

Chart 3. Synthetic indicator for trade and trade output (year to year)

Conclusion:

The synthetic indicators created for Ukrainian economy sectors on the basis of BTS indicators much better reflect the tendencies of macro indicators for industry and construction than confidence indicators suggested for EU countries. The comparison of trends for 1997-2003 allows to affirm that the proposed synthetic indicators may be used for the analysis and short-term forecasting of the above sectors.

Bibliography

1. European Economy. Supplement B. Business and Consumer Survey Results. Special Edition. — July 1991.
2. European Economy. The joint harmonised EU programme of business and consumer surveys. Reports and studies. — №6, 1997.
3. Пугачова М. В. Використання синтетичних індикаторів ділового клімату для моделювання макропоказників. // Щорічник наукових праць «Проблеми статистики». — Київ. — 2002. — № 4. — С. 163—168.
4. Charpin F., Peleraux H., A Leading Indicator for France. — Materials of 25 CIRET Conference (Paris). — October, 11—14, 2000.

Daniel Kosiorowski,
Cracow University of Economics,

DATA DEPTH CONCEPT IN MULTIVARIATE TIME SERIES

1. Introduction

Treating the socio-economic reality as a realization of some multi-dimensional stochastic model recently gains increasingly larger acceptance among both the theoretical and the practicing economists. Such models often approximate the real socio-economic systems much better than their one-dimensional equivalents. Its' construction allows for conducting holistic discourses and for making still more challenging investigative hypotheses. Usefulness of such models in explaining the socio-economic reality is dependent to a large extent upon the multi-dimensional statistical methods used for estimation and choosing specific proposition. Application of classic methods of multidimensional analysis is subject to multidimensional normality of the considered system. Such assumption is usually not met by the objects existing in reality, what is particularly visible in the area of finance¹¹. The prac-

¹¹ Theoretical background on modeling the reality in this context along with the references to particularly spacious empirical research works might be found in e.g.: Bouchaud, J-P., Potters, M. (2000).

tice of economic researches provides sense of searching for non-parametric methods of multidimensional statistics. Methods, which are useful for small samples, and at the same time, reasonably resistant to corrupt data.

From such perspective, special attention should be paid to a set of methods derived from the studies of Tukey¹ called the data depth concept. These methods offer a non-parametric alternative to classic methods. They gain importance, especially as important results appear concerning properties of the estimators² used and the approximate computational methods³, and in the face of simplicity of graphic presentation and legibility of results⁴. The concept of the depth of data is a form of generalization of the methods of one-dimensional statistics basing on order statistics, i.e. the quintiles. For example, a multidimensional median inherits after its one-dimensional equivalent the property of resistance against the outlying cases.

That concept allows for introducing the partial order relation based on the «departure from center» principle. This provides then possibility of examining the relative position of an individual analyzed, with regard to a multidimensional feature, against the central individual, what is particularly important in studies of human behaviour⁵. That concept allows for making arrangements according to the principle of the amount of dispersion around some common center of random vectors, what can be useful for studying the portfolios' risk⁶.

The goal of this work is to present and empirically study the functionality of the data depth concept's aspects particularly interesting from the practical point of view. The empirical examples concern the comparisons of portfolios consisting of companies composing the WIG20 and TECHWIG indices. We compare the classic estimators of location parameters for the multidimensional distribution of probability, namely the vector of averages and the vector of one-dimensional medians with the estimators of the maximum Mahalanobis depth and the maximum projection depth (the multidimensional median). That seems to be an important feature of our work, as the majority of dissertations concerning the application of the data depth concept uses only simulations.

2. Data Depth Concept

¹ Valuable data relating to the historical background of the issue can be found in e.g.: Liu R. Y., Parelius J. M., Singh K. (1999).

² The details of the issue are described in e.g.: Zuo, Y., Serfling, R. (2000 c).

³ The details of the issue are described in e.g.: Dyckerhoff, R. (2004).

⁴ Particularly valuable in this context is the work of Liu R. Y., Parelius J. M., Singh K. (1999).

⁵ An attempt to clarify this issue was made in the work of Kosiorowski, D. (2004).

⁶ Some practically useful results relating to this question contains e.g.: Mosler, K. (2002).

The notion of the depth of a point $x \in \mathfrak{R}^d$, $d > 1$, being a realization of some d-dimensional random vector X with the probability distribution F , is introduced basing on a special function called the **depth** or the **depth function**^{17,18}. The function of depth assigns to every point a real positive number from interval $[0,1]$ being the measure of that point's «centrality» in regard to the F distribution. Usually, a point, for which the function of depth assumes the maximum value is called ***d-dimensional median*** (an average, if there are more such points than one). In case we do not know the form of the F distribution, but we have an n-element sample with X_1, \dots, X_n , then we can replace the F distribution with its empirical version of F_n . We can get a basic perception of the depth function's properties, by taking a look at the forms of the most often used depth functions. In our considerations, we treat every element of the X_i as a column vector of $d \times 1$. We also assume, that F is absolutely continuous.

A. Mahalanobis depth (MD) at x w.r.t F is defined to be

$$MHD(F; x) = \frac{1}{1 + (x - \mu_F) \Sigma_F^{-1} (x - \mu_F)}, \quad (1)$$

μ_F, Σ_F — mean vector and dispersion matrix F , respectively.

The sample version of MD is obtained by replacing μ_F, Σ_F with their sample estimates:

$$MHD(F_n; x) = \frac{1}{1 + (x - \bar{x}) S^{-1} (x - \bar{x})}. \quad (2)$$

B. The half-space depth (HD) (Tukey depth) at x w.r.t F is defined to be:

$$HD(F; x) = \inf_H \{P(H) : H \text{ is a closed half-space in } \mathfrak{R}^d \text{ and } x \in H\}$$

The sample version of HD ($F; x$) is HD ($F_n; x$), F_n denotes empirical distribution of the sample $\{X_1, \dots, X_n\}$.

C. The simplicial depth (SD) at x w.r.t. F is defined to be

$$SD(F; x) = P_F \{x \in S[X_1, \dots, X_{d+1}]\}, \quad (3)$$

$S[X_1, \dots, X_{d+1}]$ is closed simplex formed by $d + 1$ observations from F .

¹⁷ Interesting theoretical considerations relating to the issues can be found in e.g.: Zuo, Y., Serfling, R. (2000a).

¹⁸ A proposal of axiomatization of the depth concept contains e.g.: Zuo, Y., Serfling, R. (2000a), Dyckerhoff, R. (2004).

The sample version of SD ($F; x$) is obtained by replacing F in SD ($F; x$) by empirical distribution F_n or alternatively by computing the fraction of sample random surlices containing the point x

$$SD(F_n; x) = \left(\binom{n}{d+1} \right)^{-1} \sum_{1 \leq i_1 \leq \dots \leq i_{d+1} \leq n} I_{(x \in [x_{i_1}, \dots, x_{i_{d+1}}])}. \quad (4)$$

D. Simplicial volume depth (SVD)

$$SVD^\alpha(F; x) = \left(1 + E \left[\left(\frac{\text{vol}([x, X_1, \dots, X_d])}{\sqrt{\det(\Sigma)}} \right)^\alpha \right] \right)^{-1}, \quad (5)$$

where: vol — volume, E — expected value.

E. Projection depth (PD) define the outlyingness of a point x to be the worst case outlyingness of x with respect to one dimensional median in any one-dimensional projection that is:

$$PD(F; x) = \left(1 + \sup_{\|u\|=1} \frac{|u^T x - \text{Med}(u^T X)|}{MAD(u^T X)} \right)^{-1}, \quad (6)$$

where X has distribution F , Med denotes the univariate median, MAD denotes the univariate median absolute deviation $MAD(Y) = \text{Med}(|Y - \text{Med}(Y)|)$.

The set $\{x \in \mathfrak{R}^d : D(x) = \alpha\}$ is called **α level set** or **contour of depth α**

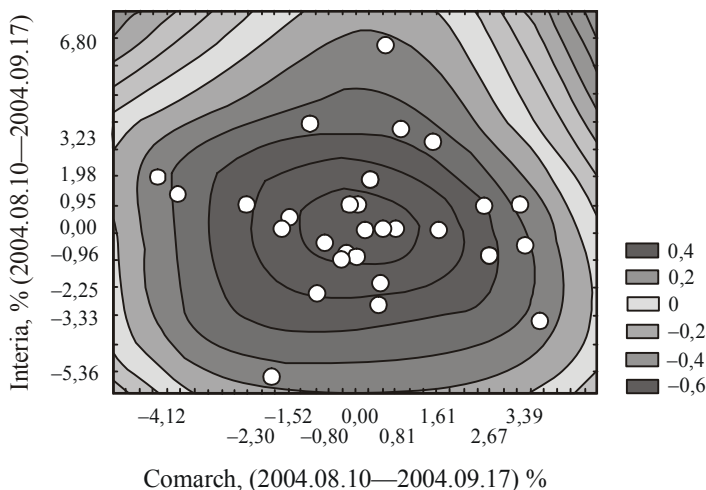
The set $D_\alpha(X) = \{x \in \mathfrak{R}^d : D(x | X) \geq \alpha\}$ is referred to as the region enclosed by contour of depth α , **α -trimmed region** or **α -central region**.

Some of the above introduced ideas are illustrated in pictures 1 to 8 showing the behaviour of two portfolios consisting of a pair of shares of companies being a part of the TECHWIG and the WIG20 indices. Fig. 1 shows a contour diagram of an empirical function of the projection depth for daily changes of share prices for a portfolio of two companies being a part of the TECHWIG index: *Comarch* and *Interia* in the period from 2004.08.10 to 2004.09.17. Fig. 2 illustrates the same portfolio in the same period, however the diagram is prepared for the Mahalanobis depth. Fig. 3 exhibits empirical projection depth for the portfolios considered in fig. 1 and 2 but for the period from 2004.07.01 to 2004.08.09. Fig. 4 was prepared for the same portfolio as in fig. 3, the same period, but for the Mahalanobis depth instead of empirical projection depth. Fig. 5 represents a contour graph of the empirical function

of projection depth for daily share price changes for the portfolio consisting of shares of two companies (being a part of WIG20 index): *Pekao* and *TPSA* and the period from 2004.08.10 to 2004.09.17. Fig. 6 displays a contour graph of the same portfolio as in fig. 5, but prepared for the Mahalanobis depth. Fig. 7, and fig. 8 demonstrate the graphs of: projection depth and the Mahalanobis depth respectively for the portfolio of *Pekao* and *TPSA*, but prepared for the period from 2004.07.01 to 2004.08.09. Table 1 includes the minimum and the maximum values for the empirical functions of depth for the considered portfolios. Table 2 shows the values of numerical characteristics of the considered portfolios. We notice, that usually there are differences between the estimates of location parameters' calculated using the vector of means from the sample, those figured out with the use of the vector consisting of one-dimensional medians and the ones computed using two-dimensional median. It is also easy to see, that the values of the two-dimensional medians differ depending on the notion of the depth used. The quality of the location parameter's estimate can be inferred from one-dimensional values of skewness and kurtosis of daily share price changes or directly from figures 1 to 8.

2004.08.10—2004.09.17

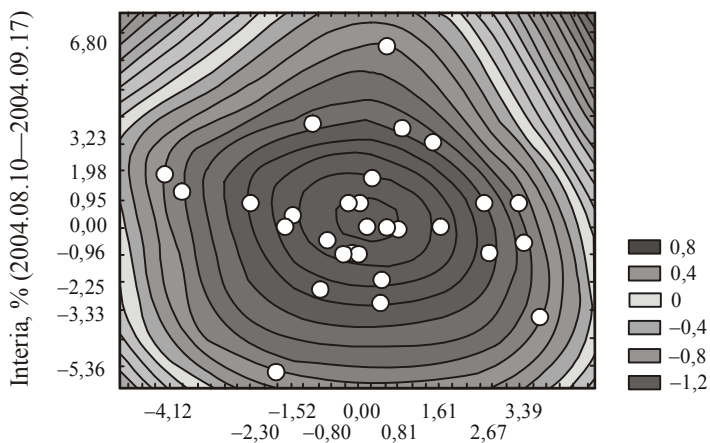
Wyglądanie najmniejszych kwadratów wazone odległościami



Pic. 1: Contur plot: *Interia* & *Comarch*
Projection depth 2004.08.10—2004.09.17

2004.08.10—2004.09.17

Wygladzenie najmniejszych kwadratyw wazone odleglosciami

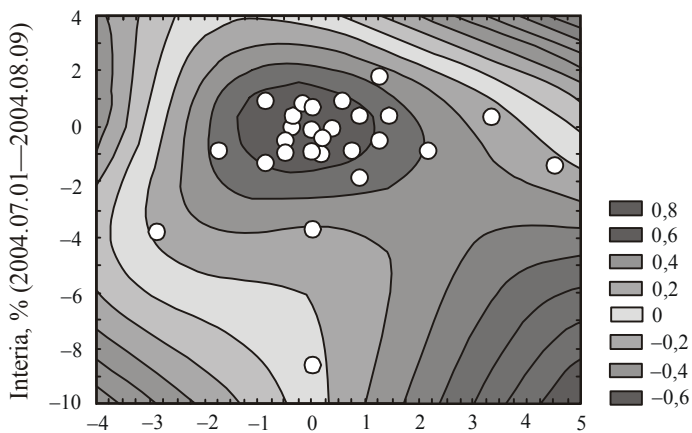


Comarch, (2004.08.10—2004.09.17) %

Pic. 2: Contur plot: *Interia & Comarch*
Mahalanobis depth 2004.08.10—2004.09.17

2004.07.01—2004.08.09

Wygladzenie najmniejszych kwadratyw wazone odleglosciami

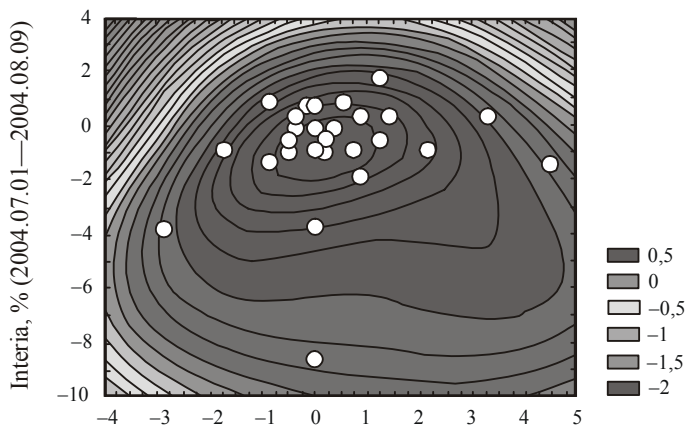


Comarch, % (2004.07.01—2004.08.09)

Pic. 3: Contur plot: *Interia & Comarch*
Projection depth 2004.07.01—2004.08.09

2004.07.01—2004.08.09

Wygladzenie najmniejszych kwadratyw wazone odleglosciami

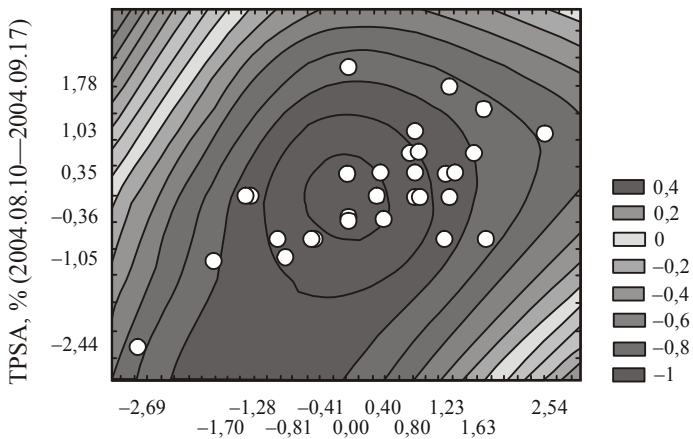


Comarch, % (2004.07.01—2004.08.09)

Pic. 4: Contur plot: *Interia* & *Comarch*
Mahalanobis depth 2004.07.01—2004.08.09

2004.08.10—2004.09.17

Wygladzenie najmniejszych kwadratyw wazone odleglosciami

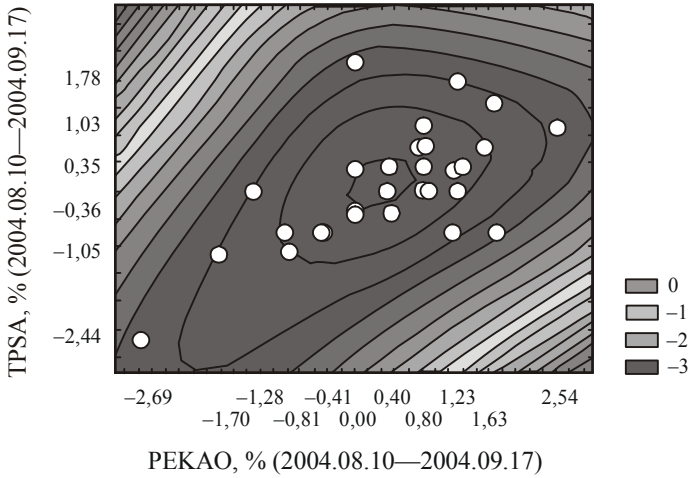


PEKAO, % (2004.08.10—2004.09.17)

Pic. 5: Contur plot: *Pekao* & *TPSA*
Projection depth 2004.08.10—2004.09.17

2004.08.10—2004.09.17

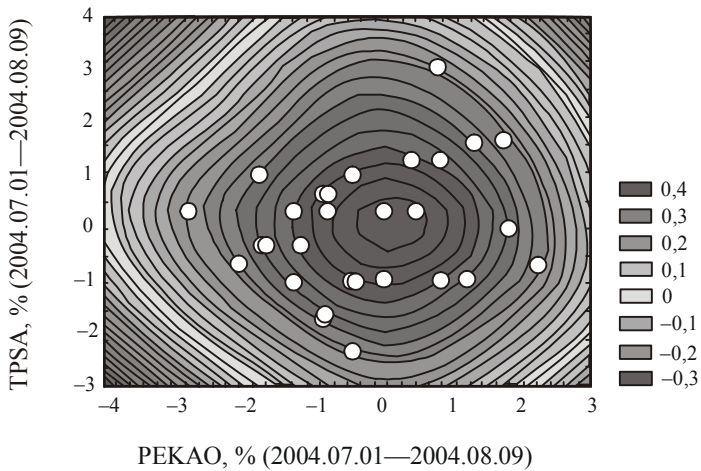
Wyglądanie najmniejszych kwadratyw wazone odleglosciami



Pic. 6: Contur plot: *Pekao & TPSA*
Mahalanobis depth 2004.08.10—2004.09.17

2004.07.01—2004.08.09

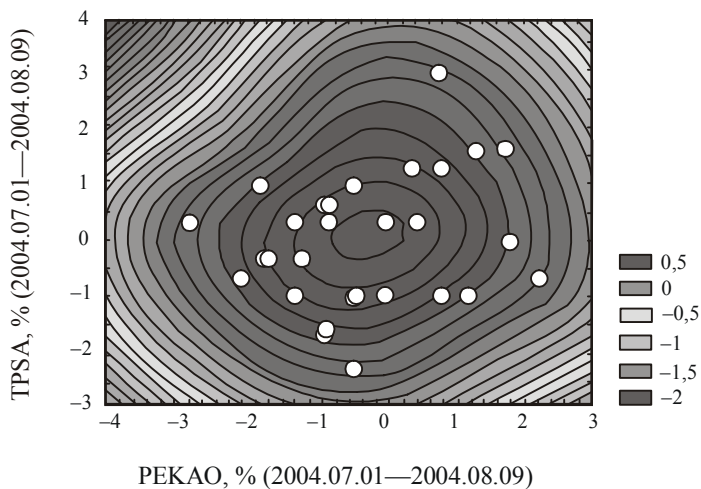
Wyglądanie najmniejszych kwadratyw wazone odleglosciami



Pic. 7: Contur plot: *Pekao & TPSA*
Projection depth 2004.07.01—2004.08.09

2004.07.01—2004.08.09

Wygladzenie najmniejszych kwadratyw wazone odleglosciami



Pic. 8: Contur plot: *Pekao & TPSA*
Mahalanobis depth 2004.07.01—2004.08.09

Tab. 1

MIN AND MAX OF SAMPLE DEPTH FUNCTION

	Minimum	Maksimum
Comarch i Interia 08.10—09.17 PD	0,136	0,857
Pekao i TPSA 08.10—09.17 PD	0,157	0,648
Comarch i Interia 08.10—09.17 MHD	0,110	0,990
Pekao i TPSA 08.10—09.17 MHD	0,097	0,984
Comarch i Interia 07.01—08.09 PD	0,075	1,000
Pekao i TPSA 07.01—08.09 PD	0,199	0,525
Comarch i Interia 07.01—08.09 MHD	0,056	0,984
Pekao i TPSA 07.01—08.09 MHD	0,126	0,895

ONE AND TWO — DIMENSIONAL DESCRIPTIVE CHARACTERISTICS

	2004.08.10—2004.09.17				2004.07.01—2004.08.09			
	Comarch&Interia		Pekao&TPSA		Comarch&Interia		Pekao&TPSA	
Mean vector	[0,188; 0,231]		[0,370; 0,073]		[0,315; -0,682]		[-0,314;-0,005]	
One dimensional Median	0,085	0	0,630	0	0	-0,445	-0,440	-0,165
Standard deviation	1,908	2,346	1,172	0,917	1,367	1,933	1,254	1,176
One dimensional skewness	-0,099	0,397	-672	-0,146	0,905	-2,568	0,261	0,410
One dimensional kurtosis	0,278	1,694	0,322	1,137	3,092	9,065	-0,543	0,299
Two dimensional Median PD	[0,17; 0]		[0; 0,35]		[0; 0]		[0; 0,33]	
Two dimensional Median MD	[0,17; 0]		[0,4; 0]		[0,18;-0,9]		[0; 0,33]	

3. Depth vs. Depth Plot

An unquestionable advantage of the data depth concept is the wealth of graphic methods available for comparing two multidimensional probability distributions. Among the many¹⁹ ideas describing special types of differences between the random vectors, one, relatively simple graph called depth vs. depth (DD — plot) deserves some particular attention. The depth vs. depth plot compares the values of the depth function for the point $x \in \mathfrak{R}^d$ under two probability distributions of F and G . Such plot, by the way, is an entirely non-parametric method of comparing two multidimensional probability distributions.

$DD(F,G)$ is the subset of \mathfrak{R}^2 defined as follows:

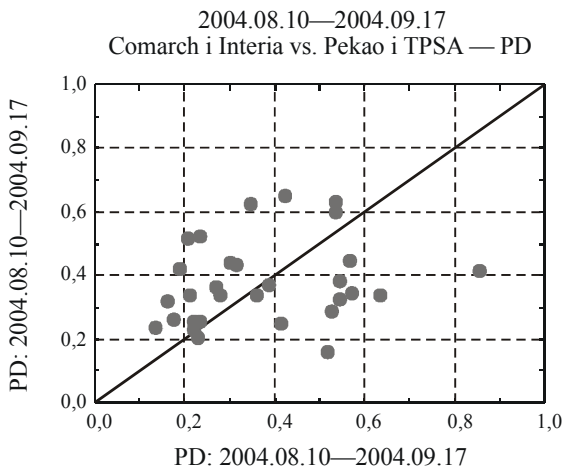
$$DD(F, G) = \{ (d(F; x), d(G; x)) : x \in \mathfrak{R}^p \},$$

If $F = G$, the DD-plot is a subset of the line segment with endpoints $(0,0), (1,1)$ — diagonal of the plot, otherwise it fills a region of \mathfrak{R}^2 whose shape is often informative about differences among F and G .

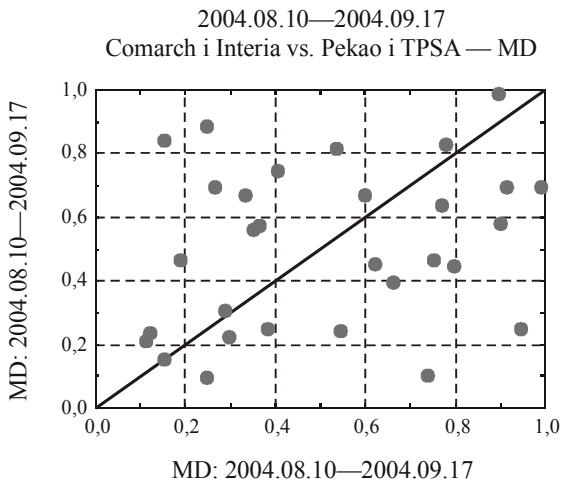
Figures 9 and 10 illustrate comparisons of two portfolios: one consisting of shares of *Comarch* and *Interia* and the other of *Pekao* and *TPSA* for the period from 2004.08.10 to 2004.09.17. Figure 9 shows comparisons prepared with the use of the projection depth, while figure 10 — with the use of the Mahalanobis depth. A difference as to the location characteristics can be observed, and is directly visible

¹⁹ About the plurality of possibilities in this sphere informs e.g. the work of Liu R. Y., Parlius J. M., Singh K. (1999).

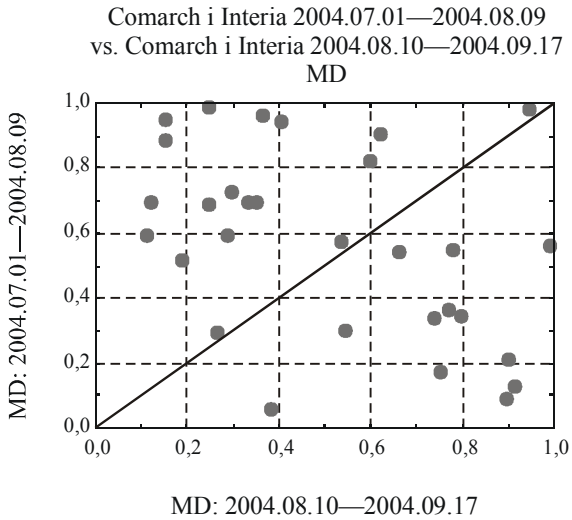
while looking at table 2. Similarly figures 11 and 12 illustrate comparisons of the above mentioned portfolios, but for the period from 2004.07.01 to 2004.08.09. In figure 13 and figure 14 the portfolios of *Pekao* and *TPSA* are compared in two periods: from 2004.08.10 to 2004.09.17, and from 2004.07.01 to 2004.08.09 with the use of projection depth and Mahalanobis depth respectively.



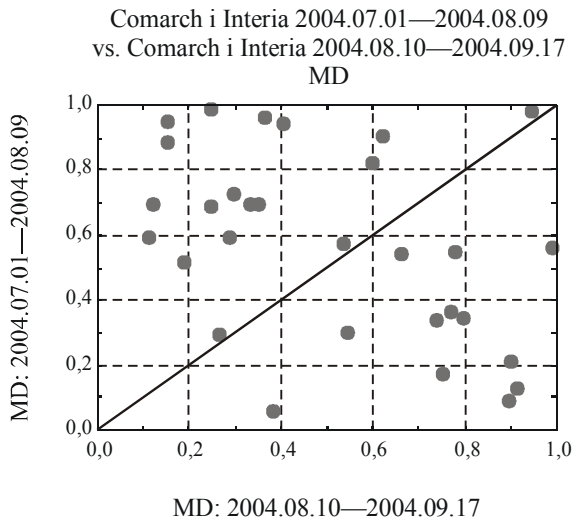
Pic. 9: DD plot *Comarch & Interia* vs. *Pekao & TPSA* Projection Depth 08.10—09.17



Pic. 10: DD plot *Comarch & Interia* vs. *Pekao & TPSA* Mahalanobis depth 08.10—09.17

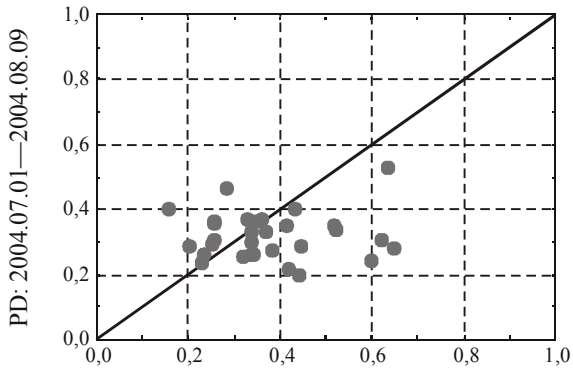


Pic. 11: DD plot *Comarch & Interia* vs. *Pekao & TPSA*
Projection depth 07.01—08.09



Pic. 12: DD plot *Comarch & Interia* vs. *Pekao & TPSA*
Mahalanobis depth 07.01—08.09

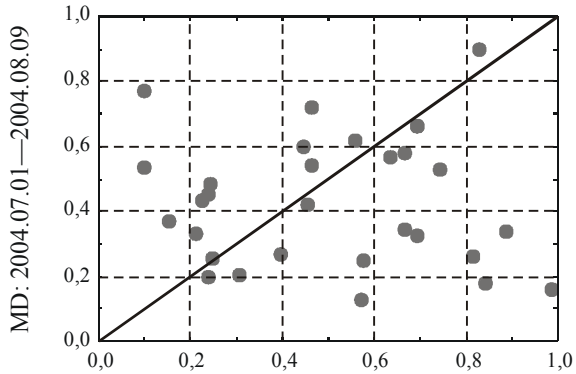
Pekao i TPSA 2004.07.01—2004.08.09
 vs. Pekao i TPSA 2004.08.10—2004.09.17
 PD



PD: 2004.08.10—2004.09.17

Pic. 13: DD plot 07.01—08.09 *Pekao* & *TPSA* vs. 08.10—09.17 *Pekao* & *TPSA* Projection depth

Pekao i TPSA 2004.07.01—2004.08.09
 vs. Pekao i TPSA 2004.08.10—2004.09.17
 MD



MD: 2004.08.10—2004.09.17

Pic. 14: DD plot 07.01—08.09 *Pekao* & *TPSA* vs. 08.10—09.17 *Pekao* & *TPSA* Mahalanobis depth

4. Computational aspects.

Inclination to accept some new statistical method should, with no doubt, depend upon whether that method can be characterized by better characteristics than the one used so far. It, certainly, depends²⁰ also on the ease of its use. That ease can refer to both the interpretation, and the calculations. The Mahalanobis depth can easily be calculated with the use of a statistical suite offering the T2 Hotelling test, and the calculations related to the simplicial depth can be reduced to solving certain number of linear equations. However, according to the author, it is worth to call special attention to the results raised in Dyckerhoff, R. (2004), enabling easy approximation of certain class of function of depth in simulation experiments; the approximation which uses one-dimensional functions of depth for calculating the multidimensional depth²¹. The idea resolves itself into the following statement: *if we accept, that a point is «central» in relation to certain multidimensional probability distribution, when all its one-dimensional projections are «central» in relation to one-dimensional projection of distribution, then we can define the multidimensional depth of the point as the minimum of all such one-dimensional projections.*

5. Summary

The quality of our comprehension of the socio-economic reality depends to a great extent upon the ability to interpret statistical data generated by multidimensional socio-economic systems existing in certain place at certain time. Interpretation is particularly difficult, when the mechanism regulating behaviour of a system slips away from the elegant theoretical models. And an attempt to interpret the reality is valuable in every respect, as the quality our comprehension of the world we live in translates into its characteristics. The data depth concept, presented in this work, constitutes according to its author, an interesting tool allowing to take up wide class of issues in research, in case of which the classic methods of multidimensional statistics fail. The non-parametric character of these methods deserves our attention, as well as relative ease of the results' interpretation and the legibility of its graphic presentation. Certain shortcoming comes from the lack of ready-to-use modules for the statistical suites available on the market that would support these methods. However, the data depth concept is a relatively new approach, so we may expect that such modules should appear soon.

²⁰ Particularly important, from the practical point of view, properties of statistical methods can be found e.g. in Zeliš A. (1997).

²¹ Theoretical background is presented in an inspiring way in the work of Zuo, Y. (2003).

Bibliography

1. *Bouchaud, J-P., Potters, M.* (2000). Theory of Financial Risks, From Statistical Physics to Risk Management. Cambridge University Press.
2. *Dyckerhoff, R.* (2004). Data depths satisfying the projection property. *Allgemeines Statistisches Archiv* 88, p. 163—190.
3. *Liu R. Y., Parelius J. M., Singh K.* (1999). Multivariate analysis by data depth: descriptive statistics, graphics and inference. *The Annals of Statistics*, Vol. 27, No. 3, p. 783—858.
4. *Kosiorowski, D.* (2004): Individual Rationality Versus Group Rationality in Statistical Modelling Issues, in: Daniel Baier and Klaus-Dieter Wernecke (eds.): *Innovations in Classification, Data Science, and Information Systems*. Proc. 27th Annual GfKl Conference. Springer-Verlag, Heidelberg-Berlin, p. 239—248.
5. *Mosler, K.* (2002). Central regions and dependency. IWAP2002, Caracas.
6. *Zuo, Y., Serfling, R.* (2000a). General notions of statistical depth function. *The Annals of Statistics* 28, p. 461—482.
7. *Serfling, R.* (2002). Multivariate Symmetry and Asymmetry. preprint, NSF Grant DMS-0103698
8. *Zuo, Y., Serfling, R.* (2000b). Nonparametric notions of multivariate «scatter measure» and «more scattered» based on statistical depth function. *Journal of Multivariate Analysis* 75 p. 62—78
9. *Zuo, Y., Serfling, R.* (2000c). Structural Properties and Convergence Results for Contours of Sample Statistical Depth Functions. *The Annals of Statistics* 28, p.483—499
10. *Zuo, Y.* (2003). Projection based depth functions and associated medians. *The Annals of Statistics* 31, p. 1460—1490
11. *Zelias A.* (1997), *Teoria Prognozy*, PWE, Warszawa.

**Ruslan Motoryn,
Yuri Hemych,**

Kiev National University of Economics

PREDICTABILITY OF MACROECONOMIC INDEXES: EFFECTIVENESS OF STATISTICAL METHODS' APPLICATIONS

1. Introduction

This paper reviews research on the effectiveness of most important statistical methods that are designed for forecasting of different macroeconomic indexes. In this paper several statistical methods have been compared by most important criterions. It was investigated the optimal forecasting methods for different branches of application. As a result it was developed some modification,