V CONGRES OF POLISH STATISTICS

Alternative approaches to modelling the impact of community well-being on individual well-being – an empirical appraisal

Włodzimierz Okrasa

Statistics Poland

July 7-11, 2025 Warsaw

OUTLINE

I. INTRO - background and problem

II. Measuring complex multidimensional phenomena over time

- community well-being – pragmatic approach: research and policy purposes;
- modeling dynamic –'diachronic' approach

→ rationale for *Multivariate Functional Principal Component Analysis/MFPCA*

III. Cross-level interaction of well-being measures - empirical illustrations:

- local deprivation and subsidies to *communes (gminas);*
- comparison of MFPCA and classic PCA approaches; -
- models of community and individual (macro-/micro-) relationships;

→ community cohesion and the role for social capital

→ multilevel modeling in spatial context

→ looking for *causality* - structural modeling.

IV. Spatial aspects of cross-level well-being interaction

- Summary and conclusions

Problem and background

The main objective of this study is to investigate modeling approaches to assess the strength and impact of the relationship (mutual influence) between community well-being and individual/household well-being based on public statistics datasets. While the latter is a challenge in itself - due to the lack of availability of a multi-source analytical database with a hierarchical (nested) data structure - the empirically demonstrated effectiveness of different, problem-specific approaches to analyze the cross-level interaction effect becomes important both from a methodological and (local) development policy perspective.

The complexity of the intertwining methodological and substantive issues involved in analysing the relationship between individual and community well-being, taking into account both temporal and spatial aspects of between-level dynamics, is addressed in two steps:

- First, at the measurement level, the Functional Data measurement approach is employed in the version of *Multivariate Functional Principal Component Analysis* (MFPCA) to deal with multidimensionality and temporality of community development (deprivation) and of individual (residents') subjective well-being.
- Secondly, having constructed a classification of both local communities (communes) and their inhabitants (for a given time period), a spatial perspective is employed – the spatial and place-based effects of community development on the resulting cross-categorization of units are assessed in terms of spatial patterns (autocorrelation and clustering tendency), spatial dependence, and spatial regression.

[*Remarks on*] Conceptualization and operationalization of community well-being (CW-B) in the evaluation policy research context

Two complementary outlooks:

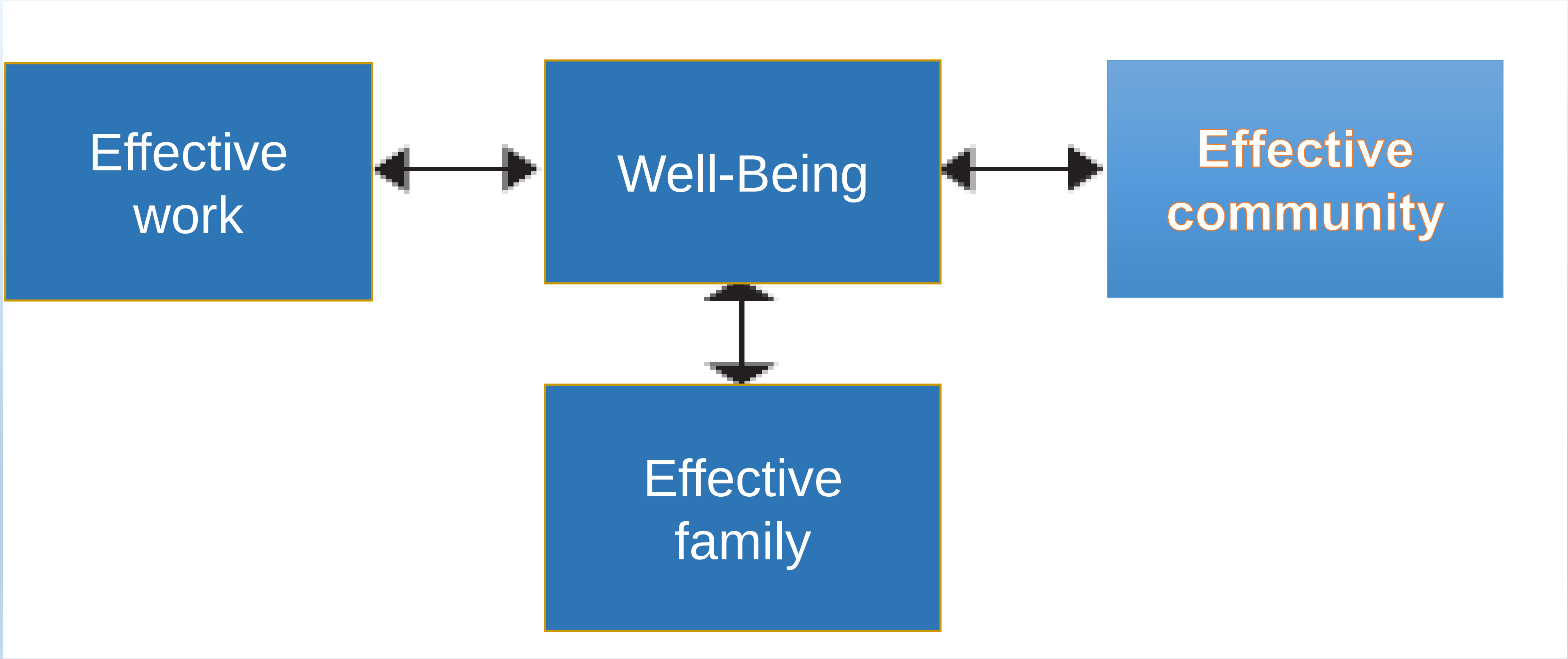
(i) ***methodological*,** starting with an overview of the main approaches (paradigms) to measuring CWB in public statistics, and

(ii) ***analytical***, checking distributional potentials of the CWB indicators for geographic targeting of public resources, including their effects for <u>reducing</u>:

- <u>local deprivation</u> (of *gminas*) and contributing to 'social progress' in the local context (*beyond GDP* as measures of social progress: OECD, 2015), and using dual type measures, objective and subjective, for the CWB;
- <u>inequalities among 'localities'</u> (NUTS5 units/LAU2 *gmina*) - implications for spatial cohesion and community

Key issues in analyzing the relationship between Community and Personal Well-Being: ***measurement – data – models***

- A **well-being measure** is presumed to be generated not only to satisfy formal requirements but primarily to guide policy, especially about local community development.
  - *Local Community*: Any configuration of individuals, families, and groups whose values, characteristics, interests, geography, and/or social relations unite them in some way (e.g., Dreher, 2016)
    → community is defined as the people living in a place such as a *neighborhood*.



Source: World Economic Forum, 2012. *Global Agenda: Well-being and Global Success* .

Methodological framework for analyzing CWB and PWB:

► accounting for *micro – macro interdependence* → *modelling multilevel relationships*

► bringing *space* into the question /equation→ *spatial (dependence) analysis.*

- Modeling multilevel *relationships* – two types of strategies:
  - cross-level *interaction*-focused approach:
    - ***decomposition of variance*** into *within* groups/differences among individuals in a community (level -1) and *between* groups (level-2) reflecting differences across comunities;

      → models for hierarchically structured data – risk of 'ecological fallacy' (Goldstein, 2003(2010); Subramanian, 2009; Sampson 2003)
  - *structural modelling of (causal) mediation mechanisms:*
    - ***decomposing total effect*** of the independent variable ('treatment') into the (natural) *direct* and *indirect* effects (Hong, 2015).

Multidimensional measures of well-being

- dimensionalization / operationalization

according to PCA and FD-PCA - some comparisons

- Multiple-source Analytical Database /MAD:- *a bottom-up approach*

  Local Deprivation and Subjective Well-Being (SWB) - data sources:

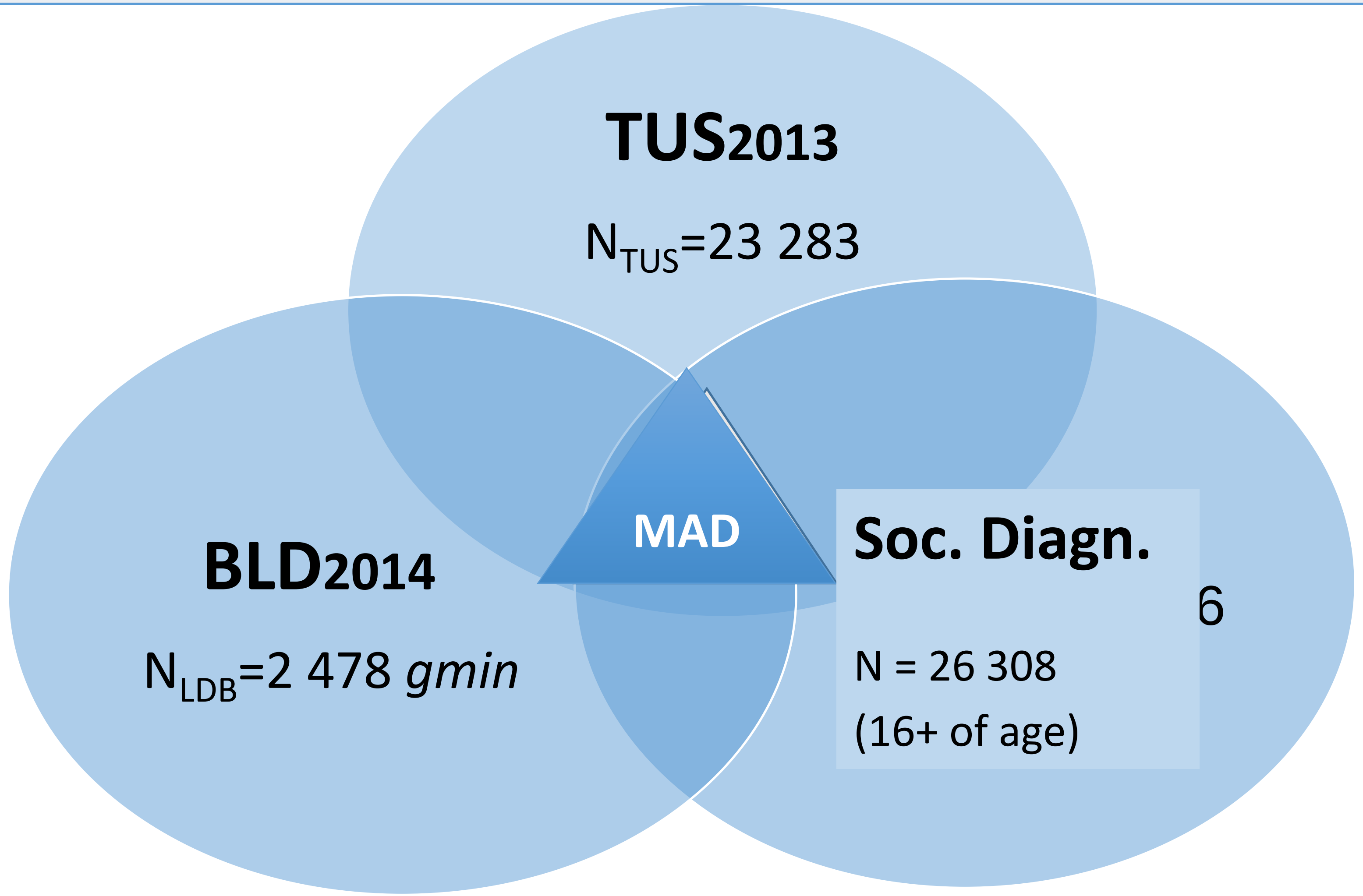
  (i) measures of local community (communes) development/deprivation and the relevant covariates*: Local Data Bank /LDB -Statistics Poland* (years 2004, 2008, 2010, 2012, 2014, 2016) for NUTS5 / LAU2, *gminas*; N = 2 478)

  (ii) subjective well-being measures based on data from nation-wide surveys:

  (a) Social Diagnosis /SD curried out in every other year (2003-2005 -...- 2015) and

  (b) Time Use Survey / TUS 2013, Statistics Poland).

Multiple-source Analytical Database / MAD
– *bottom-up* data integration, with territorial code (KODTERYT) for the commune/municipality (an ‘anchore’)



Measuring local deprivation and personal well-being

- *Multidimensional Index of Local Deprivation (MILD)*

  (i) Classic version: *Confirmatory* Factor Analysis/PCA (single-factor)

  Eleven (pre-selected) domains of deprivation - each characterized by a number of original items: *ecology – finance – economy – infrastructure – municipal utilities – culture – housing – social assistance – labour market – education – health* [altogether 67 items]

  (ii) *Functional* Principal Component Analysis *(FPCA)*

- *Personal Subjective Well-being/SWB* and *Community Subjective Well-being/CSWB*
  - SWB: individual subjective measure based on *Social Diagnosis,* using FPCA
  - SWB: individual quasi-objective - *Time Use Survey* (one-off survey data) *;*
  - CSWB: compositional - subjective*:* self-reported satisfaction with selected aspects of life (*Social Diagnosis*)

Classical principal component analysis (PCA)

(Krzysko et al., 2014)

Suppose we observe a p-dimensional random vector $\boldsymbol{X} = (X_1, X_2, ..., X_p)' \in \mathbb{R}^p$. We further assume that $\mathrm{E}(\boldsymbol{X}) = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{X}) = \boldsymbol{\Sigma}$.

In the first step we seek a variable U_1 in the form

$$U_1 =< \boldsymbol{u}_1, \boldsymbol{X} >= \boldsymbol{u}_1' \boldsymbol{X} = \sum_{i=1}^{p} u_{1i} X_i,$$

having maximum variance for all $\boldsymbol{u} \in \mathbb{R}^p$ such that $< \boldsymbol{u}, \boldsymbol{u} >= 1$.

Let

$$\lambda_1 = \sup_{\boldsymbol{u} \in \mathbb{R}^p} \mathrm{Var}(< \boldsymbol{u}, \boldsymbol{X} >) = \mathrm{Var}(< \boldsymbol{u}_1, \boldsymbol{X} >) = \boldsymbol{u}_1' \boldsymbol{\Sigma} \boldsymbol{u}_1,$$

where $< \boldsymbol{u}_1, \boldsymbol{u}_1 >= \boldsymbol{u}_1' \boldsymbol{u}_1 = 1$.

The random variable U_1 will be called the first principal component, and the vector $\boldsymbol{u}_1$ will be called the vector of weights of the first principal component.

In the next step we seek a v ariable $U_2 =< \boldsymbol{u}_2, \boldsymbol{X} >= \boldsymbol{u}_2' \boldsymbol{X}$ which is not correlated with the first principal component U_1 and which has maximum variance. We continue this process until we obtain p new variables $U_1, U_2, \ldots, U_p$ (principal components).

In general, the kth principal component $U_k =< \boldsymbol{u}_k, \boldsymbol{X} >= \boldsymbol{u}_k' \boldsymbol{X}$ satisfies the conditions:

$$\lambda_k = \sup_{\boldsymbol{u} \in \mathbb{R}^p} \mathrm{Var}(< \boldsymbol{u}, \boldsymbol{X} >) = \mathrm{Var}(< \boldsymbol{u}_k, \boldsymbol{X} >) = \boldsymbol{u}_k' \boldsymbol{\Sigma} \boldsymbol{u}_k,$$

$$< \boldsymbol{u}_{\kappa_1}, \boldsymbol{u}_{\kappa_2} >= \delta_{\kappa_1 \kappa_2}, \qquad \kappa_1, \kappa_2 = 1, ..., k.$$

The expression $(\lambda_k, \boldsymbol{u}_k)$ will be called the kth principal system of the variable $\boldsymbol{X}$ (Jolliffe (2002)).

PCA and FDPCA – *an overview*

It can be shown that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_p$ and $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_p$ are the eigenvalues and corresponding eigenvectors of the covariance matrix $\boldsymbol{\Sigma}$.

In practice this matrix is unknown, and must be estimated from the sample. Let $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ be realizations of the vector $\boldsymbol{X}$.

Then

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\boldsymbol{x}\boldsymbol{x}'.$$

Moreover, let $\hat{\lambda}_1 \geqslant \hat{\lambda}_2 \geqslant \cdots \geqslant \hat{\lambda}_p$ and $\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2, \ldots, \hat{\boldsymbol{u}}_p$ be eigenvalues and corresponding eigenvectors of the matrix $\hat{\boldsymbol{\Sigma}}$.

Then $(\hat{\lambda}_k, \hat{\boldsymbol{u}}_k)$ is called the kth principal system of the sample of the vector $\boldsymbol{X}$.

The coordinates of the projection of the ith realization $\boldsymbol{x}_i$ of the vector $\boldsymbol{X}$ on the kth principal component are equal to:

$$\hat{U}_{ik} = < \hat{\boldsymbol{u}}_k, \boldsymbol{x}_i > = \hat{\boldsymbol{u}}_k' \boldsymbol{x}_i,$$

for $i = 1, 2, ..., n$, $k = 1, 2, ..., p$. Finally, the coordinates of the projection of the i th realization $\boldsymbol{x}_i$ of the vector $\boldsymbol{X}$ on the plane of the first two principal components from the sample are equal to $(\hat{\boldsymbol{u}}_1' \boldsymbol{x}_i, \hat{\boldsymbol{u}}_2' \boldsymbol{x}_i)$, $i = 1, 2, ..., n$.

Functional Data version of the Principal Component Analyzis

- The employed functional data measurement approach – in the version of the *Multivariate Functional Principal Component Analysis (MFPCA)* - is an extension of the classic principal component analysis PCA from vector data to functional data (Górecki T., Krzyśko M., Wołyński W., 2019) with the procedure of representing data by function or curves (see Ramsay and Silverman, 2005) developed on the Besse's (1979) theoretical idea of multivariate data – where random variables take values in general Hilbert space - and its further important developments in different contexts. Of special interest here is an application to factorial methods - principal component analysis, canonical analysis - by Saporta (1981), and by Jacques and Preda (2014), who demonstrated usefulness of combining the MFPCA with Cluster Analysis.

  - The advantage of the FPCA over the classic PCA is to obtain a projection of analyzed units into one or two dimensional subspaces using information for the **whole period under study**, and to divide them into **homogenous groups** on the basis of the resulting rankings.

FPCA – *cont*

We assume that the analyzed objects characterized by variables are observed in many time points (years, months, days). Therefore, an appropriate model describing the examined objects will be *p*- dimensional random process

$$\mathbf{X}(t) = (X_1(t), \ldots, X_p(t))^\top \quad t \in I$$

Assume also that $\mathbf{X}(t) \in L_2^p(I)$ where $L_2(I)$ is a Hilbert space of integrable functions with a square on the interval *I*, and that the expected value of the process

$$\mathrm{E}(\mathbf{X}(t)) = \mathbf{0} \quad t \in I$$

From the above it follows that each component of the process can be represented in the following form:

$$X_k(t) = \sum_{b=0}^{\infty} \alpha_{kb}\varphi_b(t), \ t \in I,$$

where in the functions $\varphi_1, \varphi_2, \ldots$ form a base in space $L_2(I)$

FPCA – *cont.*

The above representation of the process requires knowledge of an infinite number of coefficients. We use an approximate representation that uses only a finite number of the first base functions. Assume that the k-th component of the process has the following representation:

$$X_k(t) = \sum_{b=0}^{B_k} \alpha_{kb}\varphi_b(t), \quad t \in I,$$

- where the number B_k determines the degree of smoothness of the function $X_k(t)$ (the smaller the value B_k the greater the degree of smoothing). Similarly to the classical case, we are looking for a randomvariable (the first functional component) U of the form:

$$U = \langle \boldsymbol{u}, \boldsymbol{X} \rangle = \int_I \boldsymbol{u}(t)^\top \boldsymbol{X}(t)dt$$

having the maximum variance for all $\boldsymbol{u}(t) \in L_2^p(I)$; (u, u)=1.

In general, the k-th functional main component fulfills the conditions:

$$\lambda_k = \sup_{\mathbf{u} \in L_2^p(I)} \mathrm{Var}(\langle \mathbf{u}, \mathbf{X} \rangle) = \mathrm{Var}(\langle \mathbf{u}_k, \mathbf{X} \rangle), \quad \langle \mathbf{u}_{\kappa_1}, \mathbf{u}_{\kappa_2} \rangle = \delta_{\kappa_1\kappa_2}, \quad \kappa_1, \kappa_2 = 1, 2, \ldots, k.$$

In the functional case, we have:

$$\langle \mathbf{u}_k, \mathbf{u}_k \rangle = \int_I u_{k1}^2(t) + u_{k2}^2(t) + \cdots + u_{kp}^2(t)dt = 1.$$

Thus, the quantity $\int_I |u_{kj}(t)|dt$ is a measure of the contribution of *j-th* component of the random process to the construction *k-th* functional principal component.

Since this process is only observed in a finite number of time moments, it is necessary to transform (smooth) discrete data into functional data (for details, see Ramsay and Silverman (2005); Gorecki, Krzysko, Wolynski (2019).

Functional version of the PCA – *cont.*

The above representation of the process requires knowledge of an infinite number of coefficients. In practice, we use an approximate representation that uses only a finite number of the first base functions.

So let's assume that the k-th component of the process has the following epresentation:

$$X_k(t) = \sum_{b=0}^{B_k} \alpha_{kb}\varphi_b(t), \; t \in I,$$

where the number B_k determines the degree of smoothness of the function $X_k(t)$ (the smaller the value B_k, the greater the degree of smoothing).

Similarly to the classical case, we are looking for a random variable (the first functional component) U of the form:

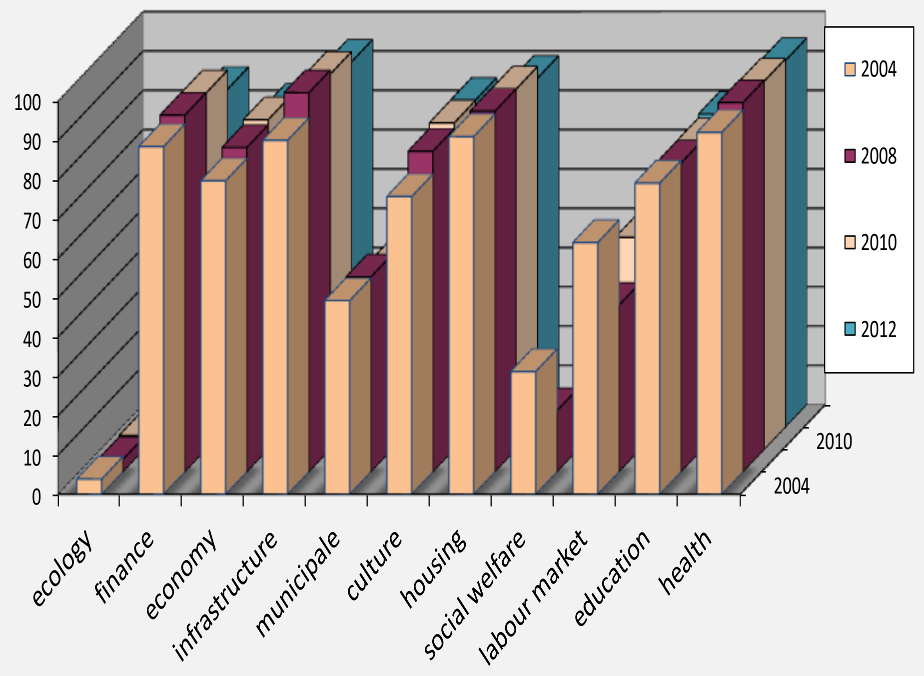
$$U = \langle \boldsymbol{u}, \boldsymbol{X} \rangle = \int_I \boldsymbol{u}(t)^\top \boldsymbol{X}(t)dt$$

having the maximum variance for all $\boldsymbol{u}(t) \in L_2^p(I)$ such that (u,u) =1.

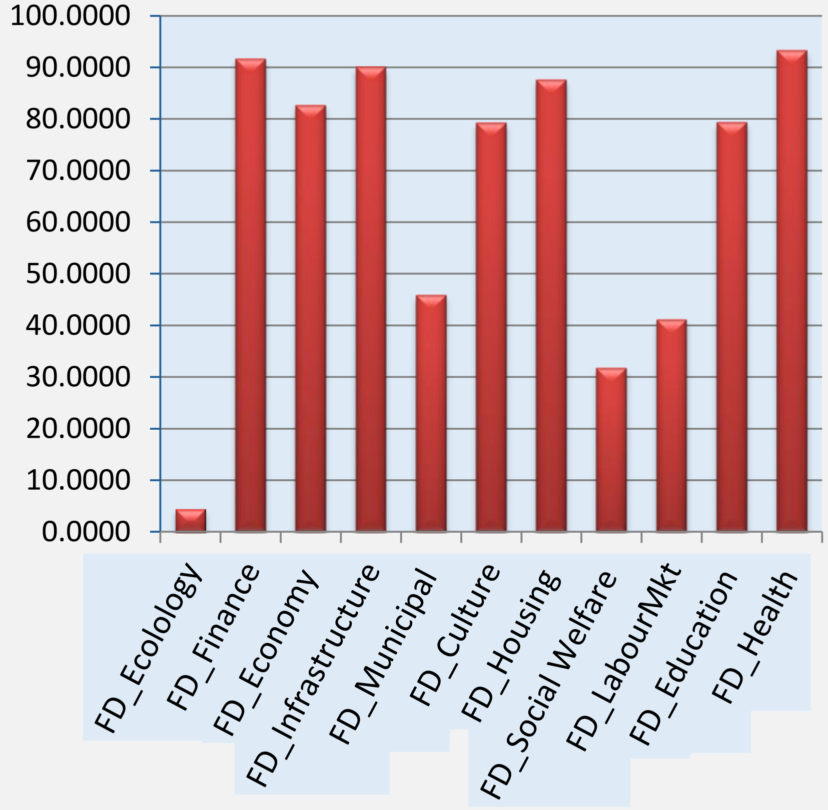
Comparison of local deprivation measures according to classic PCA and FPCA:

(FD_deprivation = $_{df}$ *de-deprivation/inverse deprivation = development*)

Multidimensional Index of Local Deprivation /MILD
– by domains (2004-08-10-14)



FD_Index of Local Deprivation /FD_MILD
by domains (2004-2016)



Influence of ***local risk*** associated with particular domains of local deprivation on selected measures of *satisfaction with …*

[After constructing the classifications of both local communities (gminas) and their residents (for a given period of time -2004 – 2014, 2016, and 2009-2015, respectively), the spatial perspective is involved (Okrasa, Krzyśko, Wołynski, 2020).]

- The *synthetic measure of satisfaction* – as an indicator of overll subjective well-being attributed to commune as a place of residents ('compositional' variable: percentage of 'satisfied' or 'very satisfied' on each scale) - is composed of the following separate scales:
  - (i) satisfaction with *living conditions,*
  - (ii) satisfaction with *living environment ,*
  - (iii)satisfaction with *social and family relations, (SMS)*

- **Local Risk** is defined as a product of a FD-scale of local deprivation in a domain ILD_d (F-Index of Local Deprivation in the domain *d*) and the respective fraction of the commune population (P_k) defined as the ratio of the ILD_d to the total size of local deprivation (FMILD):

$$\text{RiskFD\_(domain)} = \text{FD-deprivation (d-domain)} \times (P_k * (ILD_d / MILD)).$$

OLS: Local Risk associated with domain of local (inverse)deprivation as a separate predictor	Functional Data Scales of Subjective Well-Being: *satisfaction with*		
	All scales -synthetic measure of satisfaction (SMS) #	*Living conditions*	*Living environment*
FD_Risk assoc. w/ all domains	-.176 ** {-3.366)	-.100 * (-1.893)	-.107 **(-2.022)
FD_Risk assoc. w/ *ecology*	-.188 ** (-3.583)	-.098 * (-1.835)	-202 ** (-3.852)
FD_Risk assoc. w/ finance	-.194 ** (-3.7070)	-.119 ** (-2.249)	-.096 * -1.810)
FD_Risk assoc. w/ economy	-.201 ** (-3 .830)	-.123 ** (-2.311)	-.096 * (-1.812)
FD_Risk ass. w/ infrastructure	-.204 ** (-3.895)	-.123 ** (-2.327)	-.094 * (-1.760)
FD_Risk assoc. w/ culture	-.190 ** (-3.621)	-.112 ** (-2.105)	-.089 * (-1.679)
FD_Risk assoc. w/housing	-207 ** (-3.959)	-.119 ** (-2.236)	-.107 ** (-2.015)
FD_Risk ass. w/ soc.welfare policy	-.130 ** (-2.452)	-	-
FD_Risk assoc. w/ labor market	-.135 ** (-2.554)	-	-102 ** (-1.912)
FD_Risk assoc. w/ education	-.105 ** (-1.975)	-	-.125 ** {-2.349)
FD_Risk assoc. w/ health	-203 ** (-3.888)	-.123 ** (-2.310)	-.089 * (-1.677)

Significant negative effect of the risk associated with every domain of local **inverse-deprivation(development)**, esp. in the domain of economy, infrastructure and health; and in ecology for *living environment.*

Individual (subjective/quasi-objective) well-being: *Time Use Survey* dataset-based measures

- Social indicators approach – attempts to exploit TUS data (Juster; and others. e.g. Andrews 80s.); in economics (macro-indicators, Becker 1965; Nordhaus, 2009; micro-level: Kahneman and Krueger, 2006); (also used in poverty research – eg., gender effect).
    - Survey research (day reconstruction method/DRM –Statistics Poland: TUS_2013 ; N=23 000)
- Econometric research and econometric/psychometric combined approaches – Krueger and Khaneman et al.. (2008) – indicator of emotion / negative /positive affects associated with a performed activity / ***time of unpleasant state,*** **U**-index :

$$U_i = \Sigma_j\, I_{ij} h_{ij} / \Sigma_j h_{ij} \qquad (TUS_{2013}\text{: I = -1; 0; +1})$$

and $\mathbf{U = \Sigma_i(\Sigma_j\, I_{ij} h_{ij} / \Sigma_j h_{ij}) / N}$ for N-persons / group in population ;

For U-binary (-1 & 0 vs. +1), odds of U [chance of other than 'pleasant' or non-positive state vs. 'pleasant']:

Odds (prevalence) in (U) :: $\mathbf{U_i / (1- U_i)}$ → Odds U by the community level FD-measures of deprivation/development and by its selected characteristics

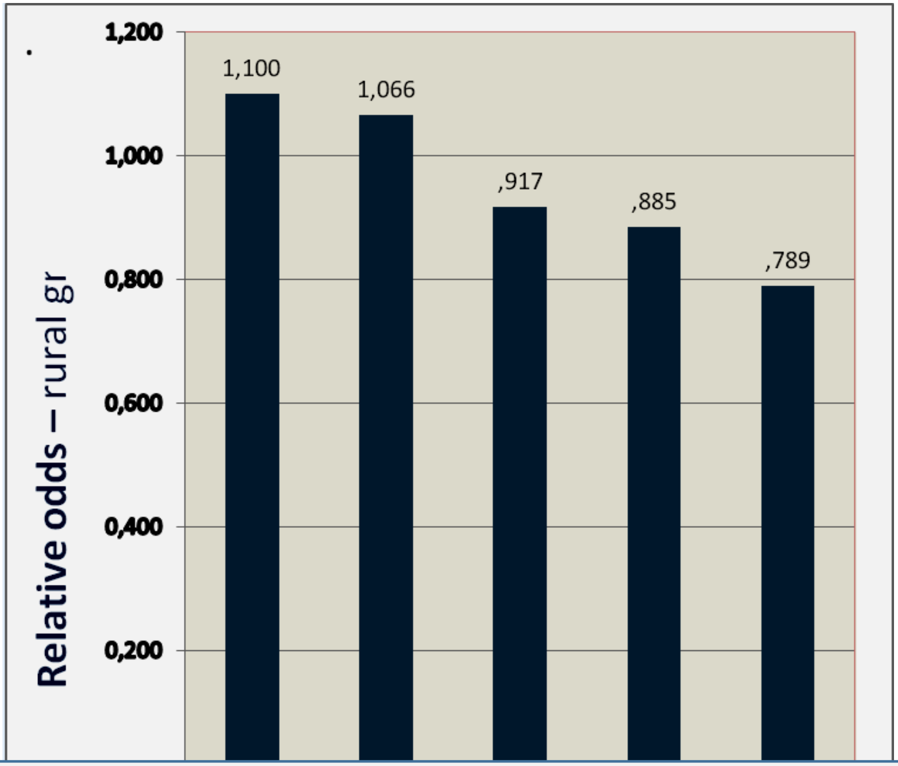
Effects of ***local deprivation*** and of ***risk** associated* with local deprivation (in Functional Data version) and of the local **community characterisitcs** for ***individual well-being***

(odds of U-unpleasant - average for a commune's residents in the TUS sample; min. 10 pers. Per)

Model	Unstandardized Coefficients		Stand. Coeff.	t	Sig.
	B	Std. Error	Beta		
• (Constant)	0,494	0,209		2,364	0,018
• FD_Local (inverse) Deprivation (2004-16)	0,000	0,000	-0,092	-2,204	0,028
• Risk assoc. w/labor market	-0,073	0,021	-0,211	-3,542	0,000
• Risk assoc. w/depr. Local economy	0,080	0,027	0,214	2,987	0,003
• Temporarily absent (from home/per 1000)	0,004	0,002	0,071	1,979	0,048
• Proportion of 'employed' to 'not-employed' in community	-0,054	0,010	-0,175	-5,631	0,000
• Number of NGOs per 1000 pers	-0,017	0,011	-0,049	-1,534	0,125
• Local authority active in revitalization	0,069	0,026	0,085	2,715	0,007
F (7, 1012) = 9,7 842; p<0.000					

Odds of experiencing 'non-positive' feeling associated with activities performed: U - index, depending on (a) the size of the living place, and (b) the level of household income

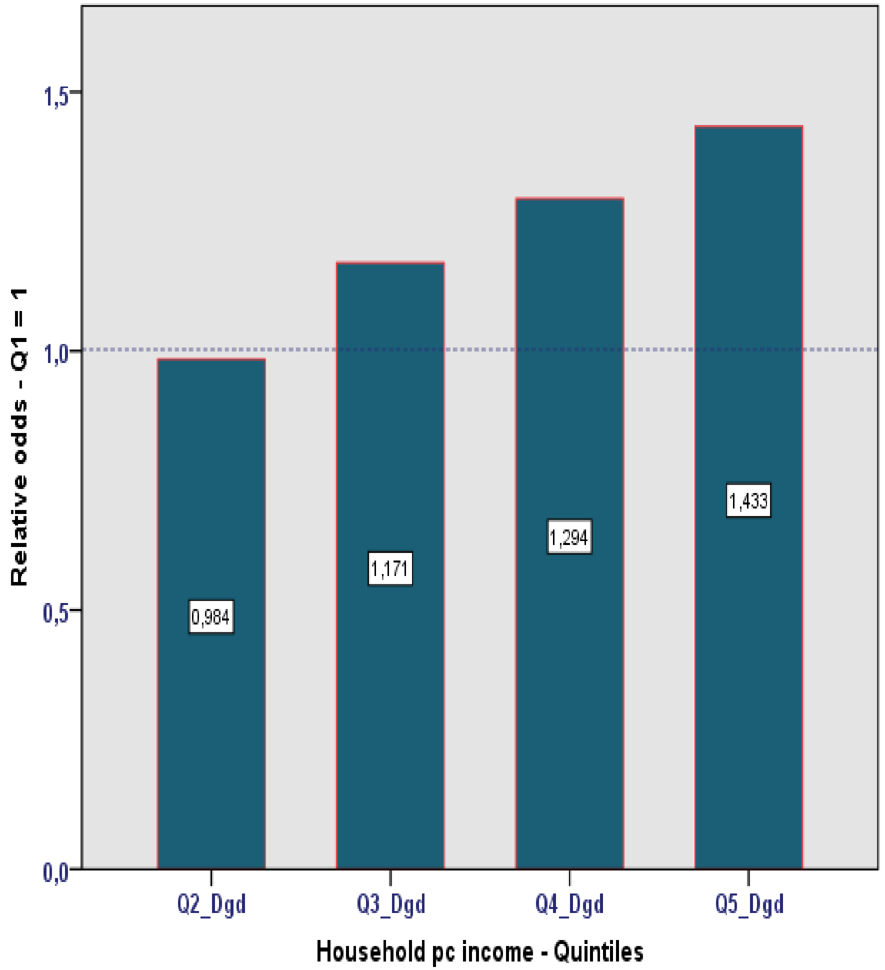
(a) the size of the living place



Size of thel iving place (commune)

>=500K **>=200-499K** **>=100-199K** **>=20-99K** **<20K**

(b) *The level of household income pc*



Community cohesion in spatio-temporal evaluation perspective :

subsidies to gminas as a deprivation-reducing (development) program.

Marginal Benefit Incidence Analysis / MBIA

using repeated spatial observations

I. Assessing incidence of subsidies at two or more years provides better insight into local policy about allocation of public resources to communes / *gminas*.

If average incidence (E_{it} / E_t) and ($\mathrm{E}_{i\ t+1} / \mathrm{E}_{t+1}$) are defined as the average share of total subsidies accrued to quintile *i* at year *t and t+1,* respectively, then the change in quintile-specific share of subsidies is given by:

$$\boxed{(\mathrm{E}_{i\ t+1} / \mathrm{E}_{t+1}) - (E_{it} / E_t)}$$

(Marginal) Benefit Incidence Analysis

using repeated spatial observations data – *contin.*

- ***Average*** odds of participation = ratio of group specific average participation rate to overall average

- ***Marginal*** odds of participation (MOP) = increment to group-specific participation rate with a change in overall participation

  → *MOP shows incidence of a change in subsidies*

  - to get it estimated

    - regression of income group specific participation rate across regions (voivodships) on average rate for region (voivodship) to get income group specific MOP.

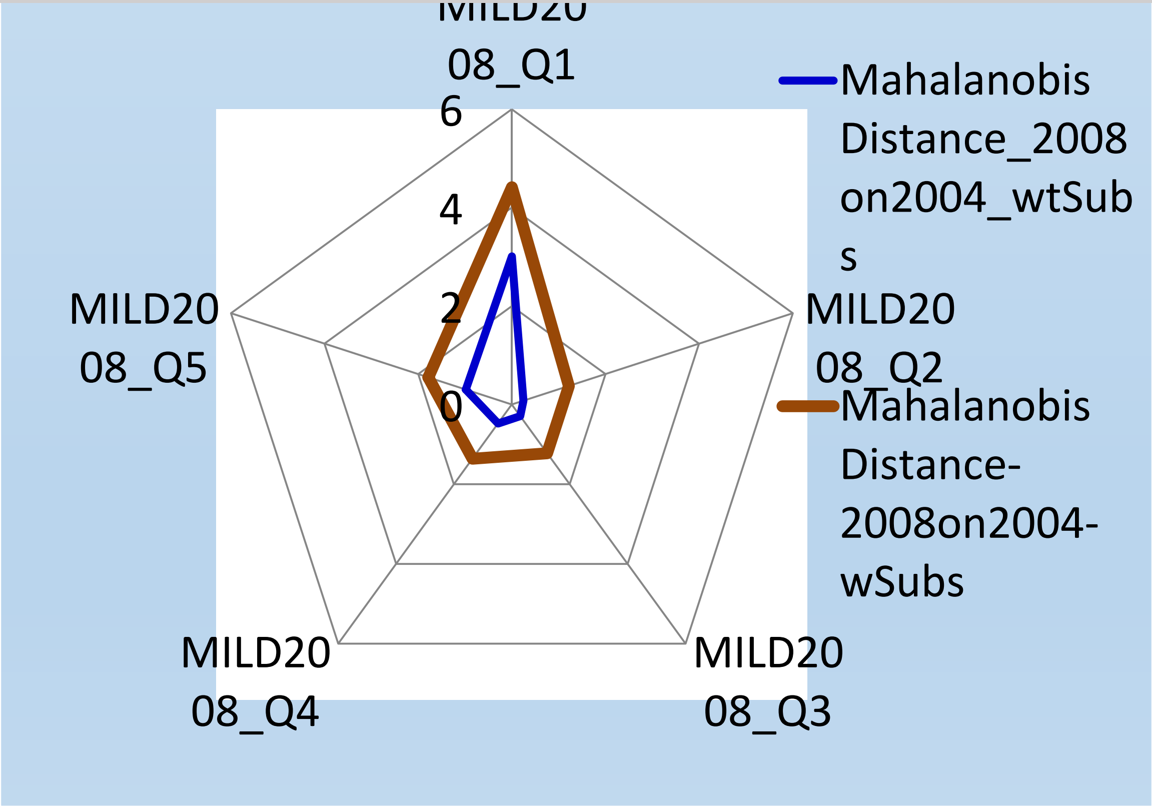
Community well-being sources: *Does public suport (subsidy) matter*

for the local development (i.e. reduction in deprivation)?

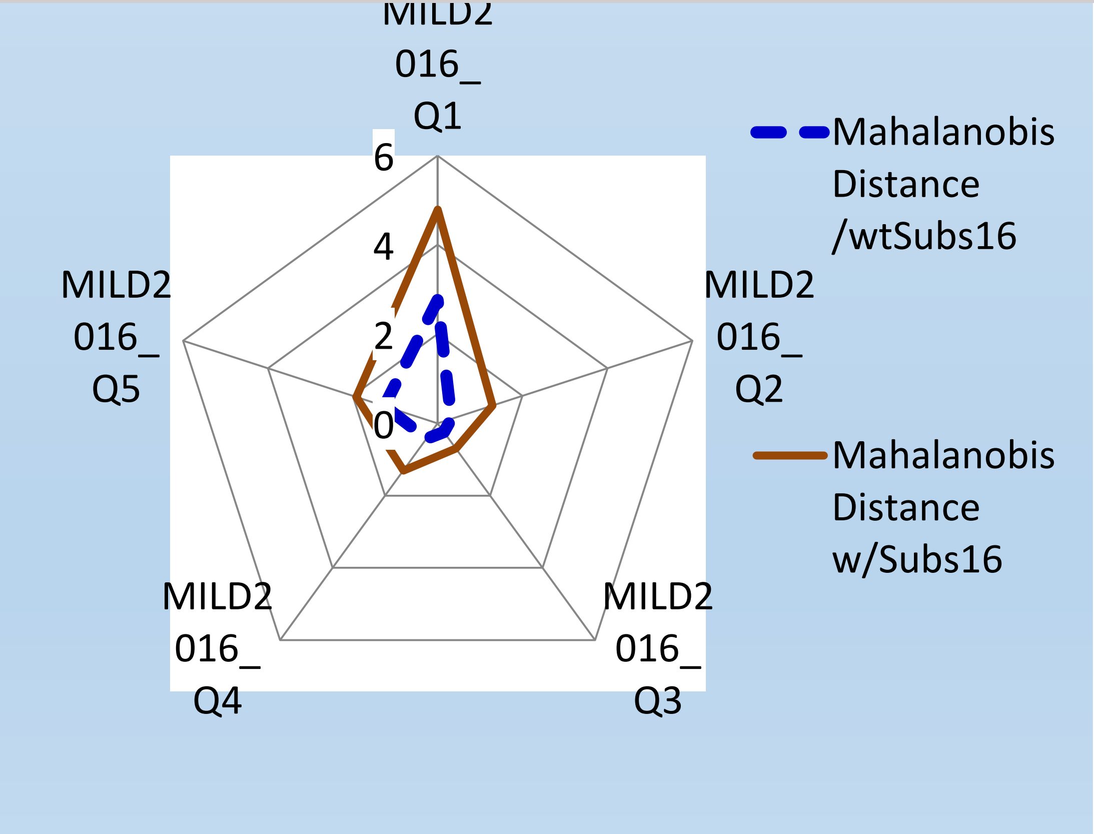
Example: **MILD in time *t+1*** predicted (regressed) on **MILD in time *t*,** *with* and *without* taking into account subsidies to gmina

Figures A, B – differences in terms of the *Mahalanobis Distance.*

A. MD for Predicted MILD_2008 on MILD_2004 - with and without subsidies

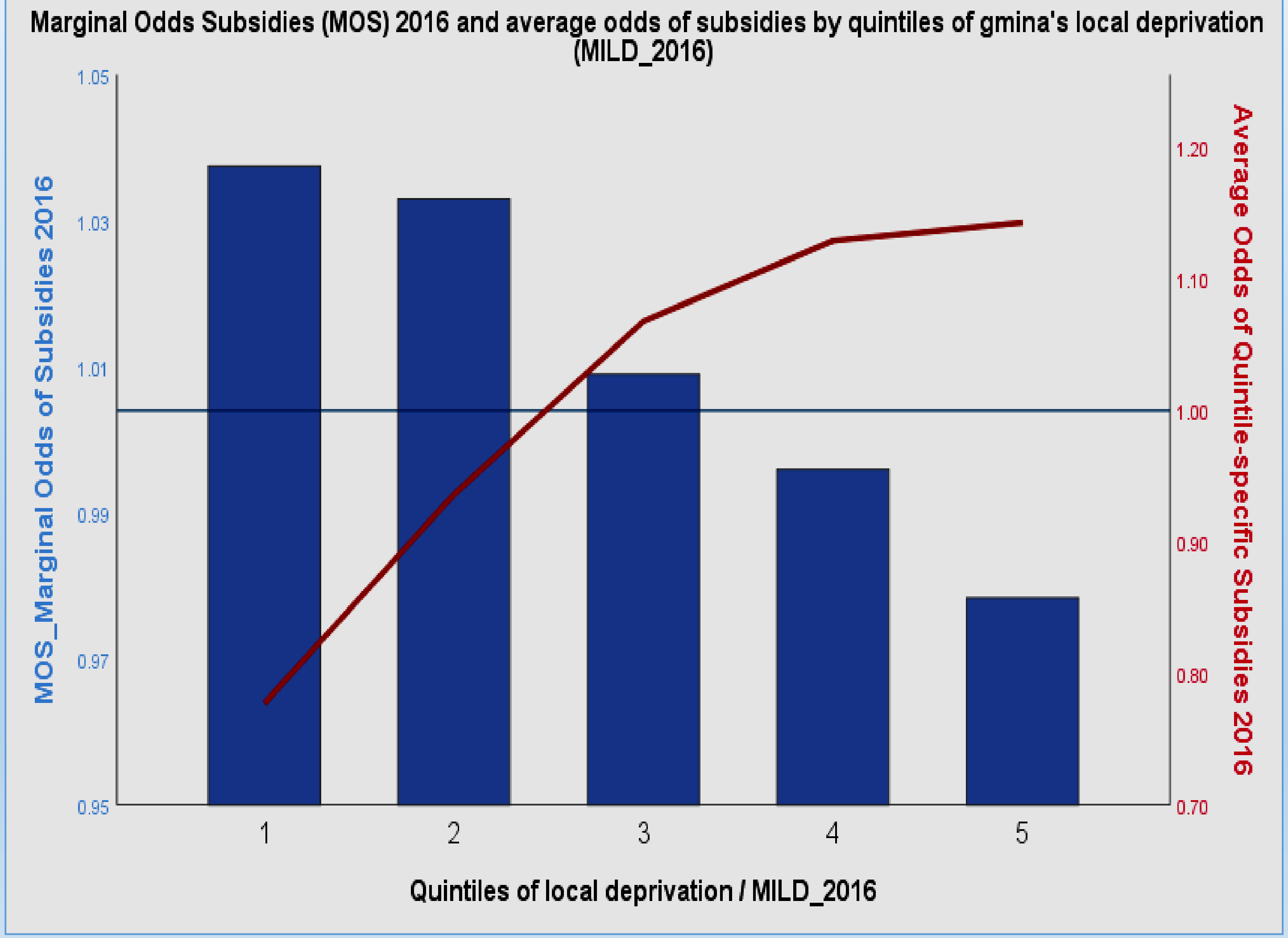


B. MD for Predicted MILD_2016 on MILD_2008 - with and without subsidies



MBIA/Marginal Benefit Incidence Analysis

Marginal Odds (participation in) Subsides /MOS 2016
and odds of quintile-specific values to NUTS5/*gminas*, across regions (voivodships)



Quintiles of MILD	**Average odds of Subsidies** 2016_Q-mean, by region	**MOS_Sub16**
1. Least deprived gminas	0,78	**1,04**
2	0,94	**1,03**
3	1,07	**1,01**
4	1,13	**1,00**
5. Most deprived gminas	1,14	**0,98**

Results – *comments on allocation of subsidies*

The model fits data well providing a robust base for making predictions of the level of subsidies being accrued to communes from the knowledge of their characteristics (predictors) included.

1) The value of ***the local deprivation (MILD) significantly influences the decision about the level of subsidies***: more deprived communes obtain bigger share of public resources (as above). It means that the applied mechanism of geographic targeting may contribute to the objectives of territorial cohesion policy.
2) Negative slope of the β_2 coefficient – for the ***relation between income inequality among communes (within county) and the level of deprivation (MILD) -*** accords with the expectations suggested by *Williamson's hypothesis* (1965) [that relation between inequality and the level of local development is shaped as an inverted U, like Kuznets' hypothesis for inequality of income distribution and GDP: gminas in more differentiatied areas (counties/powiats) are on general less deprived, and *vice-versa* – gminas in more deprived powiats tend to be less differentiated amongst themselves.

Models type I:

Cross-level operating factors of individual and community well-being:

Macro - micro influence

Multilevel modeling

Community Cohesion and the role for ***social capital:*** *Compensating variation *)*

An Experience-based Method for Valuing Social Capital

As our approach assumes that self-reported life satisfaction is a good proxy for value (utility) and estimates a utility function *U,* that depends positively on **household income, y, and on social capital SC,** it is reasonable to use a standard *compensating variation* measure of value (CoV).

[It is interpreted as the amount of money required to compensate a person for – typically, a price change - that gives rise to a loss in utility.]

Here the **compensating variation** for social capital**, CoV,** can be obtained by identifying the utility gain derived from a unit increase in social capital (eg., Anand, 2018)

Formally, a **life satisfaction equation** can be written as:

$$U^0(y^0, SC^0)=U^1(y^0+CoV,SC^1)$$

*) Anand (2018)

The role for Social Capital / Compensating Variation – cont.

Using ***a linear life satisfaction equation***, the expected utility given any particular value of social capital can be written as:

$$E(U_i|SC_i, y_i, X_i) = \beta_0 + \beta_y y_i + \beta_{sc} SC_i + \gamma' X_i + \varepsilon_i$$

- where *X* represents all additional covariates.

CoV then becomes a solution to the equation,

$$E(U_i|SC_{i0}, y_{i0}, X_i) = E(U_i|SC_{i0} + y_{i0} + CoV, X_i)$$

- which in turn implies that

$$\mathbf{CoV = \beta_{SC} / \beta_y.}$$

How *Community Cohesion (CC)* interferes with community residents' wellbeing? The *Well-Being Equation* extended by CC-related individual-level variables

`	Unstand. Coefficients		Stnd. Coeff.		
	B	**Std. Error**	**Beta**	t	**Sig.**
(Constant)	0.029	0.027		1.068	0.285
• Job-time (main and additional)	0.004	0.000	0.285	24.630	0.000
• Income of H'hold *pc* - monthly	-1.841E-05	0.000	-0.087	-6.987	0.000
• MILD_2014 Local Deprivation	0.000	0.000	0.118	6.630	0.000
• Subsidies Real < Simulated /fair	-0.011	0.002	-0.070	-6.887	0.000
• Risk assoc. w/depr. Soc.Welfare	-0.036	0.002	-0.649	-15.626	0.000
• Risk assoc. w/depr. Lab. Market	0.050	0.003	0.809	18.454	0.000
• Ratio 'in-work' to 'not-in-work'	-0.010	0.001	-0.080	-6.900	0.000
• Rural	-0.007	0.003	-0.030	-1.978	0.048
• U-R mixed	-0.014	0.002	-0.074	-5.547	0.000
• **Trust in local authority**	**-0.002**	**0.001**	**-0.032**	**-3.468**	**0.001**
• Satisfaction from the place	-0.002	0.001	-0.017	-1.898	0.058

Adjusted R Square = 0.18;
F (11, 10 095) 198.387; p< .000

CoV = -0.032/ -0.087 = 0.367 (37%)

Assessing cross-level interaction between personal and community well-being – a basic multilevel model (e.g., Subramanian. 2010)

- $\boldsymbol{y_{ij}}$; well-being of *i* individual in *j*th commune/gmin ;
- $\boldsymbol{x_{1ij}}$ predictor of indywidual (level-1) – such as: income, age, education, or satisfaction (e.g., with life in a community, family life , etc.
- predictor of level-2 / (macro-level): *Multideminsonal Index of Local Deprivation* for *j*th commune (gmina) /$MILD_j$

Model for level-1:

$$y_{ij} = \beta_{0j} + \beta_l x_{1ij} + e_{0ij}$$

where: β_{0j} – refers to x_{0ij} average score on a well-being scale in j-th commune/gmina (eg., . 'less affluent' or 'low-income', < Me, x_{0ij} =1);

β_l – average differentiation of individual well-being associated with individual material status , (x_{1ij}), across all communes; e_{0ij} – residual term for the level-1.

Treating β_{0j} as random variable: ($\beta_{0j} - \beta_0$) + u_{0j} , where u_{0j} is locally-specific associated with average value of β_0) for a specified group (eg. less satisfied with a community) and grouping them into fixed and random components ($e_{0ij} + u_{0j}$) we obtain *variance component model* or *random-intercept* model:

$$y_{ij} = \beta_0 + \beta_l x_{1ij} + (e_{0ij} + u_{0j})$$

Modeling *fixed-effect* we include a level-2 predictor – MILD -(index of local deprivation) along with individual characteristics, including *interaction* term between the two levels *:*

$$\beta_{0j} = \beta_0 + \alpha_1 MILD_{1j} + u_{0j} \quad \text{and} \quad \beta_{1j} = \beta_1 + \alpha_2 MILD_{1j} + u_{1j}$$

Accordingly, a two-level model can be specified as below:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \alpha_1 w_{1j} + \alpha_2 w_{1j} x_{1ij} + (u_{0j} + u_{1j} x_{1ij} + e_{1ij} x_{1ij} + e_{2ij} x_{2ij})$$

- where w_{1j} is a 2-level predictor. i.e. the index of local deprivation. $MILD_{1j}$.

The following model was calculated using data from *Time Use Survey* 2013 (22 695):

$$IWB(U\text{-}index)_{ij} = \beta_{00} + \beta_{10} education_{ij} + \beta_{20} income_{ij} + \alpha_1 MILD_j$$
$$+ \alpha_{11} education_{ij} * MILD_j + \alpha_{21} income_{ij} * MILD_j$$
$$+ u_{1j} education_{ij} + u_{2j} income + u_{0j} + e_{ij}$$

[It is assumed that] Such a specification of cross-level (between individual and community/gmina measures of well-being) with *interaction* effect should ensure robust estimation (e.g.. Subramanian. op. cit.. p. 521; Hox et al.. 2018).

→ Preliminary results

Multilevel regression of personal well-being – *U-index* (all activities) – on individual and commune charactersitics with cross-level interaction term; comparison of *Functional Data*-based and *classic PCA* approaches

Model with FDPCA-measures (MILDevelopment)	Std Beta	t	Model with classic PCA-measures (MILDeprivation)	Std Beta	t
Constant		13,258	Constant		5,096
Income	0,056**	6,901	Income	0,027**	4,050
Education (years of schooling)	0,075**	4,728	Education (years of schooling)	-0,045	-0,610
FD_Community Development 2004-2014	0,082*	2,304	Community Deprivation 2004-2014	-0,062*	-2,133
FD_Education *Community Development	-0,111**	-2,737	Education*Community Deprivation	0,123*	1,668
FD_Comm. Dvpt * Income	0,152**	17,778	Comm. Depriv.* Income	0,091**	13,547

$F(5,15086) = 100\,418$ ($p<.001$)

**) significant at $p < 0.01$; *) $p < 0.05$.

$F(5,22690) = 87\,196$ ($p <.001$)

Strong similarity of results obtained with FDPCA and PCA, respectively – with a more clear pattern of dependences in the first case – it does confirm the (expected) advantage of the former,

Models type II:

Structural modelling approach - *causal* mediating mechanisms: local deprivation as a factor modifying effect of an individual commune's attribute on the residences' well-being according to U-index

- Hhld Income - indpendent var. / 'treatment'
- Local deprivation / MILD – *mediating* factor

Hypothesis: The level of deprivation of a commune (gmina) affects the influence of the residents‘ subjective well-being by their material status (income)

[structural modeling (e.g. G. Hong. 2015)]:

Y - U-index (individual well-being)

Z – source of influence: HH income (average in a commune/gmina)

M - *mediator*: level of local deprivation /MILD (or ILD $_{1,\ldots,11}$)

$$M = \gamma_0 + aZ + \varepsilon_M$$

$$Y = \beta_0 + bM + cZ + \varepsilon_M$$

Substituting for M → *reduced-form* model:

$$Y = (\ldots) = \beta'_0 + c'Z + \varepsilon'_Y$$

Estimation of diffrences between coefficients of ifluence c’ – c (with local deprivation/MILD as a mediator) allows to assess indirect effect (of MILD) in estimating influence of Hhld income on individual well-being (U)

Structural (*causal*-type) modelling:

quality of living environment(*ILD*-selected domains) as a moderating factor in assessing influence of respondents' income on subjective well-being

Model / predictors	Standardized Coefficients		Difference
	Beta	t-statistics	|c'- c|
Dependent Var: U-index for all activities			
M I: ILD_economy	-.054	-1.565	0.286
Monthly income/ Mi (c’)	.072 *	2.070	
ILD_economy on Mi (c)	-.358 **	-11.807	
M II: ILD_social assistance	-.091 **	-2.824	0.007
Monthly income /Mi (c’)	-.111 **	-3.439	
ILD_soc asst. on Mi (c)	-.104 **	-3.214	
M III: ILD_labor market	-.089 **	-2.725	0.090
Monthly income /Mi (c’)	-.061 *	-1.850	
ILD_labor market on Mi (c)	-.154 **	-4.802	
M IV: ILD_health	.054	1.638	0.108
Monthly income /Mi (c’)	-.070 *	-2.137	
ILD_health on Mi (c)	-.178 **	-5.583	

The impact of respondents’ income on (subjective) well-being is modified by the level of local deprivation in selected domains (ILD); health and economy are relatively stronger than labor market and social assistance (but the latter are more important in the spatial context – see below).

Spatial aspects of cross-level relationship
- acounting for spatial heterogeneity

- Spatial autocorrelation and the tendency to geographic co-occurence of selected measures of subjective well-being along with FD_deprivation/development or 'classical' deprivation in selected domains.
- Spatial dependence of selected measures of subjective well-being - *spatial error model* of spatial regression on FD_Risk in the Labor Market and in the Local Economy, given a set of auxiliary covariates.

Two-step spatial analysis:

(1) Checking a tendency to clustering among 'spatial units' (communes/gminas) with respect to selected measures – subjective and objective – using Moran' I (global):

$$I = \frac{n}{W} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

- *where:* x_i. x_j - values of a measure at each location; ***W*** is the spatial weights matrix.

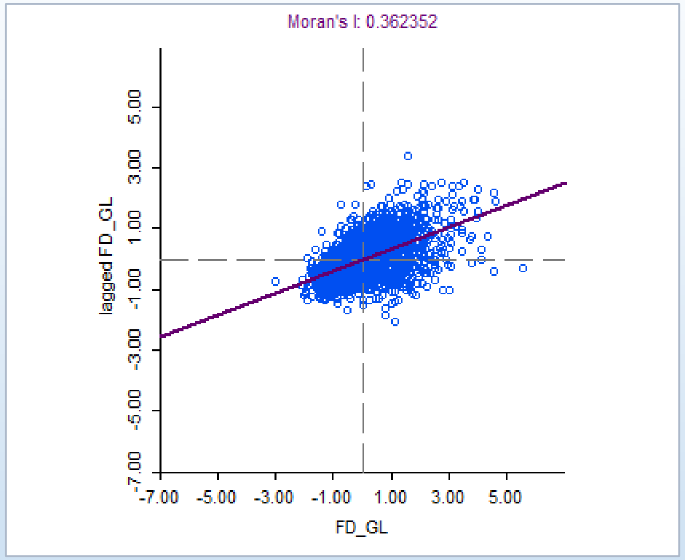
(2) Estimation of the spatial regression model parameters: (notation for individual/commune observation *i*):

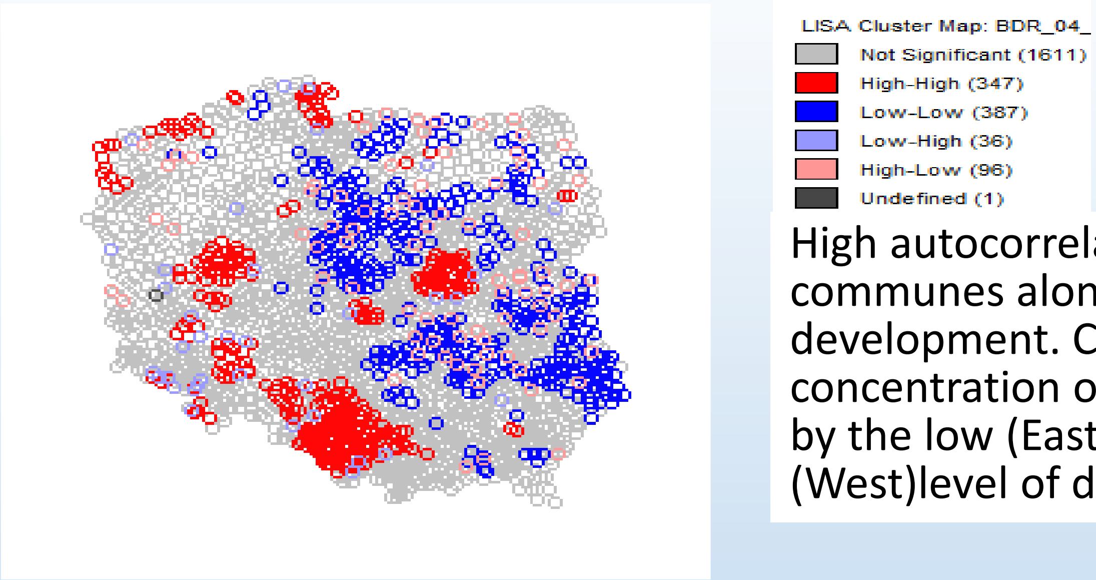
$$\boldsymbol{y_i = \rho \sum_{j=1}^{n} W_{ij}\, y_j + \sum_{r=1}^{k} X_{ir}\, \beta_r + \varepsilon_i}$$

- where: y_i – the dependent variable for observation *i*; X_{ir} *k* – explanatory variables . r = 1. …. k with associated coefficient β_r ; ***W*** matrix; *ρ* is parameter of the strength of the average association between the dependent variable for region /observations and the average of them for their neighbours; ε_i is the disturbance term – it might be assumed that ε_i is meant as either the *spatially lagged* term or *spatial error* formulation ((eg.. LeSage and Pace. 2010).

(LISA:) Scatter plots and cluster maps of local deprivation according to: (a) $FD\_MILD_{2004\text{-}2016}$ (M's I: 0.36); and (b) $MILD_{2016}$ (M's i: 0.39) - comparison

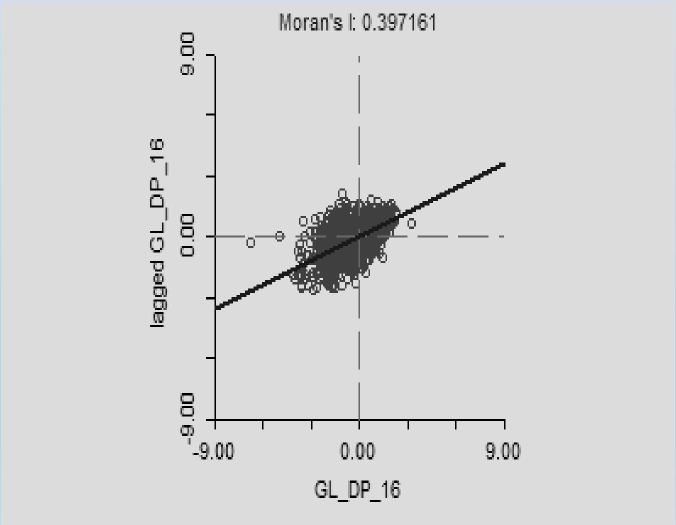
(a)

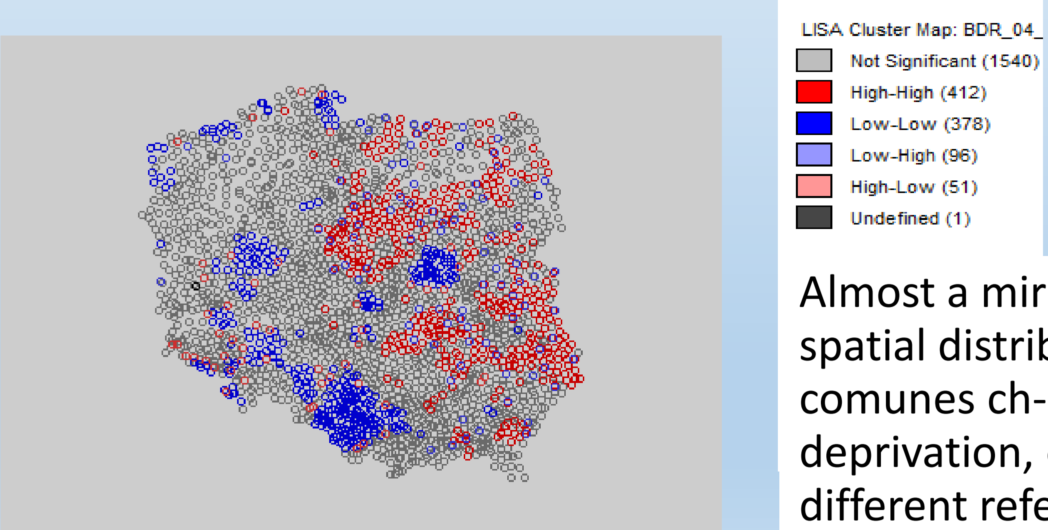




High autocorrelation of communes along the level of development. Clear pattern of concentration of clusters ch-d by the low (East) vs. high (West)level of development .

(b)

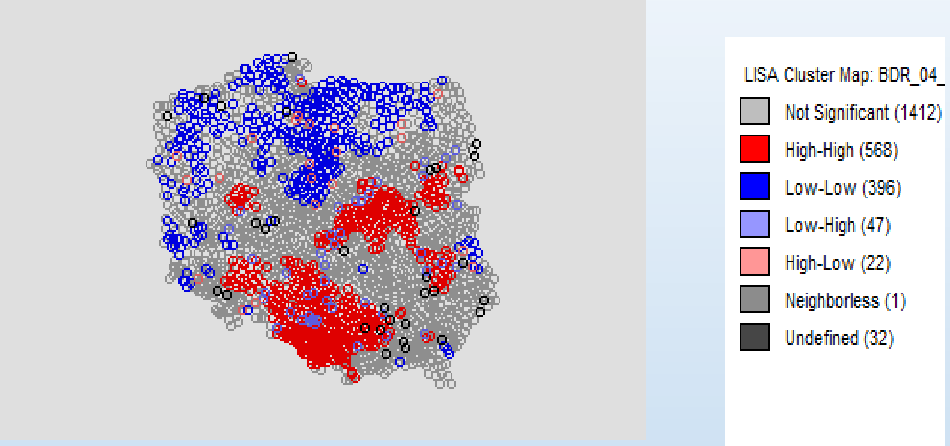




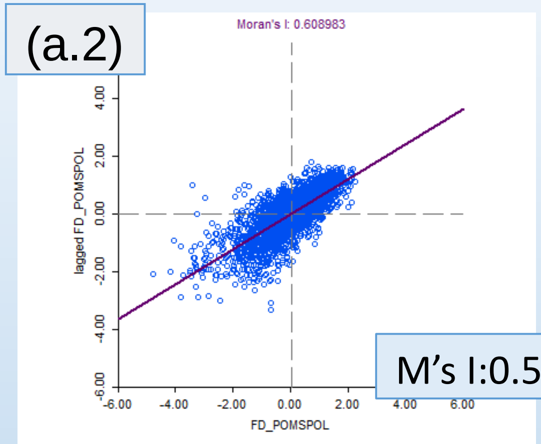
Almost a mirror pattern of the spatial distribution of comunes ch-d by the level of deprivation, despite the different reference period

Cluster maps and scatter plots of deprivation / ‘development’ in the domains of (a) *local social welfare* by FD-measure (2004-16) and FA-classic and (b) *local labour market by* FD-measure and FA.

(a.1)

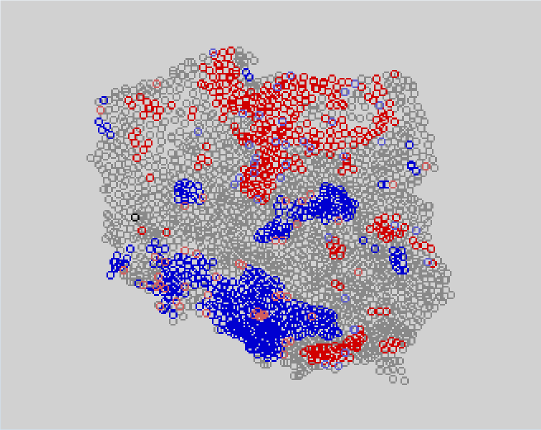


(a.2)

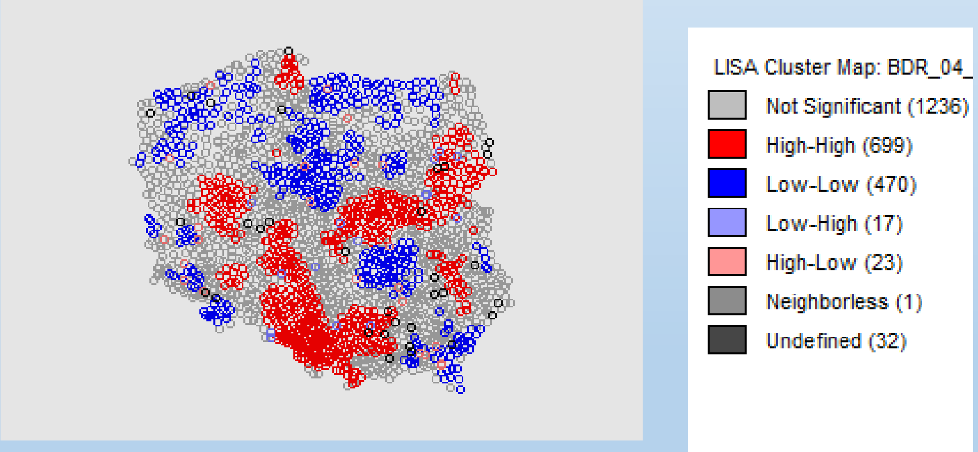


M’s I:0.56

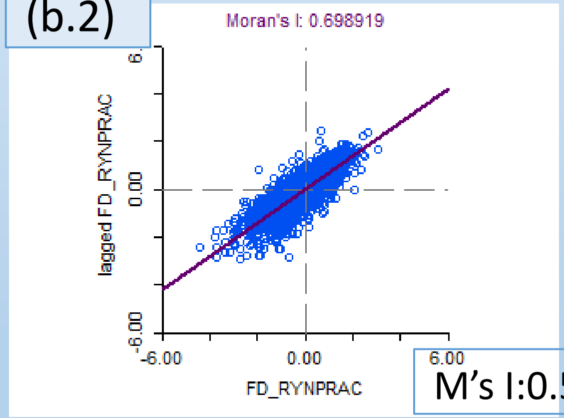
(a.3)



(b.1)

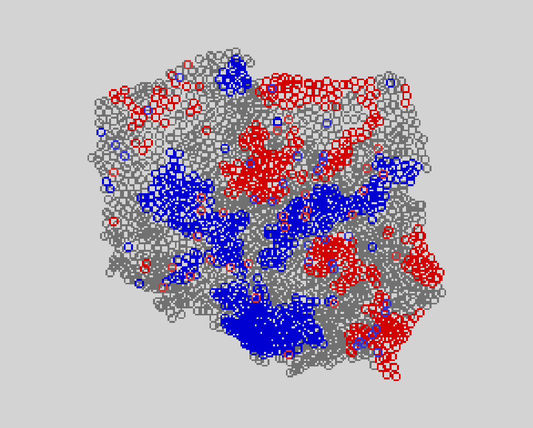


(b.2)



M’s I:0.56

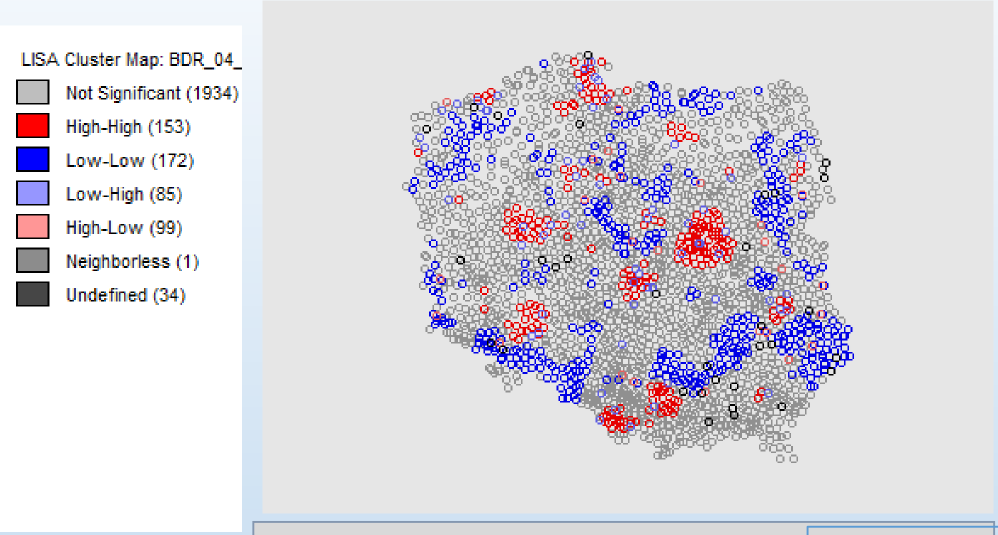
(b.3)

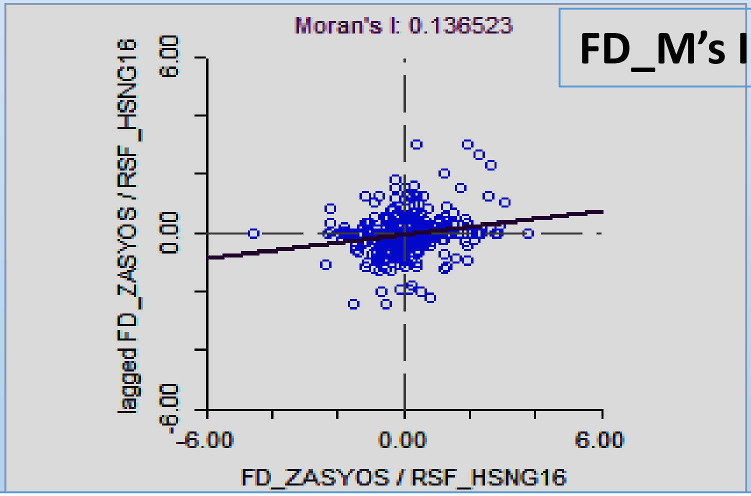


Strong autocorrelation and clear pattern of spatial clusters in each of the two domains – local social welfare and labour market – provide case for interpretation of the above relationships between risk associated with FDPCA-measure and ‘classic’ PCA measure, (a.1&a.3, and b1&b.3, respectively): the patterns are similar (but inverted values suggests different interpetation - ‘development’ (‘1’) vs. deprivation (‘3’).

Risk associated with: (a) FD_deprivation in *housing* vs. subjective W-B/*satisfaction with personal situation*; (b) deprivation in housing vs. *satisfaction with living conditions*.

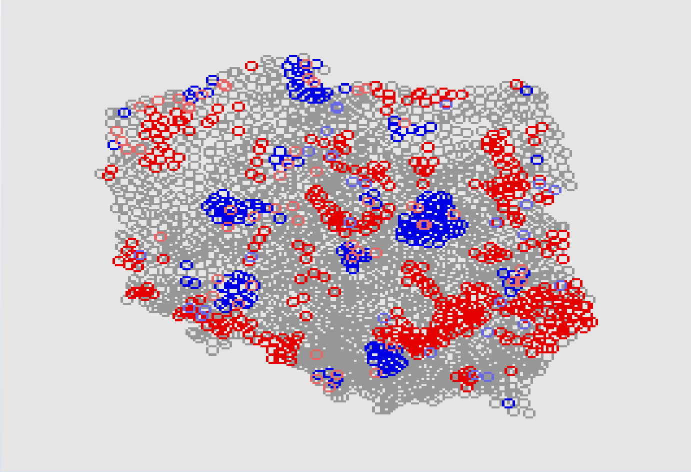
(a)

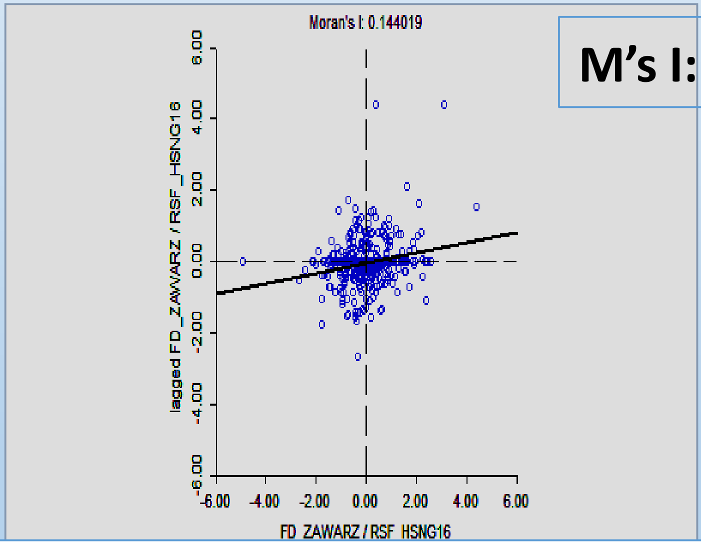




FD_M's I: 0.36

(b)





M's I: 0.34

Selected measures of subjectwive well-being –*safisfaction with personal situation* and *with living conditions* – are consistently positively related to the risk associated with deprivation in the domain of housing. The negative relations in which it remains with the FD_deprivation ('development' – see the cluster maps means (i) that relatively 'higher' risk can be iseen as favorable for geographical co-occurence of the *satisfaction with personal situation* snd the development (de-deprivation) in the domain of housing; (ii) as regards t*satisfaction with living conditions,* its lower level tends to geographically co-exist with generally the better housing situation (low deprivation) - as it is clearly seen in the case of fivemetropolitan areas

SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION (FD-measures)

Dependent -- Subjective well-being

All scales – SMS/*Synthetic Measure of Satisfaction (N 352)*

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	6.58928	4.69413	1.40373	0.16040
RiskFD_LabMkt	1.05379	0.470991	**2.2374**	**0.02526**
RiskFD_Economy	-1.4603	0.542753	-**2.69055**	**0.00713**
Subsidies FD_2016pc	0.000735	0.001080	0.68092	0.49592
NGOs per 1000_2016	-0.46272	0.2308	-**2.00488**	**0.04498**
Comm. w/revitalization	0.18080	0.381625	0.473771	0.63566
Migration_balance	0.04184	0.041461	1.00924	0.31286
LAMBDA	**0.16678**	0.0560501	2.97569	0.00292

DIAGNOSTICS FOR HETEROSKEDASTICITY

TEST	DF	VALUE	PROB
Breusch-Pagan test	6	36.8021	0.00000

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : BDR_04_16_Juneo5_2019

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	8.4296	0.00369

SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION (FD-measures)

Dependent: *Satisfaction with personal situation* (N 352)

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	2.90449	3.33088	0.87198	0.38321
RiskFD_LabMkt	0.630598	0.329713	1.91256	0.05580
RiskFD_Economy	-0.747787	0.381996	-1.95758	0.05028
Subsidies FD_2016pc	1.7133e-05	0.0007664	0.022345	0.98217
NGOs per 1000_2016	-0.262531	0.16303	-1.61032	0.10733
Comm. w/revitalization	0.009240	0.26974	0.03425	0.97267
Migration_balance	-0.008147	0.029267	-0.27837	0.78072
LAMBDA	**0.132998**	0.056856	2.3392	0.01933

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	6	30.3470	0.00003

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	4.9045	0.02679

CONCLUSIONS

Following the conceptualization of the triadic interdependence of data, measurement, and model, some observations seem worth mentioning in light of the presented empirical results:

► Bottom-up, data-driven approach to constructing Analytical Data Base encompassing individual and group/commune variables, seems to provide an alternative to the lacking appropriate (nested) data structure in analyzing cross-level relationship between the respective (development and well-being) measures, within a multidimensional framework.

► Functional Data approach to multidimensional measurement of community well-being (i.e., switching from PCA to FPCA), as well as to selected measures of subjective well-being, allows on the one side, to utilize information on long-term process of local development and, on the other, to expand the analysis towards employing a spatio-temporal framework, while clarifying the between individual (micro) and commune (macro) level relationships.

► In summary, by taking into account the dynamic aspect of the local development process (by applying the FD approach) in the analysis of its impact on the personal well-being of residents, both planning and resource allocation policies become better informed and, one might expect, more effective (for example, a given level of individual well-being can be achieved with less effort due to using such additional information than would otherwise be the case).

References

Anand, P., 2018. The Value of Individual and Community Social Resources. In : New Frontiers of the Capability Approach (pp.436-72) Publisher: Cambridge University Press.DOI:10.1017/9781108559881.019

Fischer M.M., Getis A., 2010. Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer.

Górecki T., Krzyśko M., Wołyński W., 2019. Variable Selection in Multivariate Functional Data Classification. Statistics in Transition ne series. Vol. 20 (2), pp123-138.

Hong G. 2015. Causality in a Social World: Moderation. Mediation and Spill-over. Wiley.

Hox J. J., Moerbeek M., Schoot R., van de. 2018. Multilevel Analysis: Techniques and Applications. 3rd ed., New York, Routledge.

Kim Y., Ludwigs K., 2017, Measuring Community Well-Being and Individual Well-Being for Public Policy: The Case of thr Community Well-Being Atlas, in: R. Phillips, C. Wang, (ed.), *Handbook Of Community Well-Being Reseach*. Springer.

Krueger A. B., Kahneman D., Schkade D.A., Schwartz N., Stone A., 2009. *National Time Accounting: The Currency of Life,* in : A. B. Krueger (ed), Measuring Subjective Well-Being of Nations: National Account of Time Use and Well-Being. University of Chicago Press.

LeSage J P., Pace R.K., 2010. Spatial Econometrics. [in] Fischer and Getis (2010)

OECD 2013.*O*ECD Guidelines on Measuring Subjective Well-being, OECD Publishing.

References

Okrasa W., 2017. Community Wellbeing, Spatial Cohesion and Individual Wellbeing – towards a multilevel spatially integrated framework, [in] W. Okrasa (Ed.) Quality of Life and Spatial Cohesion: Development and Wellbeing in the Local Context. Cardinal Stefan Wyszynski University Press. Warsaw.

________ , Rozkrut, D., 2024. Assessing the community and individual well-being interaction within a spatial data-driven approach. Paper at the Conference *Small Area Estimation Survey & Data Science,* Pontificia Universidad Católica del Perú, Lima, June 3 – 7 (2024).

________ , Krzyśko, M., Wołyński, W., 2020. Spatio-temporal aspects of community well-being in Multivariate Functional Data approach. [in] C.H. Skiadas, C. Skiadas (eds.), Demography of Population Health, Aging and Health Expenditures, The Springer Series on Demographic Methods and Population Analysis 50 (pp.251-273)

_________, Rozkrut D., 2018. Modelling for Improving Measurement: Strategies for Contextualization of Well-Being. IAOS2018_OECD Conference *Better Statistics for Better Lives.* Paris, Sept. 19-21.

Phillips R., and Wong C,. 2017. Handbook of Community Well-Being Research, Springer.

Stauer N., Marks N., 2009. Local Wellbeing. Can We Measure it.? https://youngfoundation.org/wp-content/uploads/2013

Subramanian S.V., 2010. Multilevel Modeling [in] Fischer M.M., Getis A., Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer