Depth Based Procedures

For Estimation ARMA and GARCH Models

Daniel Kosiorowski

Cracow University of Economics

COMPSTAT 2010 Paris 22 – 27 August 2010

I. Motivations

1. Highly effective tools of analysis and forecasting multivariate economic phenomena with stress that we are looking for a general tendency represented by a majority of objects.

- Daily updated tendency on a financial, monetary, agricultural market.
- Better understanding of a nature of a volatility
- Modeling of an anticipation of investors for a certain event, a government decision.

2. Construction of economic aggregates, indexes, ratings etc.

- Robust risk measures, measures of an allocation of an information between agents.
- A search for invariants, laws of conservation in economics.

3. Perspective problems – modeling

- Dependencies between preferences, choices of consumers, behaviors of agents
- Stresses on a financial market
- Stability of a general economic equilibrium

II. Outline

1. Robust economic analysis
2. Why depth based analysis of economic phenomena
3. Regression depth and beyond
   a) General band depth and a median path of a development of a system
   b) Robust ARMA estimator
   c) Robust GARCH estimator
4. Further inspirations – Mizera & Muller location - scale depth
5. Conclusions

ROBUST ECONOMIC ANALYSIS

John Maynard Keynes about investing in stocks (an advantage of being outlier?)

"It is the one sphere of life and activity where victory, security and success is always to the minority and never to the majority. When you find any one agreeing with you, change your mind. When I can persuade the Board of my Insurance Company to buy a share that, I am learning from experience, is the right moment for selling it."

Copernicus – Gresham law „Bad money drives out good under legal tender laws"

(a disadvantage of the outlier activity?)

An intuition that a certain behavior of an economic system is an effect of an activity of a majority of its element has a long history in economics ex. bankruptcy of a bank as an effect of a conviction of a majority of its clients

Problem: How to understand "…is an effect of an activity of a majority of its elements…

ROBUST ECONOMIC ANALYSIS

Needs:

1. A basis for continuous decision making process in a changing situation on a market / *unambiguous interpretations in different characteristics of a market uncertainty* /
2. Discovering a general tendency on a market / robustness as a fit to a majority /
3. Descriptive statistics in multivariate case /ex. multivariate skewness (as an activity of an external force), multivariate kurtosis (as a degree of beliefs of agents as to…)

Desirable properties:

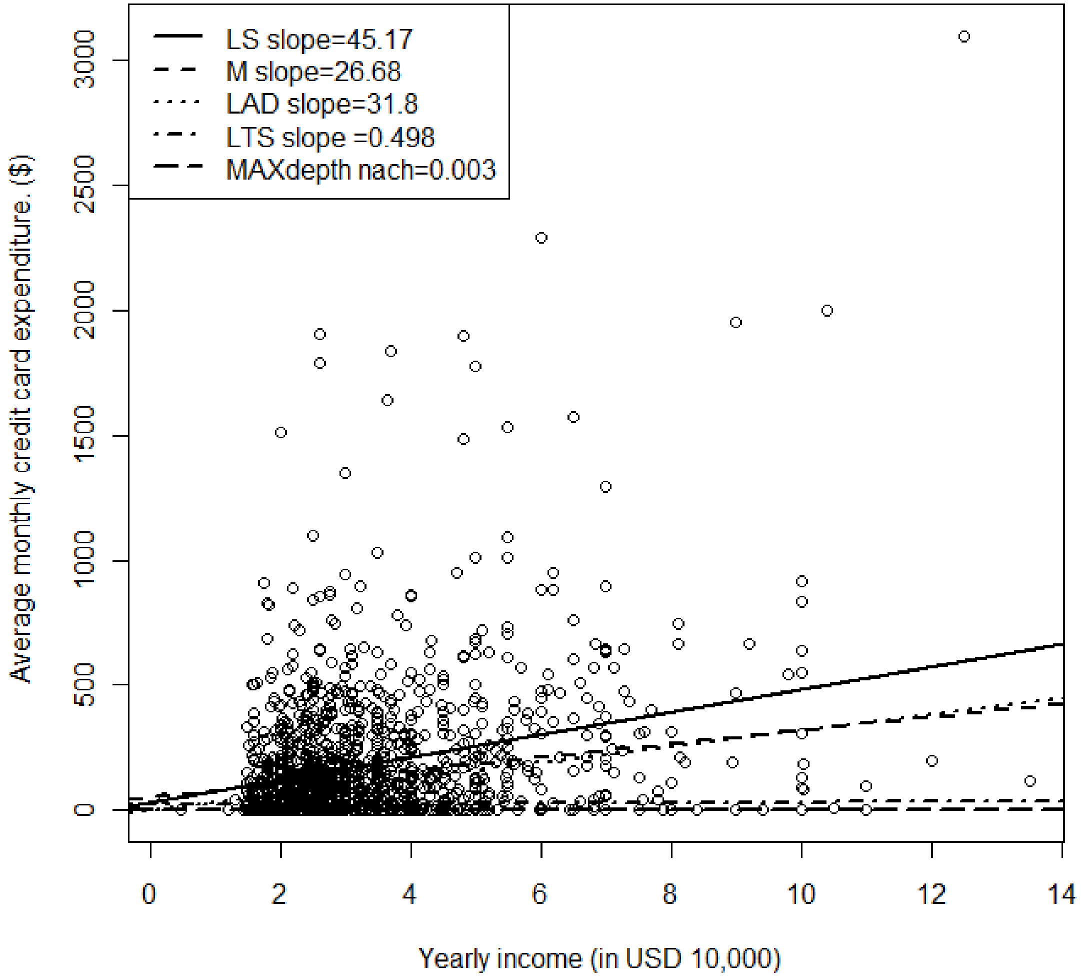
1. Good small and moderate sample behavior (20 – 100 obs), good efficiency in case of fat tailed, skewed populations, almost sure convergence and good rate of convergence.
2. Finite sample breakdown point robustness.
3. Algorithms, software, stability, approximate algorithms
4. User friendly graphical techniques of results presentation

SOLUTION:

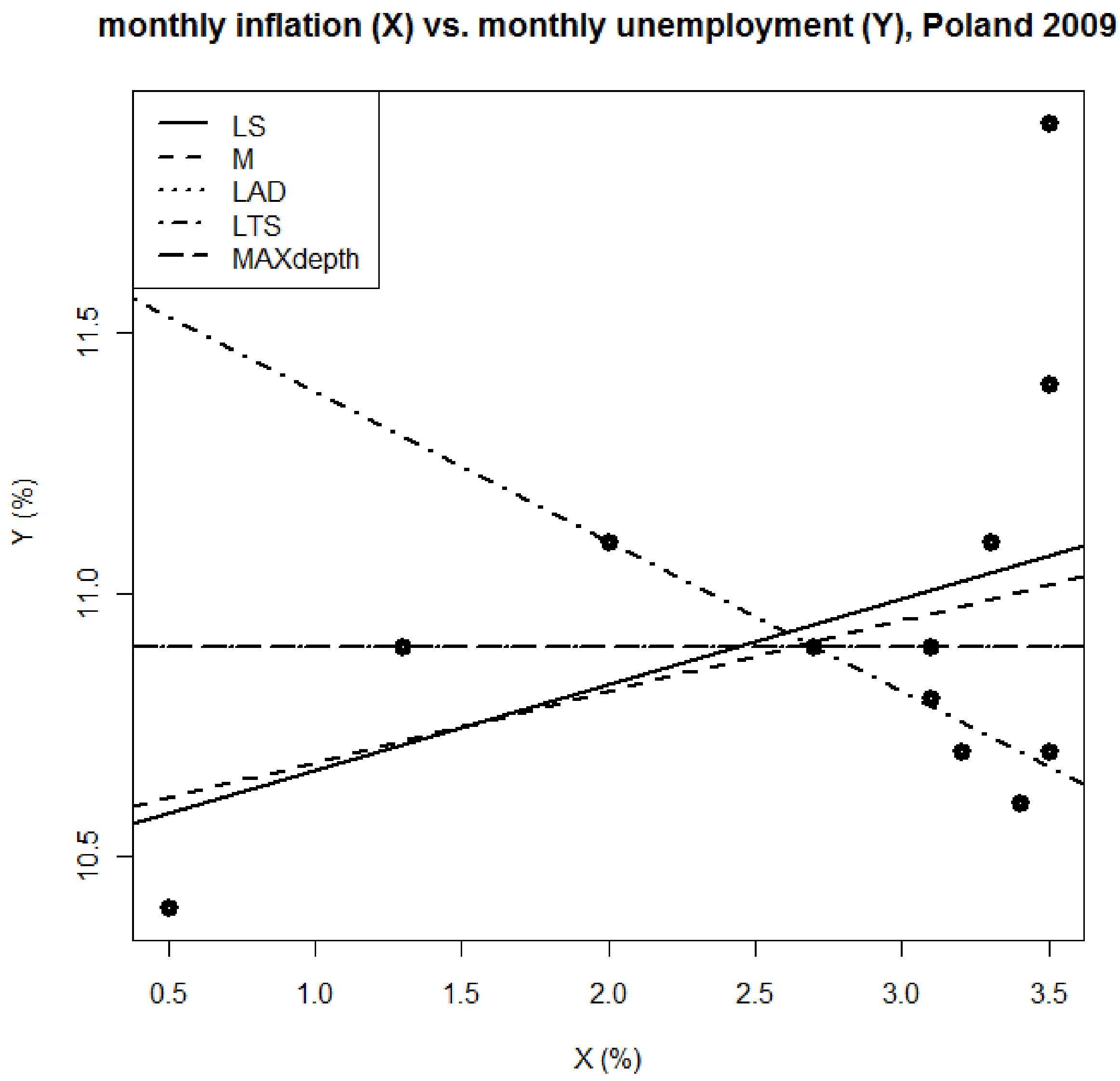
Regression depth / *a variety of economical regression based procedures, theoretical background proposed by Mizera (2002)* **/**

Projection depth / *very good statistical properties showed by Zuo (2003), effective calculation via approximate algorithm proposed by Dyckerhoff (2004)* **/**

Average monthly credit card expenditure ($) vs. yearly income (in 10 000 $) of individuals.



Source: {AER} package R project (Greene, W.H. (2003). *Econometric Analysis*, 5th edition. Upper Saddle River, NJ:Prentice Hall)


monthly inflation (X) vs. monthly unemployment (Y), Poland 2009
LS
M
LAD
LTS
MAXdepth
Y (%)
11.5
11.0
10.5
0.5
1.0
1.5
2.0
2.5
3.0
3.5
X (%)

3. REGRESSION DEPTH

Famous concept introduced by Rouseeuuw and Hubert (Rouseeuuw & Hubert 1998) generalized by Mizera (Mizera 2002) and Mizera & Muller (Mizera & Muller 2004), studied among others by Rouseeuuw, Van Aelst and Van Driessen, Bai & He (Bai & He (1999)) and …

Good points

- A variety of possible applications in economics
- An interesting criterion of fit similar to economical understanding of Pareto's optimality
- Large model generating data (heteroscedasticity, autocorellations, skewness of errors)
- Existence of computational algorithms via free software

Bad points

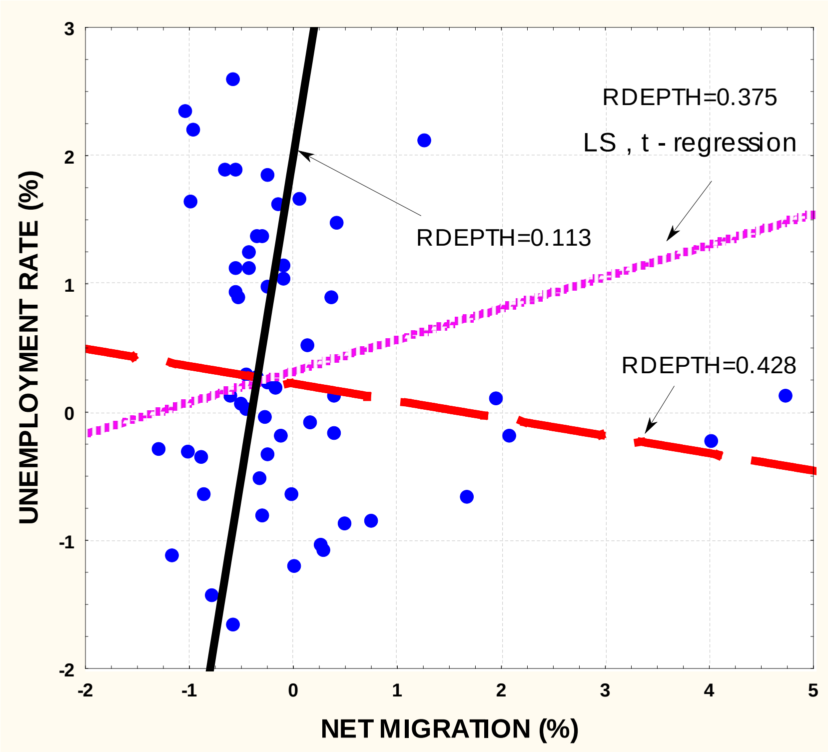
- Computation complexity, small sample theory

***Dilemmas** :*

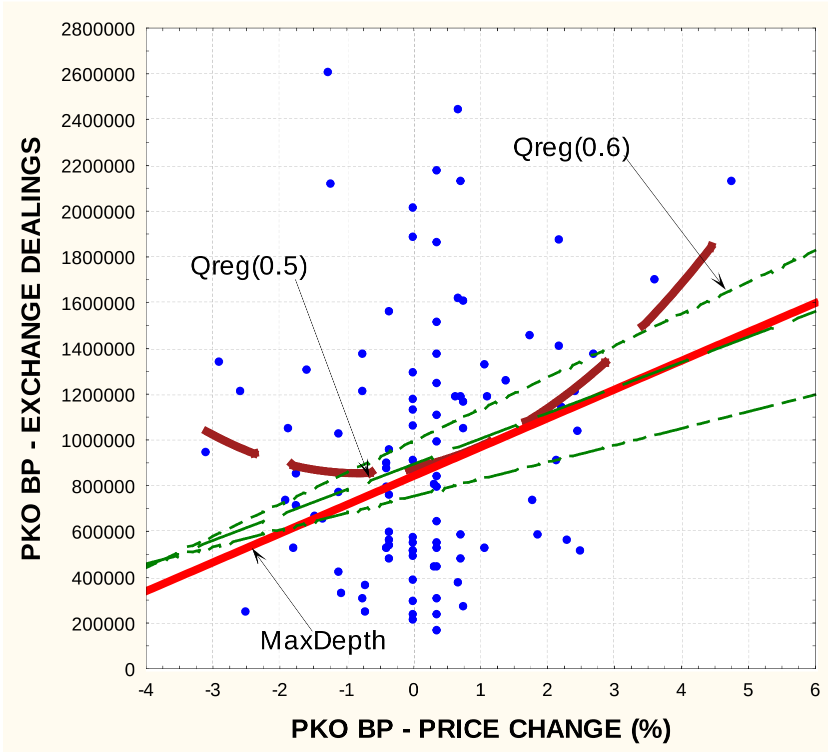
Modeling by means of a linear robust regression or more complicated nonlinear regression.

Better fit to data but weak economical basis or better economical background and weak fit.

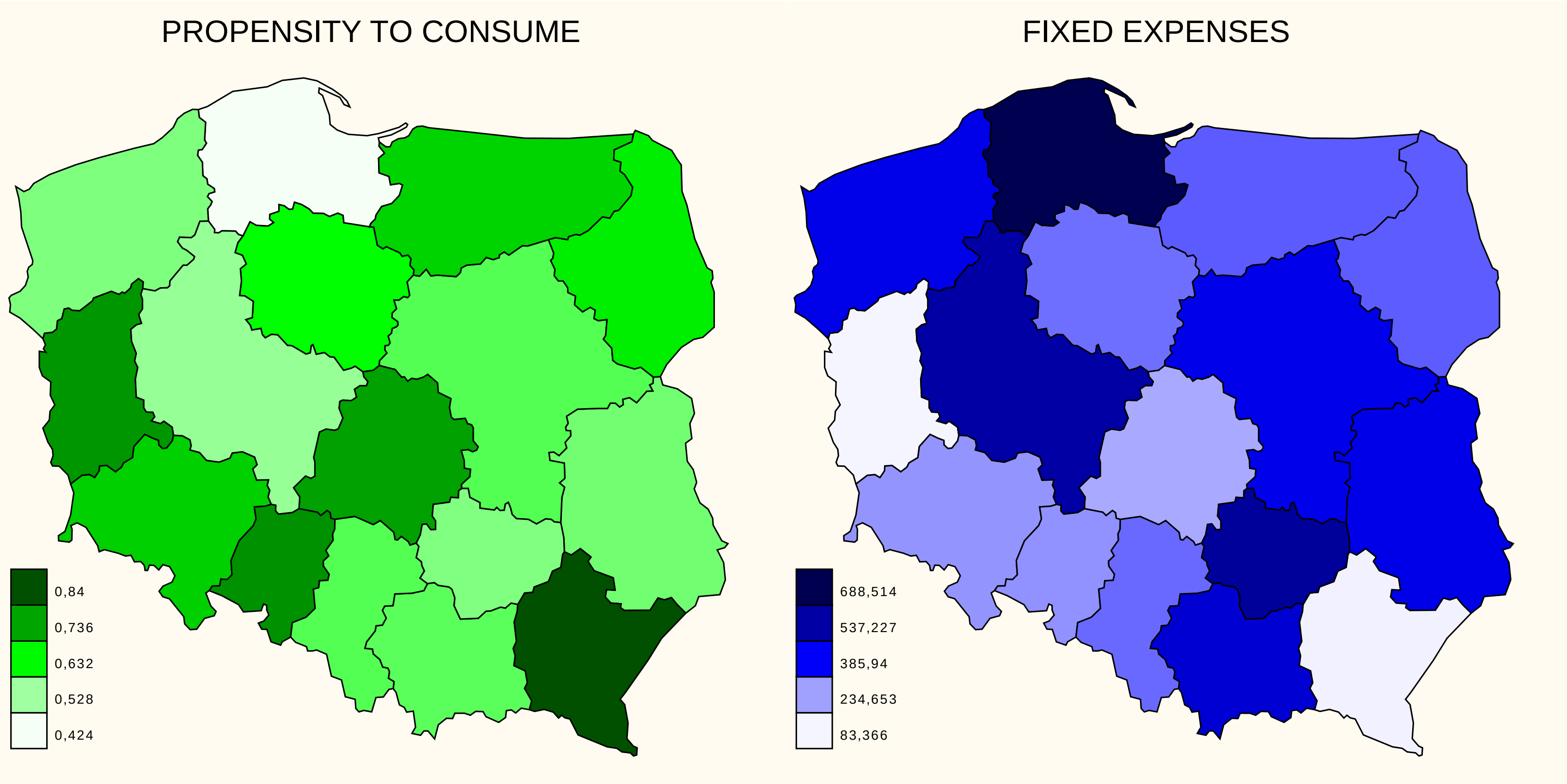
Unemployment rate (%) vs net migration (%) in polish border subregions (powiats) in 2005 year



Exchange dealings vs price change of share PKO BP bank



Propensity to consume and fixed expenses - slopes and intercepts in maximal regression depth estimation of consumption vs household's available income in polish district in 2006 year.



DEPTH BASED ANALYSIS OF ECONOMIC TIME SERIES

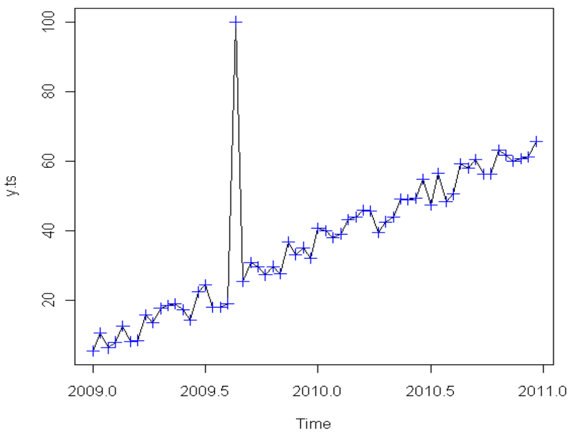
Time series analysis has a special position in economics (GDP growth, inflation dynamics, effectiveness of an intervention on a currency market etc.)

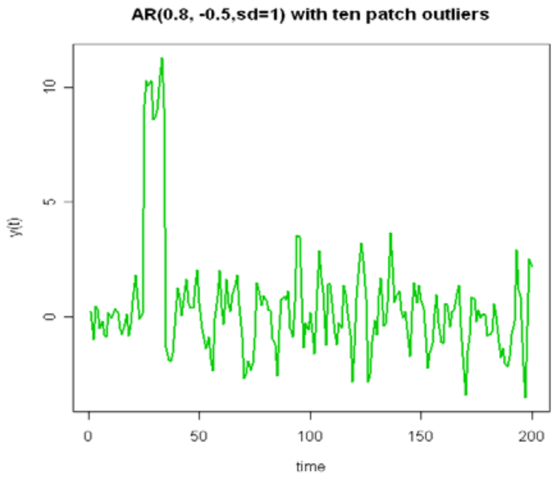
Robust time series analysis seems to be especially challenging because of temporal dependence in the data, various types of outliers (outlying time series and outliers in time series of various types see ex Maronna, Martin & Yohai (2006))

Difficulties in modeling economical time series: (short samples, changes of model generating data, insufficient apriori knowledge, difficulty in understanding of changes in behaviors of agents)

Insufficient theoretical background for GARCH rather than ARMA modeling of economic systems. In practice we are only looking for a description of a general tendency: simpler model better model (GARCH(1,1), AR(1))

Several types of outliers in case of time series, outliers may be isolated or occur in patches.





Suppose that GARCH(m,r) or ARMA(p,q) series is given by y_t. A observed series corresponding to the isolated additive outliers (AO) is $y_t + v_t u_t$, where y_t, v_t and u_t are independent processes, $P(u_t = 1) = \varepsilon$ / an outlier occurs/, $P(u = 0) = 1 - \varepsilon$, $\varepsilon \in [0, 0.5)$.

Other types of outliers: *replacement outliers, innovation outliers (depend on a considered model). Notice that in case of time series we can discriminate between outlying time series and outliers in time series.*

Dilemmas:

Robust procedures and simple linear model or more complicated nonlinear model, switching regime model.

Better fit to data but weak economical basis or better economical background and weak fit (Support Vector Machine regression or ARMA modeling)

GARCH modeling of dependencies between price of an asset and its dispersion or comparative analysis using ex Mizera-Muller location – scale depth.

Robust measures on a model level /MAD as a measure of risk, projection median as an attractor of workers skills/ or only robust estimation, testing procedures /MAD as an estimate of SD, projection median as an estimate of center of the workers skills/.

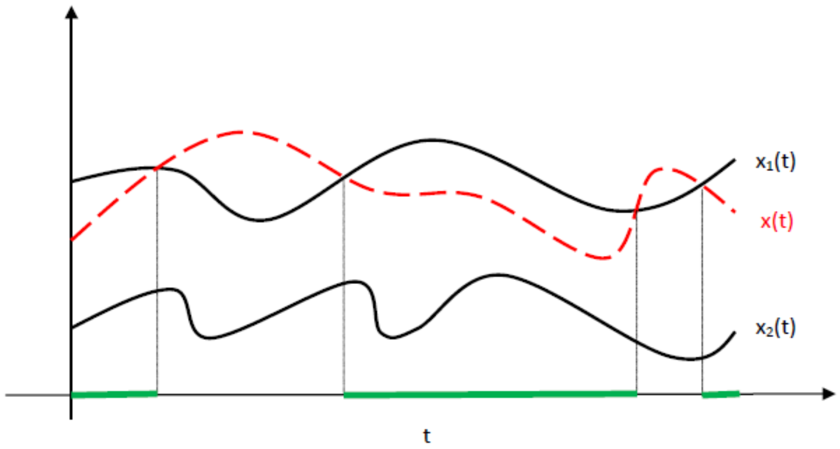
GENERAL BADN DEPTH – LOOKING FOR A TYPICAL PATH OF A DEVELOPMENT

An idea of depth for functions was proposed by Fraiman, Muniz i Lopez (2006) and Pintado & Romo (2006).General band depth (Pintado & Romo 2006) seems to be a valuable method of indicating a central type of an evolution of an economic system.

It is much easier to analyze a development of one company instead of whole branch.

Which path is typical, which is outlying?

- A company in a branch stock index / price, dealings /
- A country of a certain region / GDP, inflation dynamics, economic development path /
- Sales of a certain product after a promotion
- Analysis of effectiveness of a regional politics



For any function x in $\{x_1,...,x_n\}$, $j \geq 2$ let

$$A_j(x) \equiv A(x;x_{i_1},x_{i_2},...,x_{i_j}) \equiv \left\{ t \in [0,1] : \min_{r=i_1,...,i_j} x_r(t) \leq x(t) \leq \max_{r=i_1,...,i_j} x_r(t) \right\},$$

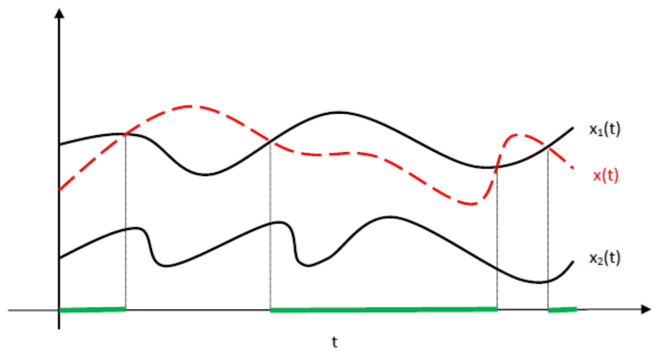
be the set of points in the interval $[0,1]$, where the function x is inside the band determined by the observations $x_{i_1},x_{i_2},...,x_{i_j}$.

If λ is the Lebesque'a measure on the interval $[0,1]$, $\lambda(A_j(x))$ is the proportion of time that x is inside the band.

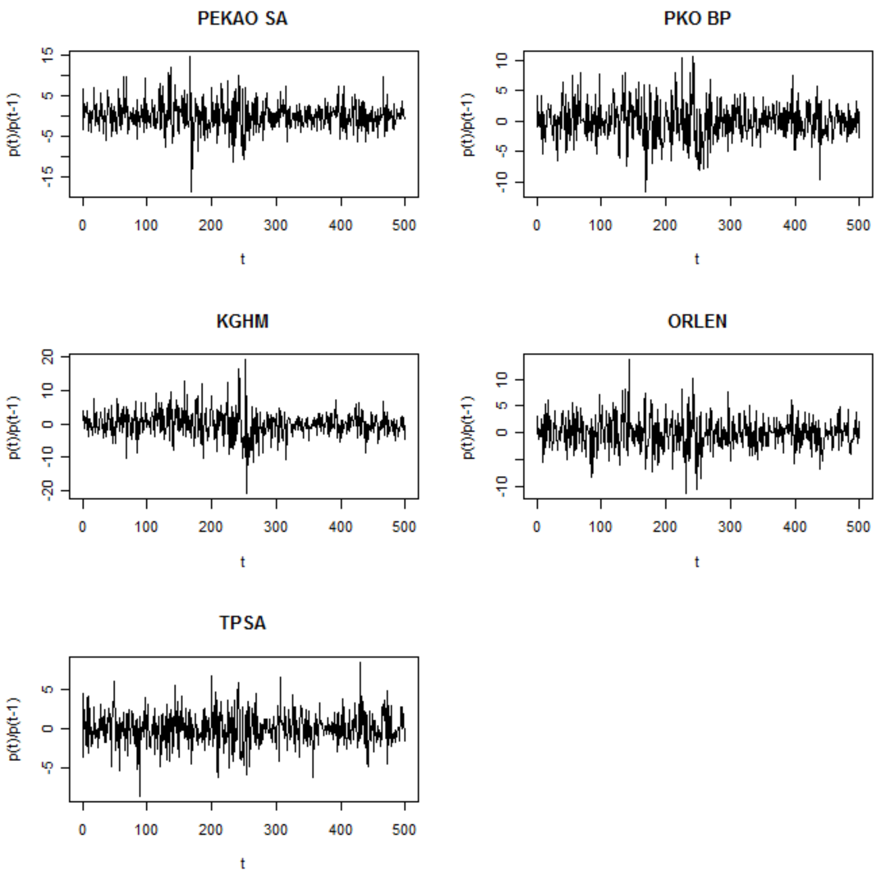
Calculating $GS_n^{(j)}(x) = \binom{n}{j}^{-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_j \leq n} \lambda(A(x;x_{i_1},x_{i_2},...,x_{i_j})),\ j \geq 2$

Pintado & Romo (2006) define generalized band depth as

$$GS_{n,J}(x) = \sum_{j=2}^{J} GS_n^{(j)}(x),\ J \geq 2.$$



Prices of stocks changes - exchange quotations of five biggest companies of polish stocks index WIG20



Pintado & Romo general band depth of five exchange quotations of biggest companies of polish stock index WIG20

Company	depth
PEKAO SA	0.2075
PKO BP	0.2295
KGHM	0.1674
PKN ORLEN	0.2083
TPSA	0.1872

REGRESSION DEPTH BASED ESTIMATOR OF ARMA MODEL

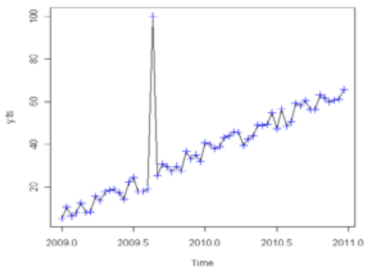
A time series $\{x_t; t = 0, \pm 1, \pm 2, ...\}$ is generated by ARMA(p,q) model if it is stationary and

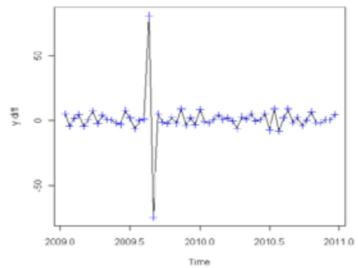
$$x_t = \alpha + \phi_1 x_{t-1} + ... + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + ... + \theta_q w_{t-q},$$

with $\phi_p \neq 0$, $\theta_q \neq 0$, and $\sigma_w^2 > 0$, the parameters p and q are called the autoregressive and moving average orders, respectively, $\alpha = \mu(1 - \phi_1 - \cdots - \phi_p)$, where μ denotes a mean of x_t, $\{w_t; t = 0, \pm 1, \pm 2, ...\}$ is (typically) Gaussian white noise sequence.

CLASSICAL ESTIMATES: Maximum Likelihood, Least Squares, Yule – Walker equations (see Maddala 2006, Box, Jenkins and Reinsel 1994).

Problems: propagation of an effect of outlier ex. estimation on a base of a lagged time series, differencing of a time series etc.





ROBUST PROPOSITIONS:

a) Maximum likelihood + residuals diagnostics -- propositions suffer from a masking problem

b) Robust estimators ex. M- estimator or S – estimator. Because of the fact that one outlier can affect several residuals – propositions have low BP (but here also filtered residuals propositions – asymptotically biased or lack of a asymptotic or small sample theory)

RECENT PROPOSITIONS:

M – estimators with bounded propagation of outliers (Muler, Pena and Yohai (2007)) – generalizations of the MM - estimates introduced by Yohai for regression. Their estimates are consistent and asymptotically normal under a perfectly observed ARMA model.

Problems: complicated calculation and lack of software

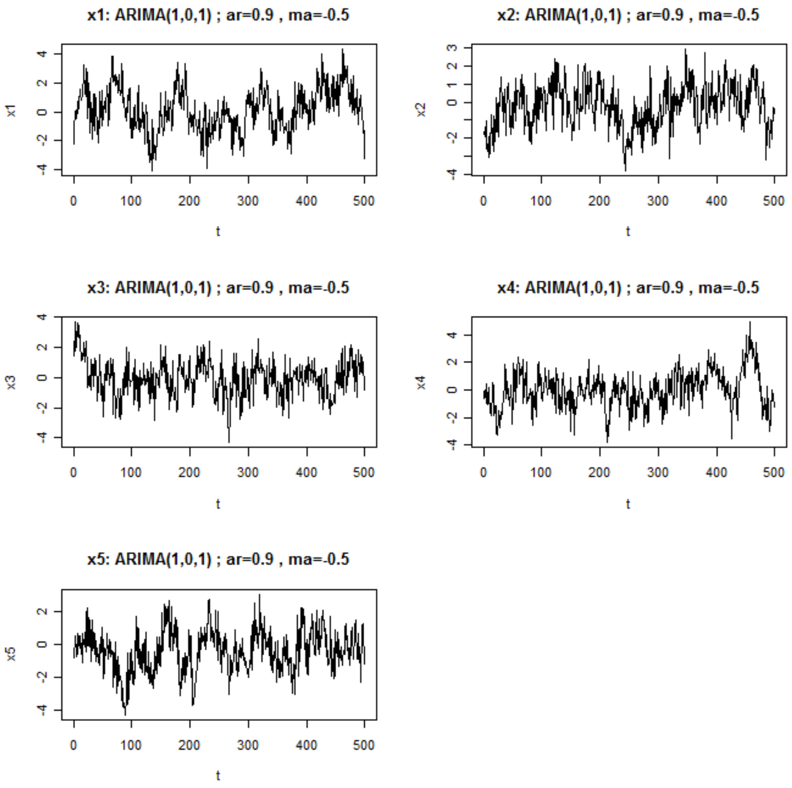
PROPOSITION 1: Let $\mathbb{X} = \{x_1, x_2, ..., x_T\}$, $2 < T$, denote a time series generated by ARMA(p,q) model containing additive outliers. We obtain estimates of the parameters of ARMA(p,q) model, $0 < p + q << T$ in a two step procedure:

STEP 1: We calculate maximum depth estimates of AR(p) part of the underlying process by choosing $\phi_1, ..., \phi_p$ as the maximal regression depth estimates applied to a data sets $\mathbb{Y} = \{x_1, ..., x_{T-p}\}$, $\mathbb{X}_1 = \{x_2, ..., x_{T-p+1}\}, ..., \mathbb{X}_p = \{x_{p+1}, ..., x_T\}$ created from $\mathbb{X}$ elements.

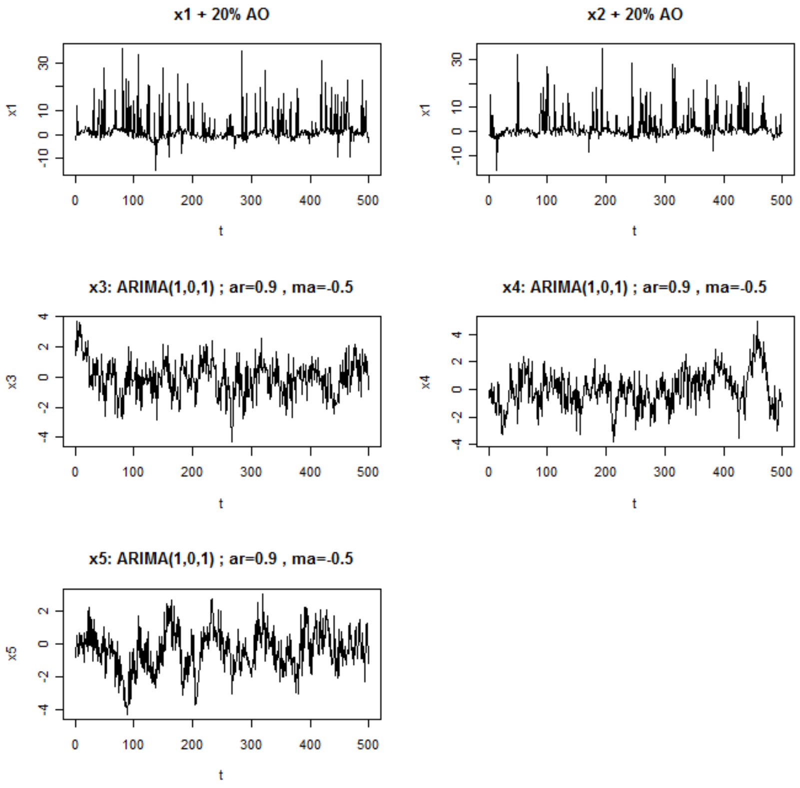
STEP 2: We add MA(q) part to the estimated in the step 1 AR(p) part by minimizing a robust measure of a dispersion between observed and generated by the model values e.g. for MAD (median absolute deviation).

(*We could use ex.* ***one step ahead prediction*** *and minimize MAD of differences between observed and predicted by the candidate ARMA model data. The AR part is fixed in the first step, the minimization concerns MA candidates*).

Simulated trajectories from ARMA (1,1) model, trajectories without additive outliers.



Simulated trajectories from ARMA (1,1) model, two trajectories contain 20% of additive outliers.



STATISTICAL PROPERTIES OF THE PROPOSITION:

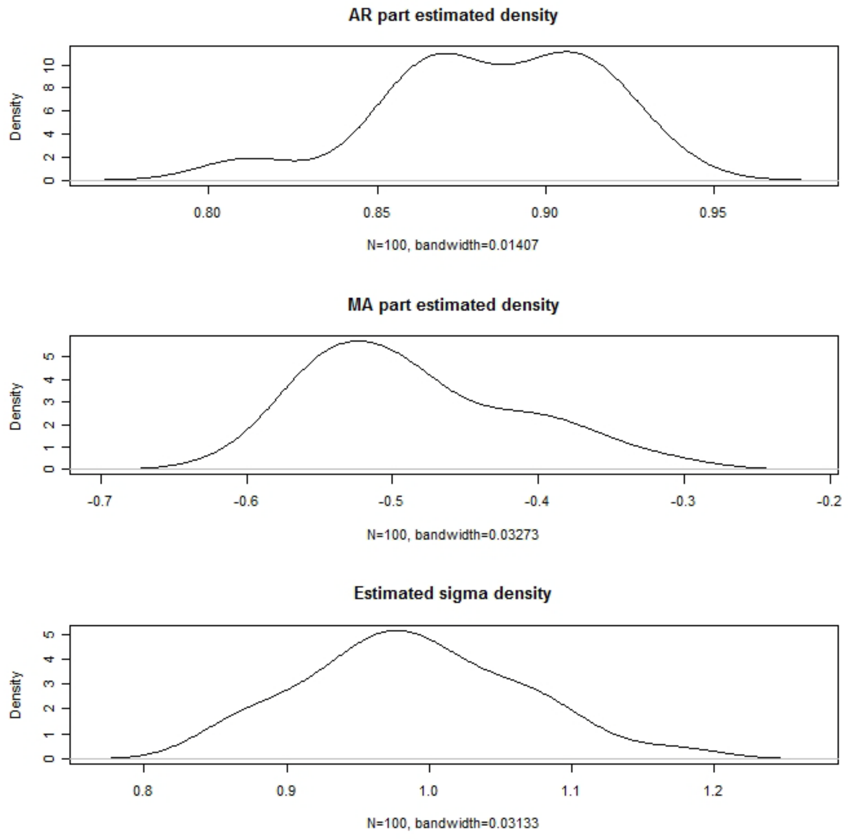
Various simulated data sets /skewed t errors, data containing <20% of additive outliers indicate on:

- High breakdown point of a procedure >10% for ARMA(1,1), AR(1)
- Satisfying efficiency at ARMA(1,1) normal model for errors
- Stability of estimation
- Good performance in case of skewed t errors

VERY GOOD PROPERTIES IN CASES OF SIMPLE MODELS

AR(1), ARMA(1,1) and sample sizes 100 observations

Kernel density estimation of the proposed parameters estimators of the ARMA(1,1) with $\phi_1 = 0.7$, $\theta_1 = -0.5$, $\sigma = 1$. Each of the simulated trajectories contained 5% of the additive outliers.



REGRESSION DEPTH BASED ESTIMATOR OF GARCH MODEL

If x_t is the value of a stock at time t, then the return or relative gain, y_t, of the stock at time t is $y_t = \frac{x_t - x_{t-1}}{x_{t-1}}$, definition implies that $x_t = (1 + y_t)x_{t-1}$.

Typically for financial series, the return y_t, does not have a constant variance and highly volatile periods tend to be clustered together.

ARCH (Engle (1982)), GARCH (Bollerslev (1986)) models and other volatility models describe behavior of y_t series.

GARCH(p,q) model was developed by Bollerslev (1986). For conceptual origins, technical details, properties, classical estimates (conditional ML) see ex. Tsay (2005).

(A) $$y_t = \sigma_t \varepsilon_t$$

(B1) $$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2$$ ARCH(1)

(B2) $$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + ... + \alpha_m y_{t-m}^2$$ ARCH(m)

(B3) $$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$ GARCH(1,1)

(B4) $$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{m} \alpha_j y_{t-j}^2 + \sum_{j=1}^{r} \beta_j \sigma_{t-j}^2$$ GARCH(m,r)

where $\alpha_i \geq 0$, $1 \leq i \leq m$, $\beta_i \geq 0$, $1 \leq i \leq r$ and $\alpha_0 > 0$, ε_t is (typically) standard Gaussian white noise. We assume $\sum_{i=1}^{m} \alpha_i + \sum_{i=1}^{r} \beta_i < 1$ for a strict stationary of the process.

APPLICATIONS OF THE GARCH MODELS IN THE ECONOMICS

(*) Better theoretical understanding of an investment, inflation process? NOT

(**) Better abilities to foresee a crisis on a certain financial market NOT

(***) A practical need of forecasting of a future risk of a stock having its history YES

PRACTICAL NEEDS:

To give an insight to the economist into the relation between future volatility of a priced investment and observed volatility and observed value of the return.

$$\hat{\sigma}^2_{t+1} = \hat{\alpha}_0 + \sum_{j=1}^{m} \hat{\alpha}_j y^2_{t+1-j} + \sum_{j=1}^{r} \hat{\beta}_j \hat{\sigma}^2_{t+1-j},$$

To discover a general tendency, to provide robust indications of a development direction in a changing situation on a market rather than to provide a precise asymptotic theory.

Models usually estimated by maximum conditional likelihood assuming that distribution of one observation conditionally to the past is normal – are very sensitive to the presence of outliers.

ROBUST PROPOSITIONS: a variety of methods, for overview see ex. Muler, Yohai (2007)

- Constrained M- estimates Mendes & Duarte (1999)
- Diagnostic procedures for detecting outliers Careno, Pena & Ruiz (2001)
- S – estimates Sakata & White (1998)
- Conditional distribution given the past is a heavy tailed distribution instead normal distribution.

Propositions based on predictors of the conditional variance are very sensitive to outliers, they are rather low efficiency.

RECENT PROPOSITIONS: Constrained M-estimators proposed by Muler & Yohai (consistency and asymptotic normality) and bounded M-estimates / additional mechanism that bounds the propagation of the effect of one outlier on the subsequent predictors of the conditional variances / - also consistent and asymptotically normal, robust to small fraction of AO.

It is well known that GARCH(m,r) admits a non – Gaussian ARMA(max{m,r},r) model for the squared process y_t^2 (see ex. Shumway & Stoffer (2006) for details).

In case ARCH(1) we have

$$y_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + v_t,$$

In case GARCH(1,1) we have

$$y_t^2 = \alpha_0 + (\alpha_1 + \beta_1) y_{t-1}^2 + v_t - \beta_1 v_{t-1}, \text{ (*)}$$

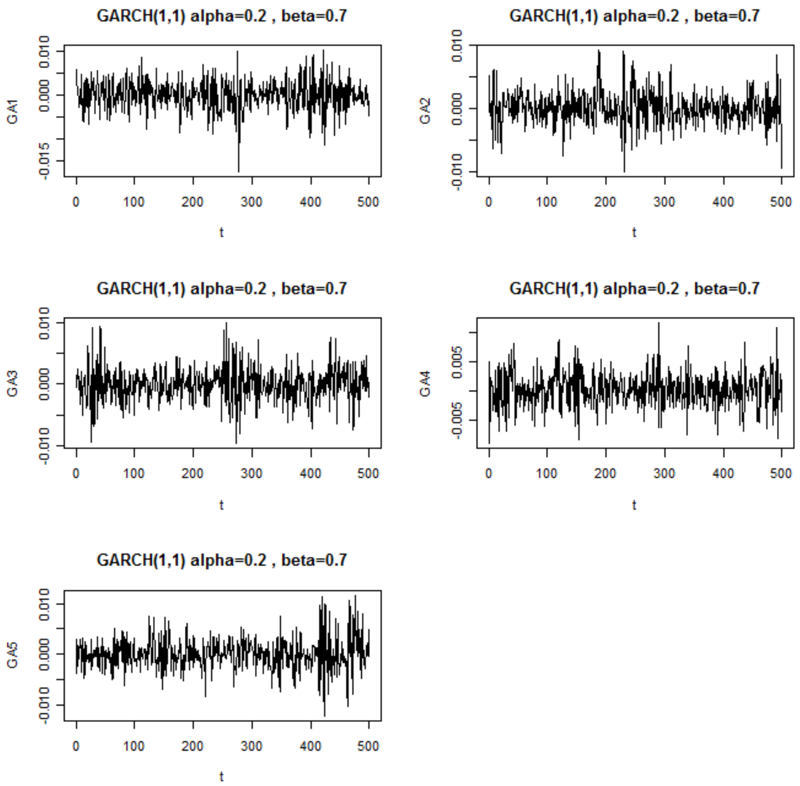
where $v_t = \sigma_t^2(\varepsilon_t^2 - 1)$, $\varepsilon_t \sim N(0,1)$ so $\varepsilon_t^2 - 1$ is a shifted χ_1^2 random variable .

Standard ARMA based identification and inference procedures may be directly applied to the above process, although the heteroscedasticity in the innovations, $\{v_t\}$, render such an approach **inefficient**.

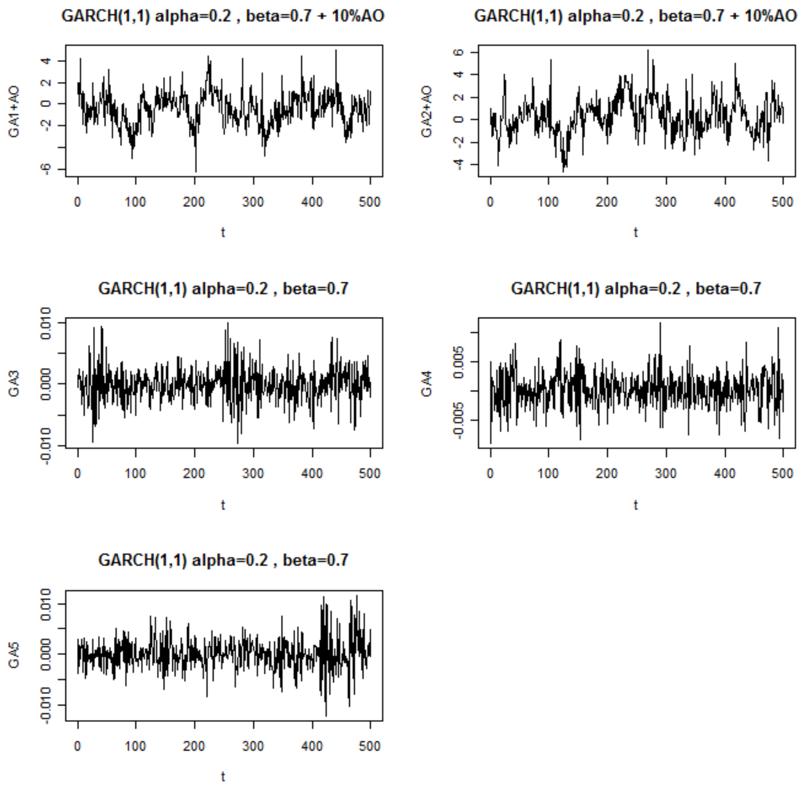
BUT IN CASE OF AN EXISTENCE OF OUTLIERS EFFICIENCY OF THE STANDARD METHODS DECREASES DRAMATICALLY. MAXIMUM REGRESION DEPTH ESTIMATE CAN EASILY COPE WIT THE HETEROSCEDASTICITY OF THE INNOVATIONS

PROPOSITION 2: Let $\mathbb{X} = \{x_1, x_2, ..., x_T\}$, $2 < T$, denote a time series generated by GARCH(1,1) model containing AO outliers. We obtain an estimate of the model parameters applying proposition 2 to the ARMA(1,1) based estimation of squared process (*).

Simulated trajectories from GARCH (1,1) model, trajectories without additive outliers.



Simulated trajectories from GARCH (1,1) model, trajectories without and with 10% of additive outliers.



STATISTICAL PROPERTIES OF THE PROPOSITION:

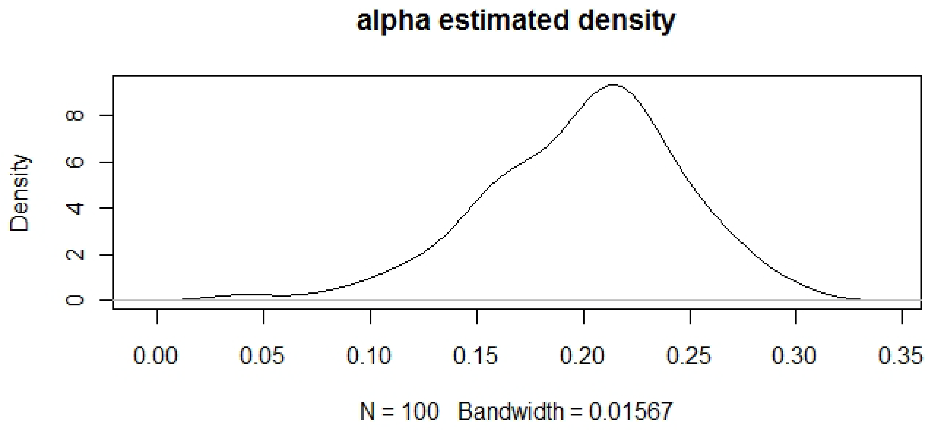
Various simulated data sets /data containing <15% additive outliers indicate on:

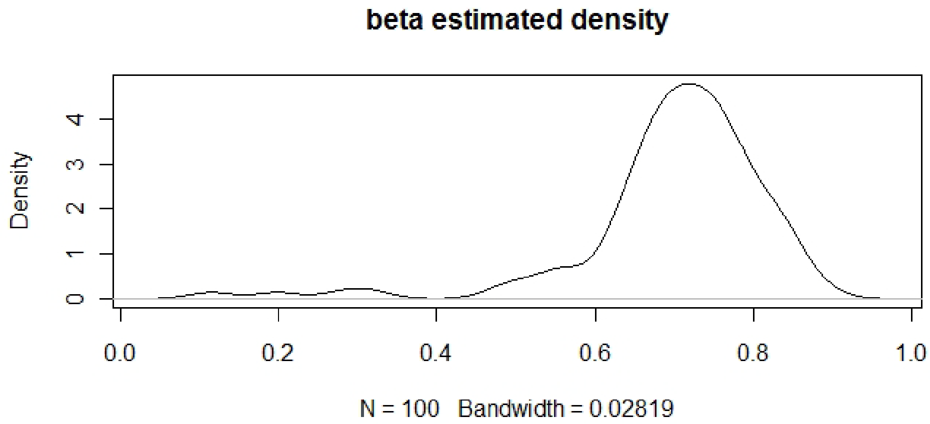
- Good breakdown point of a procedure >5%
- Satisfying efficiency at GARCH(1,1) conditional normal model
- Stability of estimation
- Good performance in case of skewness of errors

VERY GOOD PROPERTIES IN CASES OF SIMPLE MODELS

GARCH(1,1) and sample sizes 100 observations

Kernel density estimation of the proposed parameters estimators of the GARCH(1,1) with $\alpha_1 = 0.2$, $\beta_1 = 0.7$. Two of each five simulated trajectories contained 10% of the AO.





FUTURE PERSPECTIVES - MIZERA & MULLER LOCATION – SCALE DEPTH

Mizera (2002) proposed a general theoretical approach of halfspace depth; Mizera & Muller (2004) proposed a location – scale depth.

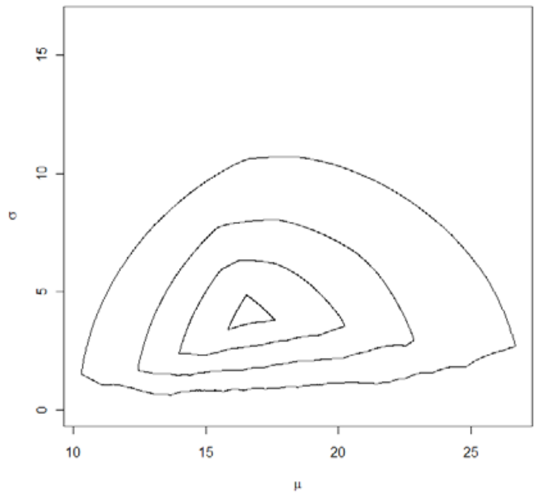
a) A likelihood–based approach enabling to incorporate a priori knowledge into a depth based estimation procedure via known density function.

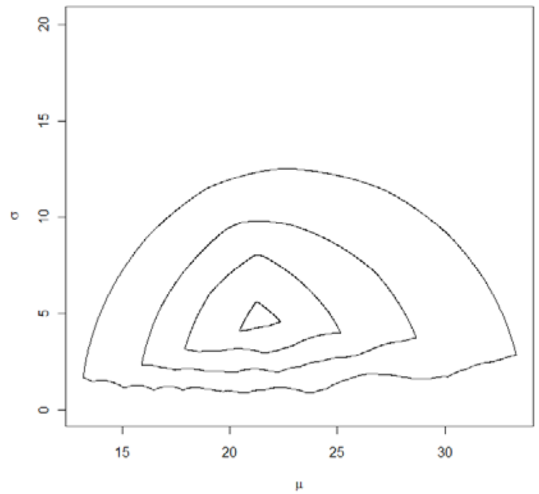
b) A variety of economical application **of joint analysis of location and scale**

- risk (sd of return) and expected return
- dispersion of incomes (energy of system) and average income (wealth of system)
- **robust estimators of GARCH, M-GARCH models**

Ex: An evaluation of an agent's position in a community depends on his distance to a center (location) and a dispersion of positions (scale) and his belief as to random mechanism generating positions (density).

Unemployment rate in polish subregions - powiats





Unemployment 2000

StudentMed=(16.6;3.97) ,

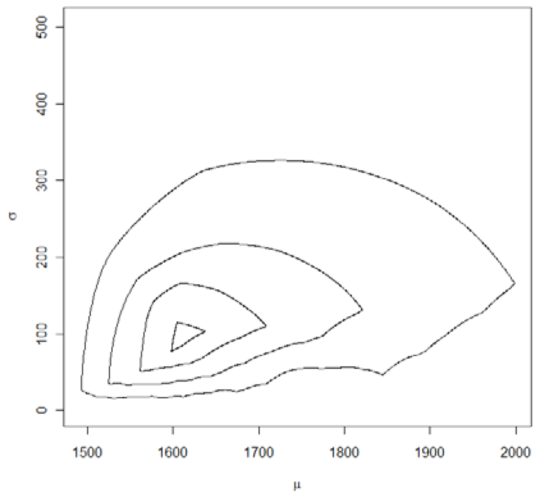
$$\bar{x} = 17.75, \; sd = 6.55$$

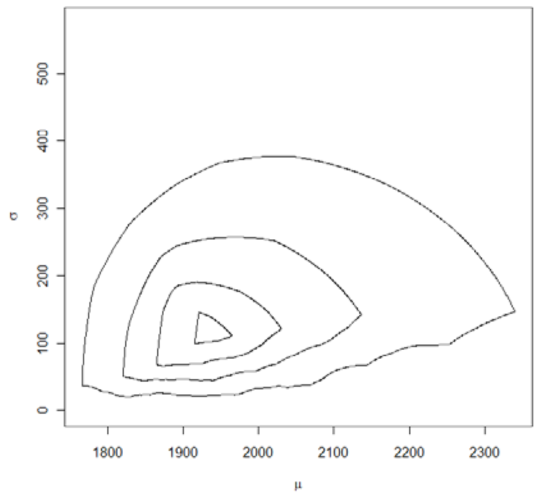
Unemployment 2005

StudentMed=(21.46;4.89),

$$\bar{x} = 22.42, \; sd = 7.72$$

Average salary in polish subregions - powiats





Average salary 2000

StudentMed=(1613.5;99.7),

$$\bar{x} = 1706.16, \; sd = 251.2$$

Average salary 2005,

StudentMed=(1931.43;112.85),

$$\bar{x} = 2026.36, \; sd = 301$$

2000 year

2005 year

CONCLUSIONS

Although at present an exact small sample theory or an asymptotic theory of the proposed ARMA and GARCH models estimators is not known for the author – the simulation studies indicate on their promising statistical properties.

Propositions are simple and user friendly. They are appropriate for the robust economic analysis purposes, especially in cases of simple AR(1), ARMA(1,1), ARCH(1), GARCH(1,1) cases.

Acknowledgements: The author was partially supported by the N N111 193036 grant from the Polish Ministry of Science and Higher Education.

SELECTED REFERENCES

BAI, Z. D., & HE, X. (1999). Asymptotic distributions of the maximum depth estimators for regression and multivariate location. The Annals of Statistics (27), 1616 - 1637.

LOPEZ-PINTADO, S. and ROMO J. (2006): Depth-based classification for functional data : In: Liu R.Y., Serfling R., Souvaine D. L. (Eds.):Series in Discrete Mathematics and Theoretical Computer Science, AMS, vol. 72, 103 - 119.

MARONA,R. A., MARTIN R. D., YOHAI V. J. (2006): Robust Statistics – Theory and Methods.John Wiley \& Sons, Chichester.

MULER, N.,YOHAI, V. J. (2007): Robust estimates for GARCH models, Technical Report Instituto de Calculo Facultaded de Cinecias Exactas y Naturales Universidad de Buenos Aires.

MULER, N., PENA, D. and YOHAI V. J. (2009) : Robust estimation for ARMA models, Annals of Statistics 37 (2),816-840.

ROUSSEEUW, P. J., HUBERT, M. (1998) : Regression Depth, Journal of the American Statistical Association, 94, 388 - 433.

SERFLING, R. (2006): Depth Functions in Nonparametric Multivariate Inference: In: Liu R.Y., Serfling R., Souvaine D. L. (Eds.):Series in Discrete Mathematics and Theoretical Computer Science, AMS, vol. 72, 1 - 15.

VAN AELST, S., ROUSSEEUW, P. J. (2000):Robustness Properties of Deepest Regression, J. Multiv. Analysis, 73, 82-106.

ZUO, Y. (2003). Projection Based Depth Functions and Associated Medians. The Annals of Statistics , 31 (5), 1460 - 1490.

ZUO, Y., & SERFLING, R. (2000). Structural properties and convergence results for contours of sample statistical depth functions. The Annals of Statistics (28), 483 - 499.