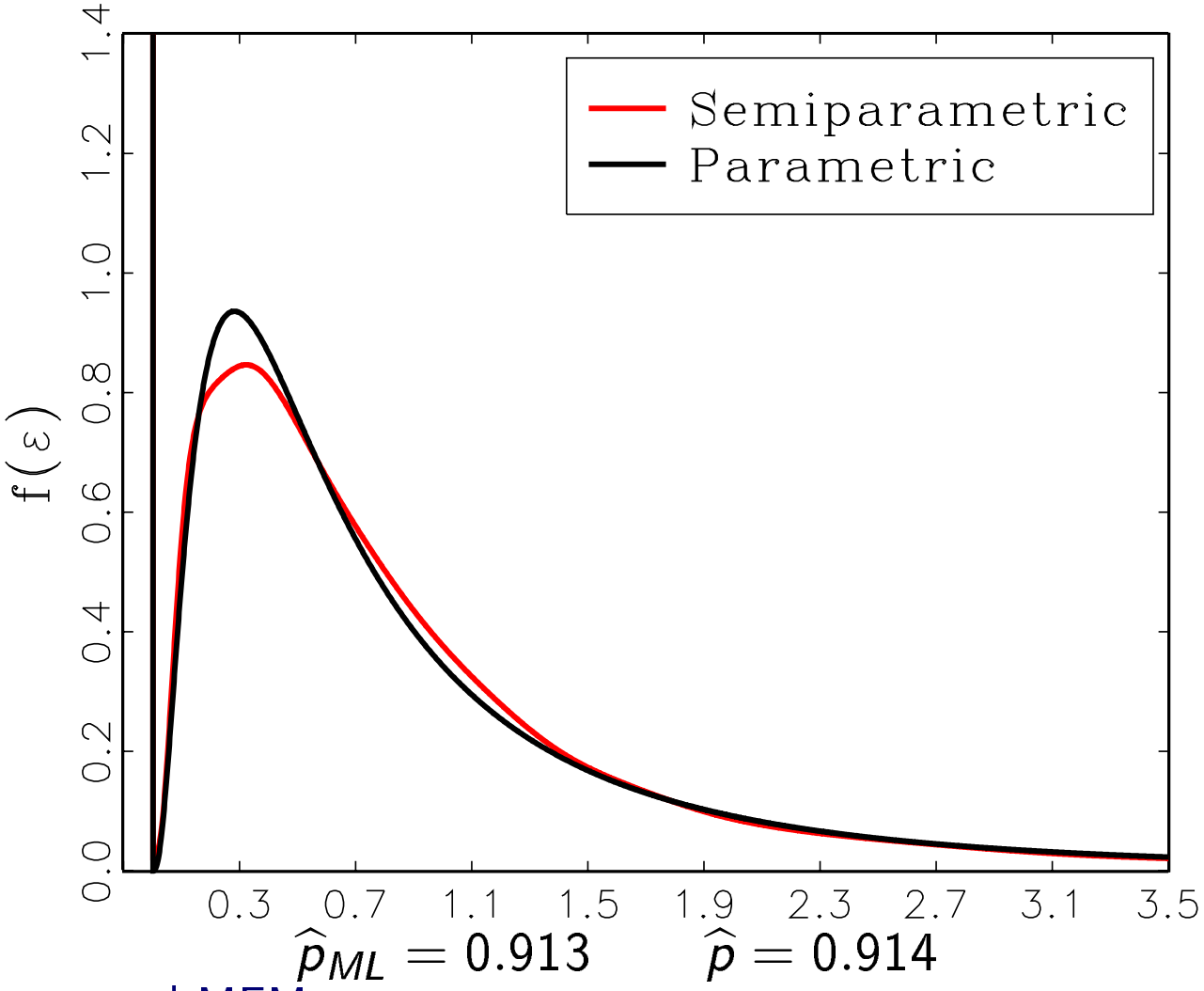


Figure: PDF of $\varepsilon_t$ implied by ML-estimates of DGF-MEM and semiparametric benchmark (IBM)



# Extension by a Binary Choice Model

- Define binary trade indicator:

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

- Probability $p$ is time-varying:

$$p_t = P\left(\varepsilon_t > 0 | \mathcal{F}_{t-1}\right) = P\left(\mathcal{I}_t = 1 | \mathcal{F}_{t-1}\right) = p(\mathcal{F}_{t-1}; \pi)$$

$\Rightarrow$ Binary choice model

- 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.

# Specification of $p_t$

- Autologistic model (Cox, 1958):

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

where

$$w_t = \theta_0 + \sum_{i=1}^{l} \theta_i \, y_{t-i} + g_t, \quad \text{and } g_t = \sum_{i=1}^{d} \gamma_i \, \mathcal{I}_{t-i}.$$

- Natural choice: $y_{t-i} = x_{t-i}$.
- Rydberg & Shephard (2003): to reduce multicollinearity define

$$\text{where } m_t = \begin{cases} x_t - \mathcal{I}_t & \text{if } x_t \geq \mathcal{I}_t \\ 0 & \text{else.} \end{cases}$$

Table: Estimation Results - LDGF-MEM - Logit Part

| | **Boeing** | | | **IBM** | | |
|---|---|---|---|---|---|---|
| | Coef. | T-Stat. | P-Val. | Coef. | T-Stat. | P-Val. |
| $m_{t-1}$ | 0.057 | 2.741 | 0.006 | 0.105 | 3.235 | 0.001 |
| $\mathcal{I}_{t-1}$ | 0.495 | 11.594 | 0.000 | 0.593 | 7.720 | 0.000 |
| $\mathcal{I}_{t-2}$ | 0.291 | 6.664 | 0.000 | 0.490 | 6.305 | 0.000 |
| $\mathcal{I}_{t-3}$ | 0.246 | 5.580 | 0.000 | 0.367 | 4.678 | 0.000 |
| $\mathcal{I}_{t-4}$ | 0.205 | 4.621 | 0.000 | 0.136 | 1.620 | 0.105 |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $\mathcal{I}_{t-9}$ | 0.114 | 2.537 | 0.011 | 0.220 | 2.704 | 0.007 |
| $\mathcal{I}_{t-10}$ | 0.196 | 4.426 | 0.000 | 0.335 | 4.183 | 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. $U(0,1)$.

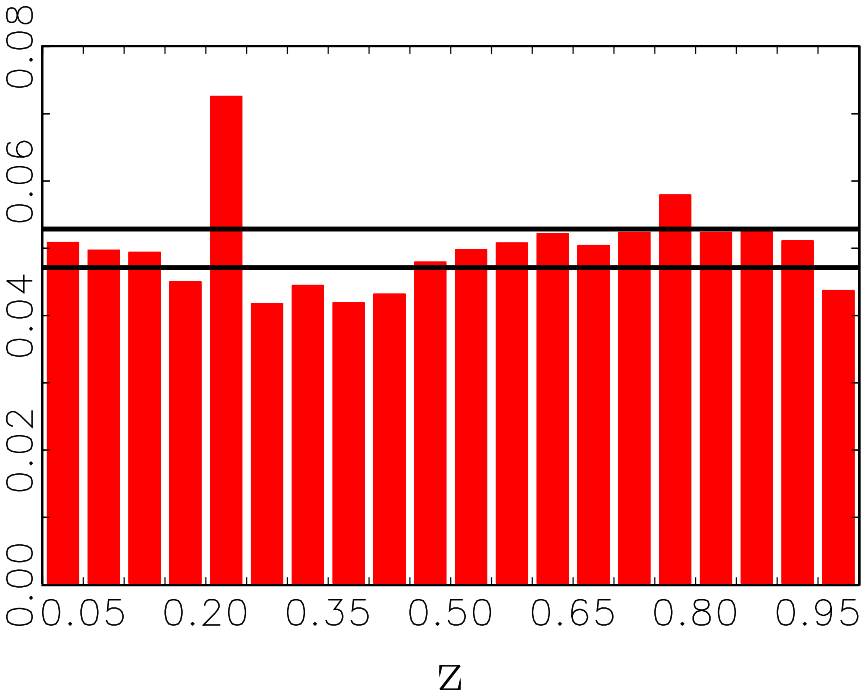
# Randomized PIT

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

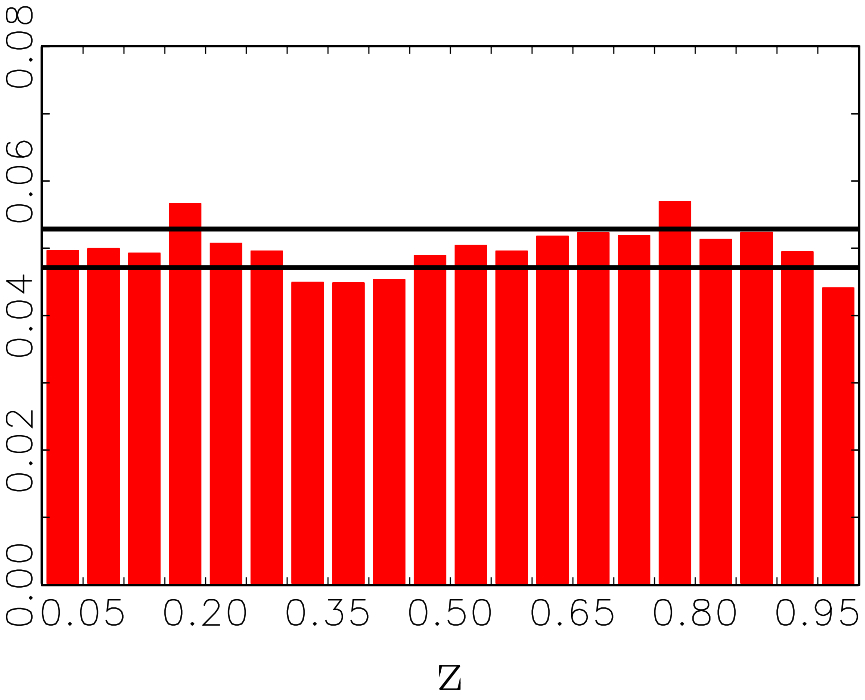
$$z_t = \begin{cases} U_t \, Q_t(x_t) & \text{if } x_t = 0 \\ Q_t(x_t) & \text{if } x_t > 0, \end{cases}$$

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

Figure: Histogram of Z (Boeing)

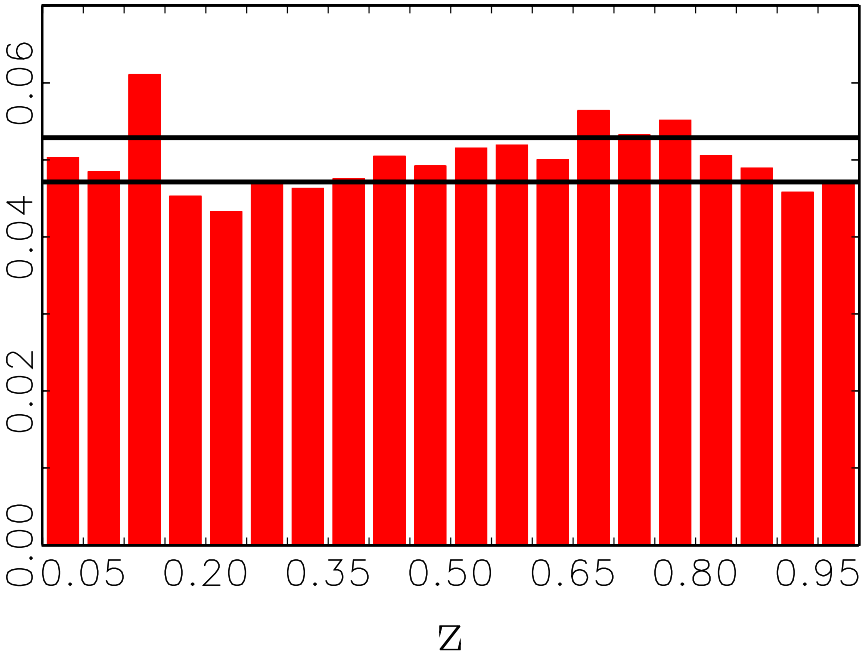



(a) DGF-MEM

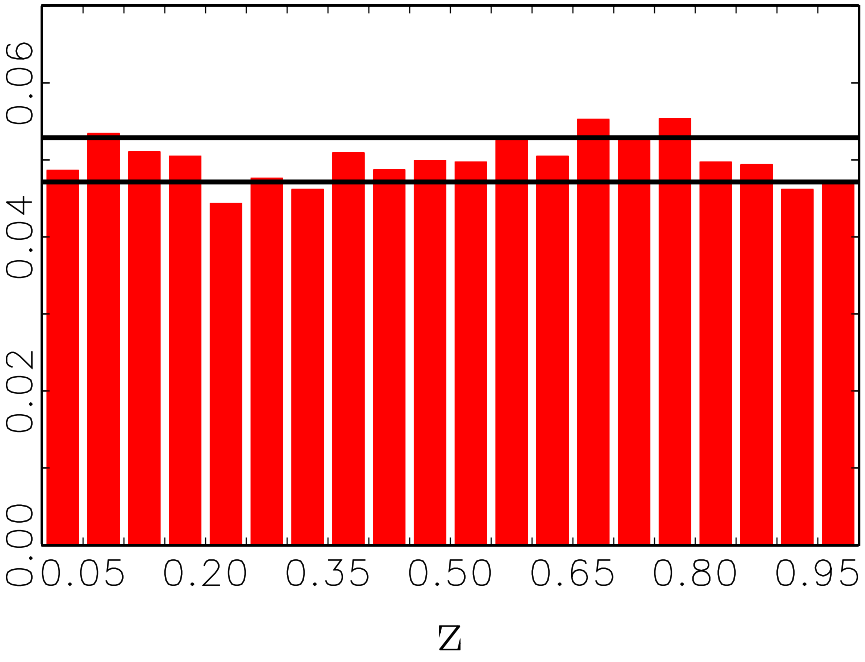



(b) LDGF-MEM

Figure: Histogram of Z (IBM)




(a) DGF-MEM




(b) LDGF-MEM

Table: $\chi^2$-Test for Uniformity

| | **Boeing** | | **IBM** | |
|---|---|---|---|---|
| | $\chi^2$ | P-Val. | $\chi^2$ | P-Val |
| DGF-MEM | 275.357 | 0.000 | 105.033 | 0.000 |
| LDGF-MEM | 68.230 | 0.000 | 50.379 | 0.001 |

- Test for parametric density based on semiparametric benchmark.
- Binary choice component:
  - Improved specification: e.g. GLARMA.
  - (A second) semiparametric benchmark.
- Simple (economic) application.

# KDE Approach

- Support of $g_\varepsilon(\varepsilon_t)$ is bounded from below: boundary bias!
- Possible remedy: use of asymmetric kernels.
  $\Rightarrow$ Gamma kernel (Chen, 2000):

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

  where

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

- KDE for $g_\varepsilon(\varepsilon_t)$:

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

# Wald-Wolfowitz-Runs-Test

- $H_0$: 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 $H_0$: $\quad Z = \frac{R - E(R)}{V(R)} \overset{a}{\sim} N(0,1)$.

Table: Runs-Test-Results

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

Table: Ljung-Box-Statistics for Autologistic Residuals

| | **Boeing** | | **IBM** | |
|---|---|---|---|---|
| | Stat. | P-Val. | Stat. | P-Val |
| $LB_{20}$ | 76.016 | 0.000 | 103.588 | 0.000 |
| $LB_{50}$ | 307.729 | 0.000 | 321.067 | 0.000 |
| $LB_{100}$ | 545.430 | 0.000 | 561.089 | 0.000 |

Definition of residuals:

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