Indeks 371262
e-ISSN 2657-9545
ISSN 0033-2372

Przegląd Statystyczny

Statistical Review

Vol. 71 No. 4 2024

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

INFORMATION FOR AUTHORS

Przegląd Statystyczny. Statistical Review publishes original research papers on theoretical and empirical topics in statistics, econometrics, mathematical economics, operational research, decision science and data analysis. The manuscripts considered for publication should significantly contribute to the theoretical aspects of the aforementioned fields or shed new light on the practical applications of these aspects. Manuscripts reporting important results of research projects are particularly welcome. Review papers, shorter papers reporting on major conferences in the field, and reviews of seminal monographs are eligible for publication based on the Editor-in-Chief's decision.

Since 1st May 2019, the journal has been publishing articles in English.

Any spelling style is acceptable as long as it is consistent within the manuscript.

All works should be submitted to the journal through the Editorial System (https://www.editorialsystem.com/pst).

For details of the submission process and editorial requirements, please visit https://ps.stat.gov.pl/ForAuthors.

Przegląd Statystyczny
Statistical Review

Vol. 71 No. 4 2024

EDITORIAL OFFICE'S ADDRESS

Statistics Poland (GUS), al. Niepodległości 208, 00-925 Warsaw, Poland

Language editing: Scientific Journals Division, Statistics Poland
Technical editing and typesetting: Statistical Publishing Establishment – team supervised by Maciej Adamowicz

Zakład Wydawnictw Statystycznych

Printed and bound: Statistical Publishing Establishment
al. Niepodległości 208, 00-925 Warsaw, Poland, zws.stat.gov.pl

Website: ps.stat.gov.pl

ISSN 0033-2372
e-ISSN 2657-9545
Index 371262

Information on the sales of the journal: Statistical Publishing Establishment
Phone no.: +48 22 608 32 10, +48 22 608 38 10

Order no. 174/2024

CONTENTS

Przegląd Statystyczny. Statistical Review, 2024, vol. 71, 4, 1–16 https://doi.org/10.59139/ps.2024.04.1
Received: 30.03.2025, revised: 11.06.2025, accepted: 11.06.2025

Spectral clustering and principal component analysis as tools for variable transformation in symbolic interval-valued data ensembles

Marcin Pełka[a]

Abstract. The selection, weighting and transformation of variables are essential phases of the modelling process. Two approaches can be applied to improve a model's accuracy: the selection of variables and the transformation of variables. In symbolic data analysis, two different approaches can be adopted: principal component analysis (PCA) and spectral clustering. In all the cases, we initially start with a set of symbolic variables and, after transformation, we obtain either classical variables (single numeric values) or symbolic variables that can be used in various models. The paper presents and compares PCA and spectral clustering for symbolic data when dealing with the problem of variable transformation. Artificial data with a known cluster structure were used to compare both single and ensemble clustering approaches. The results suggest that spectral clustering achieves better results for single and ensemble models.

Keywords: symbolic data analysis, ensemble learning, spectral clustering, principal component analysis

JEL: C63, C87, C90

1. Introduction

Machine learning techniques are very useful in dealing with discrimination tasks, as they are able to address various problems and are usually relatively accurate. In many cases, a one-digit prediction error can be achieved for different test sets (Meyer et al., 2003). Generally, it can be said that machine learning methods have reached a high level of complexity, adaptiveness, etc., and they can detect the relationships and rules that occur in a dataset.

Wolpert and Macready (1997) show that the search for the best method for solving all problems associated with machine learning is useless, as no such method exists. The choice of the method should be problem-based and related to a particular classification problem. In many cases combining (aggregating) different models (methods) can prove a good solution. Models that combine the results obtained from different models are known as ensemble learning (see for example Kuncheva, 2014; Polikar, 2012; Sagi & Rokach, 2018; Zhou, 2021). Hybrid models propose a similar solution (see for example: Ardabili et al., 2019; Tsai & Chen, 2010).

[a] Wrocław University of Economics and Business, Faculty of Economics and Finance, ul. Komandorska 118/120, 53–345 Wrocław, Poland, e-mail: marcin.pelka@ue.wroc.pl, ORCID: https://orcid.org/0000-0002-2225-5229.

In general, the main goal of cluster analysis is to obtain relatively homogeneous clusters, i.e. groups of objects that are similar when considering the variables used in cluster analysis. Usually, we would like the clusters to be isolated and cohesive (Gnanadesikan et al., 1995, p. 3). The key issue that has a major impact on the clustering process is the method of variable selection, as it affects the information that will be provided to the model.

Selecting as many variables as possible seems to be inefficient and time-consuming. As we consider the variable selection for clustering, we want the final clusters to be relatively homogeneous (Gnanadesikan et al., 1995; Guyon & Elisseeff, 2003). This can be achieved by:

- selecting weights for variables (representing the variables' 'importance' in clustering);
- variable selection, where from initial n variables, we select m $(m \leq n)$. This can be seen as a special case of weighting variables, where the selected ones are assigned weight 1 and those not selected 0;
- replacing the initial n variables with new variables. This is known as variable transformation. Such new variables might have some known and desirable properties.

This paper presents and compares two methods for variable transformation for symbolic data: principal component analysis (PCA) and spectral clustering. In the empirical part of the paper, datasets with a known number of clusters are used for single and ensemble clustering methods to compare how the proposed transformation methods affect the results of clustering. All the simulations were done using the R software.

This paper is organised as follows: Section 2 introduces the main aspects of symbolic data and shows PCA and spectral clustering as techniques for symbolic variable transformation. Section 3 presents the artificial datasets and ensemble clustering for symbolic data. The results for simple and ensemble clustering that are compared according to the adjusted Rand index are described in Section 4, while Section 5 summarises the key findings.

2. Variable transformation for symbolic data

Each symbolic object can be described by different variables (see Table 1 to view some examples). These variables can be (Billard & Diday, 2006; Bock & Diday, 2000):

1. Quantitative (numerical values):
   a) numerical single-valued variables
   b) numerical multi-valued variables
   c) interval-valued variables
   d) histogram variables;

2. Qualitative (categorical values):
   a) categorical single-valued variables
   b) categorical multi-valued variables
   c) categorical modal variables.

Table 1. Examples of symbolic variables

Symbolic variable	Realisations	Variable type
price of a car (in EUR)	(19,000; 23,000); (20,000; 35,000); (22,000; 37,000); (32,000; 47,000)	interval-valued (non-disjoint)
engine capacity (in ccm)	(1,000; 1,200); (1,300; 1,400) (1,500; 1,800); (1,900; 2,200)	interval-valued (disjoint)
chosen car colour	{red, black, blue, yellow} {magenta, white, grey, violet}	categorical multi-valued
preferred car brand	{Toyota (0.7); Audi (0.3)} {Skoda (0.6); VW (0.3); Other (0.1)}	categorical modal
distance travelled daily [in km]	<10, 20> (0.65); <21, 30> (0.35) <10, 20> (0.40); <21, 30> (0.60)	histogram
sex of a person	{male, female}	classical (nominal)
age of the customer	20, 30, 40, 55, 24, 35, 47	classical (ratio)

Source: author's work based on: Billard and Diday (2006), Bock and Diday (2000).

Symbolic data allow us to consider the uncertainty and variability in the data, enabling the description of objects in a new, more complex way. New methods, however, are necessary to analyse this type of data.

When dealing with the issue of variable transformation for symbolic interval-valued data, two approaches can be adopted:
a) PCA for symbolic data (PCA-SDA);
b) spectral clustering for symbolic data (SPEC-SDA).

The well-known PCA for classical data involves the following steps (Hair et al., 2010; Krzanowski, 2000):
a) obtaining correlation (or covariance) matrix $(\mathbf{R})$ for standardised data;
b) calculation of eigenvalues and eigenvectors for $\mathbf{R}$;
c) sorting eigenvalues and eigenvectors in ascending order and selecting the first s of them. As a result, a reduced matrix is obtained;
d) multiplying the initial data matrix by the reduced eigenvalue matrix.

When dealing with interval-valued symbolic data, several approaches (algorithms) can be applied in the case of PCA-SDA.

The first proposals where the mode (the value that appears most often in a set of data values) or the average used as representative of the interval-valued data were introduced by Nagabhushan et al. (1995). Cazes et al. (1997) and Chouakria et al. (2000) proposed vertices PCA (VPCA) and centres PCA (CPCA), where the vertices of the hyperboxes or centres of the interval were used as representatives of interval-valued symbolic data.

The VPCA was improved by Lauro et al. (2000) and Douzal-Chouakria et al. (2011) by introducing a label matrix and allowing for trivial intervals and generalised weight functions.

Palumbo and Lauro (2003) proposed a midpoint and radii PCA (MRPCA) by introducing a radius to the CPCA method. D'Urso and Giordani (2004) devised a way to use least squares for MRPCA, while Gioia and Lauro (2006) proposed the application of interval algebra for all the calculations. Le-Rademacher and Billard (2012) showed how to apply covariance to extend the classical PCA. Wang et al. (2012) introduced the complete information-based PCA (CIPCA). Chen et al. (2015) defined a covariance matrix for probabilistic symbolic data and presented a new PCA based on this variance-covariance structure.

Zuccolotto (2006) suggested describing objects by estimated means of a p-dimensional variable. Oliveira et al. (2017) proposed the use of truncated versions of symbolic principal components that apply a strict subset of the original symbolic variables as a way to improve the interpretation of symbolic principal components. Ichino (2011) introduced a new quantification method for symbolic PCA. The quantile method is applied for histogram and nominal multi-value types and other types of symbolic data at the time. Su and Wu (2024) suggested the adaptation of the symbolic PCA method for time series data.

In this paper, the CPCA, MRPCA and also methods based on the covariance matrix will be applied for dimensionality reduction in ensemble clustering methods.

In the CPCA, the $\mathbf{X}_C$ $(N \times p)$ matrix is calculated from the symbolic data matrix, where symbolic interval-valued data is replaced (substituted) by its midpoint (centre):

$$\mathbf{X}_C = \begin{bmatrix} x_{11}^c & \cdots & x_{1p}^c \\ \vdots & \ddots & \vdots \\ x_{N1}^c & \cdots & x_{Np}^c \end{bmatrix}, \tag{1}$$

where the centre is calculated as $x_{ij}^c = \frac{\underline{x}_{ij} + \overline{x}_{ij}}{2}$, with $\underline{x}_{ij}$ being the lower bound of the j-th symbolic interval-valued variable, and $\overline{x}_{ij}$ being the upper bound of the j-th symbolic interval-valued variable.

Matrix $\mathbf{X}_C$ contains the coordinates of the N hyper-rectangles. The well-known classical PCA is applied to this matrix. Then, all the vertices of each hyper-rectangle are projected in the obtained subspace and the lower-dimensional rectangles (if we extract only two principal components) are constructed with segments covering all the projections. This method assumes that the hyper-rectangle can be represented well by its centres and then the obtained subspace optimising the projection of the centres should also be optimal for the hyper-rectangles.

The PCA for symbolic interval-valued data can be additionally done by using ranges (radii) and midpoints (centres). In this case, the covariance matrix can take the following form:

$$\mathrm{Cov}(\mathbf{X}) = \frac{1}{n}(\mathbf{X}^C)^\mathrm{T}(\mathbf{X}^C) + \frac{1}{n}\Delta([\mathbf{X}])^\mathrm{T}\Delta([\mathbf{X}]) + \frac{1}{n}[\mathbf{X}^C\Delta([\mathbf{X}]) + \Delta([\mathbf{X}])^\mathrm{T}\mathbf{X}^C], \quad (2)$$

where $\mathbf{X}^C$ is the matrix of the midpoints (centres) and $\Delta([\mathbf{X}])$ is the standard variance-covariance matrix calculated for single-valued data.

Two independent PCAs should be singly exploited on those two matrices which, however, do not cover the whole variance. A solution to this problem is reflected in the formula: $\mathbf{X}^C\Delta([\mathbf{X}]) + \Delta([\mathbf{X}])^\mathrm{T}\mathbf{X}^C$. It takes into account the residual variance simultaneously and it allows for a logical, graphical representation of data. This is a well-known PCA on the interval midpoints whose solutions are given by $\mathbf{X}^C\Sigma^{-1}\mathbf{u}_m^c = \lambda_m^c\mathbf{u}_m^c$, with λ_m^c being defined under the usual orthonormality constraints. Similarly to the PCA that is based on midpoints, the solutions are obtained for ranges $\Delta([\mathbf{X}])\Sigma^{-1}\mathbf{u}_m^r = \lambda_m^r\mathbf{u}_m^r$, with the same orthonormality constraints for λ_m^r and $\mathbf{u}_m^r$.

Palumbo and Lauro (2003) suggest maximising the convergence coefficient between the midpoints and radii proposed by Tucker:

$$f(T) = \frac{t\prime_l\Delta([\mathbf{X}])_l^\mathrm{T}\mathbf{X}^C}{(t\prime_l\Delta([\mathbf{X}])_l^\mathrm{T}\Delta([\mathbf{X}])_l t)^{1/2}((\mathbf{X}^C)^\mathrm{T}\mathbf{X}^C)^{1/2}}, \quad (3)$$

where $[t_1, \ldots, t_l, \ldots, t_p]$ is the rotation matrix.

Furthermore, the PCA for symbolic interval-valued data can be done by using a covariance. In this case, the total sum of products (*SPT*) is decomposed into two components: the sum of products within (*SPW*) and the sum of products between (*SPB*), and these products are connected to the covariance:

$$nCov_{j_1,j_2} = SPT_{j_1,j_2} = SPW_{j_1,j_2} + SPB_{j_1,j_2}, \quad (4)$$

where $CovW_{j_1,j_2} = \frac{SWW_{j_1,j_2}}{n} = \frac{1}{n}\sum_{i=1}^{n}\frac{(\overline{x}_{ij1}-\underline{x}_{ij1})(\overline{x}_{ij2}-\underline{x}_{ij2})}{12}$ and $CovB_{j_1,j_2} = \frac{SBB_{j_1,j_2}}{n} =$ $= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\overline{x}_{ij1}-\underline{x}_{ij1}}{2} - \overline{X}_{j1}\right)\left(\frac{\overline{x}_{ij2}-\underline{x}_{ij2}}{2} - \overline{X}_{j2}\right)$.

The covariance approach for PCA utilises all the information in the symbolic data. The Cov matrix is decomposed into $CovW$ and $CovB$ matrices. This allows for a deeper insight into the PCA results for traces of these matrices.

Spectral clustering is not a new clustering method, but rather a new way of preparing the dataset for other clustering methods (e.g. k-means, hierarchical clustering, etc.). Finite-sample properties of spectral clustering were shown by Ng et al. (2002), Shi and Malik (2000).

Spectral clustering has the advantage of performing effectively in the presence of non-Gaussian clusters. Additionally, this approach is free from the drawback of the presence of local minima. The results obtained via spectral clustering in many cases outperform other well-known clustering methods (Luxburg, 2007). What is more, spectral clustering can detect clusters of different shapes, as it makes no assumptions according to the shape of clusters (Luxburg, 2007).

Spectral clustering, however, has its disadvantages. The choice of a good similarity graph is a challenging task and, usually, it entails a fully connected graph. Spectral clustering can also be unstable under different choices of the parameters for the neighbourhood graphs. Another problem is the selection of the kernel for spectral clustering. Many different kernels can be applied and each of them can lead to different outcomes. The Gaussian kernel tends to be used most often (see Karatzoglou, 2006, where the application of different kernels is presented).

Another issue in spectral clustering is the selection of a good σ parameter. This parameter should minimise the inter-cluster distances for a given number of clusters. Karatzoglou (2006) proposed an efficient algorithm for finding the optimal σ parameter.

The spectral clustering algorithm for symbolic data involves the following steps:

1. Let $\mathbf{X}$ be the symbolic data table with n rows and m columns. Let u be the number of clusters;
2. Let $\mathbf{A} = [a_{ik}]$ be an affinity matrix for the objects. This $\mathbf{A}$ matrix can be calculated in many ways and its elements can be defined as:

$$A_{ik} = \exp(-\sigma \cdot d_{ik}) \quad \text{for } i \neq k, \tag{5}$$

where σ is the scaling parameter that minimises the sum of the inter-cluster distances for a given number of clusters u, and d_{ik} is the distance between the i-th and k-th object;

3. Calculation of the Laplacian: $\mathbf{L} = \mathbf{D}^{1/2}\mathbf{A}\mathbf{D}^{1/2}$ (with $\mathbf{D}$ being the weight matrix with sums of each row from $\mathbf{A}$ on the diagonal);
4. Calculation of eigenvectors and eigenvalues of $\mathbf{L}$;
5. First u eigenvectors create the $\mathbf{E}$ matrix. Each eigenvector is treated as a column of $\mathbf{E}$, thus $\mathbf{E}$ has $n \times u$ dimensions;

6. Normalisation of **E** according to $y_{ij} = \frac{e_{ij}}{\sqrt{\sum_{j=1}^{u} e_{ij}^2}}$;
7. Finally, the **Y** matrix is the starting point for some clustering algorithms (i.e. k-means, hierarchical clustering).

The only difference between spectral clustering for classical and symbolic data lies in the applied distance measure. For details concerning distance measures for symbolic data, see Billard and Diday (2006) or Bock and Diday (2000).

3. Ensemble clustering for symbolic data

In general, ensemble learning methods are based on aggregated, combined results obtained from different models (clustering methods). These results can be seen as different points of view on the same dataset. Ensemble techniques have been applied with success in the context of supervised learning as they lead to improved accuracy and stability of algorithms (Breiman, 1996).

In ensemble clustering, we combine the results of N different models $(P_1, \dots, P_n)$ into one final clustering (aggregated clustering, ensemble clustering), i.e. P^* with k clusters (Fred & Jain, 2005).

There is a formal mathematical proof showing that in the case of ensemble learning in supervised tasks, the error reached by the ensemble is lower than any of the errors of the base models that form the ensemble (Gatnar, 2008).

Ensemble clustering can be seen as the solution to the problem with the selection of the clustering method. In this case, different clustering methods enable us to take 'different point of view' into account. Ensemble methods can prove effective when dealing with too few or too many data. If too many data occur, we can divide them into smaller, easier-to-learn partitions, and if there is a small amount of data, the same data can be used many times via bootstrapping techniques. Ensemble learning makes it also possible to deal with complex data or data too difficult to cluster. In this case, ensemble learning enables the data to be 'cut' into smaller, easier-to-learn parts, which is also known as the 'divide and conquer' approach. When dealing with many real-life problems involving decision-making, it is normal to consider information from many sources, known as information fusion (see for example Kuncheva, 2014; Zhou, 2012).

In the case of symbolic data, the following paths for ensemble clustering can be distinguished:

1. Clustering based on multiple relational matrices proposed by de Carvalho et al. (2012). The idea is based on various distance matrices (that can be obtained from different distance measures, subsets of objects or subsets of variables). These relational matrices are used to calculate relevance weight vectors. The relevance weight vectors and distance matrices are then used to group a set of objects into final clusters;

2. Applying one of the well-known ensemble clustering methods to symbolic data:
   a) proposal made by Leisch (1999), where many different clustering results are used to obtain cluster centres. Then these centres are used to obtain the final clusters. At the end, all objects are assigned to the nearest cluster. Medoids (cluster representatives) are used for symbolic data
   b) adaptation proposed by Dudoit and Fridlyand (2003), where cluster labels are permutated so that there is maximum overlap with the original clustering of these observations
   c) Hornik's (2005) idea to minimise the distance between the set of all the possible consensus clustering elements and all elements of the ensemble clustering;
3. Applying one of the consensus functions (Fred & Jain, 2005):
   a) hypergraph partitioning, which assumes that clusters can be represented as edges on a graph. Their vertices correspond to the objects to be clustered. Each edge describes a set of objects belonging to the same cluster. The problem of consensus clustering is reduced to finding the minimum cut of a hypergraph
   b) the voting approach, where we permute cluster labels in such a way that the best agreement between the labels of two partitions is obtained. All the partitions from the cluster ensemble must be relabelled according to a fixed reference partition
   c) mutual information which assumes that the objective function of a clustering ensemble can be formulated as the mutual information between the empirical probability distribution of labels in the consensus partition and the labels in the ensemble. A generalised definition of mutual information is usually applied in this approach
   d) co-association-based functions, where the main assumption is that objects which belong to the same cluster ('natural cluster') are co-located in the same clusters in different data partitions. The elements of the co-association matrix are defined as: $C(i,j) = \frac{n_{ij}}{N}$, where n_{ij} is the number of times that objects i and j are grouped in the same cluster (together) among all N base partitions
   e) finite mixture models, where the main assumption is that the output labels are modelled as random variables drawn from a probability distribution described as a mixture of multinomial component densities. The objective of consensus clustering is formulated as a maximum likelihood estimation.

In the empirical part, Leisch's (labelled LE), Hornik's (labelled HE), Dudoid and Fridlyand's (labelled DFE) and the co-clustering matrix (CCE) are used to obtain the final partitions (clusters) with the application of the Silhouette (Rousseeuw, 1987) clustering index to find the final number of clusters.

Although this index has some limitations, like bias toward convex or spherical clusters (see Dudek, 2020), high dimensions reduce its effectiveness (see Tomašev & Radovanović, 2016). This index is sensitive to noisy variables.

4. Single and ensemble clustering results

To compare how PCA and spectral clustering for symbolic data handle different shapes of clusters, the `cluster.Gen` function from the `clusterSim` package for the R software was used (Walesiak & Dudek, 2024). The `cluster.Gen` function allows the generation of various cluster shapes. To generate symbolic interval-valued data, the data for each model is generated twice, thanks to which sets A and B are obtained. Minimum value x_{ij}^{A}, x_{ij}^{B} is treated as the lower bound of the symbolic variable and the maximum is treated as the upper bound. The following simulation paths were used:

a) different PCA for symbolic interval-valued data were applied, then ensemble clustering methods were used (path P_1);
b) spectral clustering with different distance measures ($\sigma = 2$ in all models) was used, and the final $\mathbf{Y}$ matrix was applied for ensemble clustering (path P_2);
c) both PCA and spectral clustering were used with different initial settings (path P_3).

The following clustering methods were applied: partitioning around medoids (PAM), hierarchical-clustering (single-link), dynamic clustering for symbolic data (SClust) and clustering based on the distance matrix (DClust). Both `SClust` and `DClust` are functions of the `symbolicDA` package of R.

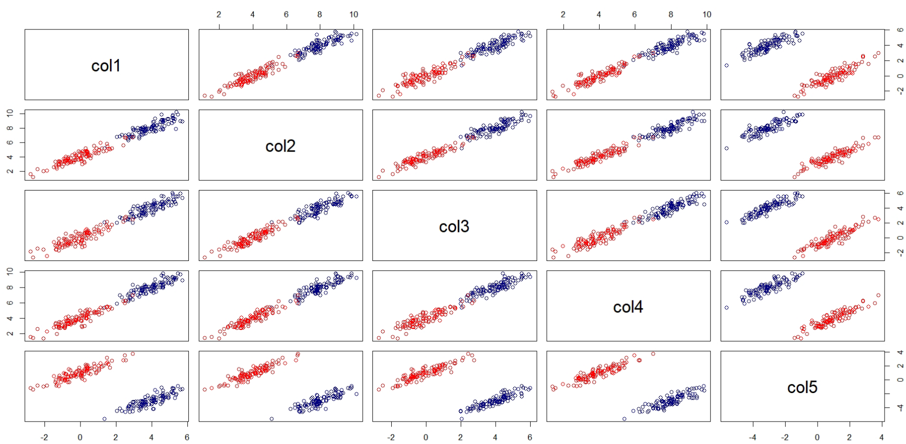
The following datasets with known cluster structures were prepared with the application of the `cluster.Gen` function from the `clusterSim` package of the R software:

a) set I: 100 objects in two well-separated clusters (see Figure 1) in five dimensions with means $(4, 8, 4, 8, -3)$, $(0, 4, 0, 4, 1)$ and covariance matrices $\Sigma_1(\sigma_{jj} = 1, \sigma_{jl} = 0.9)$, $\Sigma_2(\sigma_{jj} = 1, \sigma_{jl} = 0.5)$, $\Sigma_3(\sigma_{jj} = 1, \sigma_{jl} = -0.7)$, $\Sigma_4(\sigma_{jj} = 1, \sigma_{jl} = 0.84)$, $\Sigma_5 \begin{bmatrix} 1 & -0.45 & -0.45 & -0.45 & -0.45 \\ -0.45 & 1 & -0.56 & -0.56 & -0.56 \\ -0.45 & -0.56 & 1 & -0.58 & -0.58 \\ -0.45 & -0.56 & -0.58 & 1 & -0.74 \\ -0.45 & -0.56 & -0.58 & -0.74 & 1 \end{bmatrix}$;
b) set II: 100 objects in five not well-separated clusters in five dimensions (see Figure 2) with means $(5,5,5,5,5)$, $(-3,3,-3,3,-3)$, $(0,0,0,0,0)$, $(-5,-5,-5,-5,-5)$

and covariance matrices $\sum_1 \begin{bmatrix} 1 & -0.9 & -0.9 & -0.9 & -0.9 \\ -0.9 & 1 & -0.7 & -0.7 & -0.7 \\ -0.9 & -0.7 & 1 & -0.85 & -0.85 \\ -0.9 & -0.7 & -0.85 & 1 & -0.9 \\ -0.9 & -0.7 & -0.85 & -0.9 & 1 \end{bmatrix}$, $\sum_2(\sigma_{jj} = 1, \sigma_{jl} = 0.9)$, $\sum_3(\sigma_{jj} = 3, \sigma_{jl} = 1.5)$, $\sum_4(\sigma_{jj} = 1, \sigma_{jl} = 0)$, $\sum_5(\sigma_{jj} = 1, \sigma_{jl} = 0.2)$;

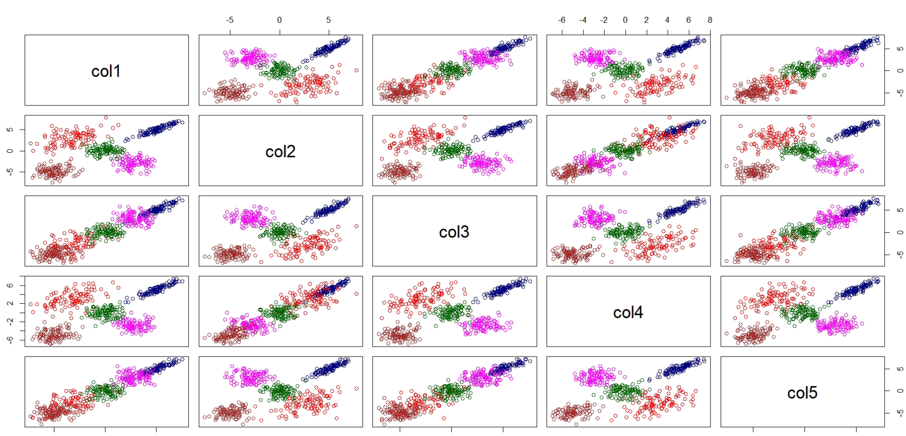
c) set III: 100 objects in three well-separated clusters in five dimensions (see Figure 3) with means $\sum_1(\sigma_{jj} = 1, \sigma_{jl} = 0)$, $\sum_2(\sigma_{jj} = 1, \sigma_{jl} = -0.9)$, $\sum_3(\sigma_{jj} = 1, \sigma_{jl} = 0.9)$, $\sum_4(\sigma_{jj} = 3, \sigma_{jl} = 1)$, $\sum_5(\sigma_{jj} = 1, \sigma_{jl} = -0.5)$.

Figure 1. Two well-separated clusters in five dimensions (set I)



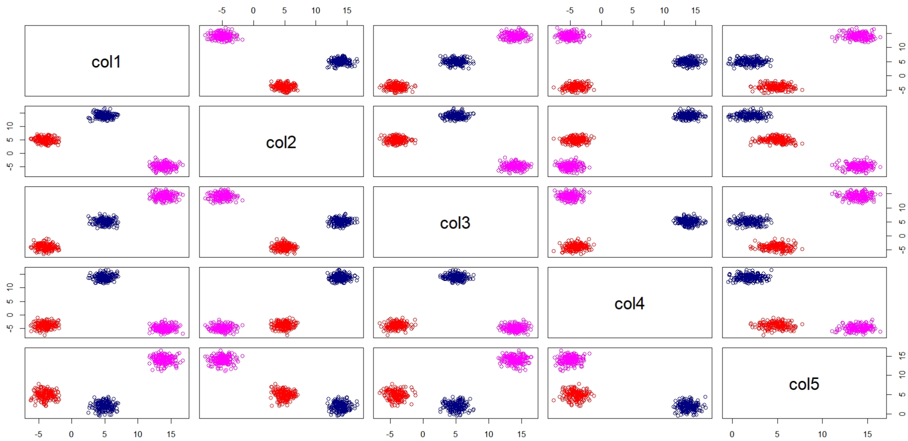
Source: author's work based on the R software.

Figure 2. Five not well-separated clusters in five dimensions (set II)



Source: author's work based on the R software.

Figure 3. Three well-separated clusters in five dimensions (set III)



Source: author's work based on the R software.

To compare how PCA and spectral clustering perform for symbolic data, 50 simulations were done and the average adjusted Rand index (Rand, 1971) was calculated. The average adjusted Rand index values are shown in Table 2 for single clustering methods and in Table 3 for ensemble clustering results.

Table 2. Simulation results for single models: all paths and datasets

Single model	Dataset	Adjusted Rand index value	Calculation time
pam (cPCA)	I	0.9984	2.95938 mins
	II	0.9862	3.68677 mins
	III	1	3.50588 mins
pam (mrPCA)	I	1	2.94649 mins
	II	0.9982	3.92382 mins
	III	0.7349	3.67878 mins
pam (covPCA)	I	0.9943	4.12835 mins
	II	1	2.66233 mins
	III	1	57.90654 secs
pam (specl & H)	I	0.9959	2.96048 mins
	II	0.9801	2.23560 mins
	III	0.9978	4.96317 mins
pam (specl & U_2)	I	1	2.90862 mins
	II	0.9993	3.01923 mins
	III	1	4.97362 mins
pam (specl & SO_1)	I	0.9424	2.60051 mins
	II	0.9720	7.33325 mins
	III	1	5.86112 mins
DClust (cPCA)	I	1	3.58567 mins
	II	0.5912	10.95904 mins
	III	1	1.93726 mins

Table 2. Simulation results for single models: all paths and datasets (cont.)

Single model	Dataset	Adjusted Rand index value	Calculation time
DClust (mrPCA)	I II III	0.7723 0.3434 1	1.72763 mins 10.71638 mins 3.84637 mins
DClust (covPCA)	I II III	0.9882 0.5864 0.5567	1.72323 mins 10.70673 mins 3.89015 mins
DClust (specl & H)	I II III	1 0.8065 0.8899	2.54698 mins 16.12656 mins 11.30388 mins
DClust (specl & U_2)	I II III	0.9899 0.7422 0.7597	13.63857 mins 16.00192 mins 5.76498 mins
DClust (specl & SO_1)	I II III	0.9616 0.5486 1	2.61785 mins 15.94209 mins 5.74380 mins

Note. pam – partition around medoids, DClust – dynamic clustering based on the distance matrix, cPCA – centres PCA, mrPCA – midpoints and radii PCA, covPCA – covariance-based PCA, specl – spectral clustering, H – Hausdorff distance, U_2 – Ichino-Yaguchi distance, SO_1 – de Carvalho distance.

Source: author's work based on the R software.

Table 3. Simulation results for ensemble models: all paths and datasets

Ensemble model	Dataset	Adjusted Rand index value		
		Path P_1	Path P_2	Path P_3
LE	I II III	0.8728 0.6533 0.8672	0.8989 0.7623 0.9123	0.9014 0.9765 0.9826
HE	I II III	0.9836 0.9123 0.9635	1 1 1	1 1 1
DFE	I II III	0.9927 0.9563 0.9972	1 0.9991 1	1 0.9831 1
CCE	I II III	0.9873 0.9654 0.9862	1 0.9864 1	1 1 1

Note. LE – Leich's ensemble, HE – Hornik's ensemble, DFE – Dudoit and Fridlyand's ensemble, CCE – co-clustering matrix ensemble.

Source: author's work based on the R software.

Dataset II (five not well-separated clusters in five dimensions) was the most challenging to cluster for all of the methods. However, the classical pam method combined with either the PCA or the spectral approach for symbolic data outperformed the DClust method designed for symbolic data.

When we look at the ensemble results (Table 3), we can see that Hornik's ensemble model, as well as Dudoit and Fridlyand's ensemble models achieved the highest average values of the adjusted Rand index across all model types. Similarly to the single model results, the most challenging task was to detect clusters by the ensemble models in Dataset II. Nevertheless, Hornik's, and Dudoit and Fridlyand's ensemble models perform most efficiently.

5. Conclusions

Two different approaches for symbolic data transformation have been shown in the paper for ensemble learning with this type of data. The first approach uses the well-known PCA applied for symbolic data, the second one utilises spectral clustering. Additionally, a combination of PCA and the spectral approach was used in the ensembles.

However, PCA for symbolic data is limited to symbolic interval-valued data only, while spectral clustering for symbolic data can handle various symbolic data types, as it requires only an appropriate distance measure for symbolic data. Notably, there are many different symbolic distance measures suitable for various symbolic variable types.

Ensemble clustering for symbolic data enables the integration of different clustering results ('points of view') to achieve a single, improved and more stable clustering outcome. In ensemble clustering, the key steps such as variable selection, variable weighting and variable transformation remain critical, as in the case of a single clustering method.

The results indicate that complex datasets (e.g. those with intricate cluster structures, outliers or noisy variables) are challenging for single symbolic clustering methods based on PCA. However, PCA combined with spectral clustering, as well as spectral clustering alone performs most effectively with such datasets (as measured by the Adjusted Rand Index). Similar trends are observed with ensemble clustering methods for symbolic data. Specifically, symbolic ensemble clustering techniques such as Hornik's, and Dudoit and Fridlyand's methods generally outperform a co-clustering (co-occurrence) matrix and Leisch's ensemble methods when dealing with complex data structures.

References

Ardabili, S., Mosavi, A., & Várkonyi-Kóczy, A. R. (2019). Advances in Machine Learning Modeling Reviewing Hybrid and Ensemble Methods. In A. R. Várkonyi-Kóczy (Ed.), *Engineering for Sustainable Future. Selected papers of the 18th International Conference on Global Research and Education Inter-Academia – 2019* (pp. 215–227). Springer. https://doi.org/10.1007/978-3-030-36841-8_21.

Bock, H.-H., & Diday, E. (2000). *Analysis of Symbolic Data. Explanatory Methods for Extracting Statistical Information from Complex Data*. Springer. https://doi.org/10.1007/978-3-642-57155-8.

Billard, L., & Diday, E. (2006). *Symbolic Data Analysis. Conceptual Statistics and Data Mining*. Wiley. https://doi.org/10.1002/9780470090183.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1007/BF00058655.

de Carvalho, F. A. T., Lechevallier, Y., & de Melo, F. M. (2012). Partitioning hard clustering algorithms based on multiple dissimilarity matrices. *Pattern Recognition*, *45*(1), 447–464. https://doi.org/10.1016/j.patcog.2011.05.016.

Cazes, P., Chouakria, A., Diday, E., & Schektrman, Y. (1997). Extensions de l'Analyse en Composantes Principales à des données de type intervalle. *Revue de Statistique Appliquée*, *45*(3), 5–24. http://eudml.org/doc/106421.

Chen, M., Wang, H., & Quin, Z. (2015). Principal component analysis for probabilistic symbolic data: a more generic and accurate algorithm. *Advances in Data Analysis and Classification*, *9*, 59–79. https://doi.org/10.1007/s11634-014-0178-2.

Chouakria, A., Diday, E., & Cazes, P. (2000). Vertices Principal Components Analysis With an Improved Factorial Representation. In A. Rizzi, M. Vichi & H.-H. Bock (Eds.), *Advances in Data Science and Classication* (pp. 397–402). Springer. https://doi.org/10.1007/978-3-642-72253-0_54.

Dudek, A. (2020). Silhouette Index as Clustering Evaluation Tool. In K. Jajuga, J. Batóg, M. Walesiak (Eds.), *Classification and Data Analysis. SKAD 2019. Studies in Classification, Data Analysis, and Knowledge Organization.* Springer. https://doi.org/10.1007/978-3-030-52348-0_2.

Douzal-Chouakria, A., Billard, L., & Diday, E. (2011). Principal component analysis for interval-valued observations. *Statistical Analysis and Data Mining*, *4*(2), 229–246. https://doi.org/10.1002/sam.10118.

Dudoit, S., & Fridlyand, J. (2003). Bagging to improve the accuracy of a clustering procedure. *Bioinformatics*, *19*(9), 1090–1099. https://doi.org/10.1093/bioinformatics/btg038.

D'Urso, P., & Giordani, P. (2004). A least squares approach to principal component analysis for interval valued data. *Chemometrics and Intelligent Laboratory Systems*, *70*(2), 179–192. https://doi.org/10.1016/j.chemolab.2003.11.005.

Fred, A. L. N., & Jain, A. K. (2005). Combining multiple clustering using evidence accumulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(6), 835–850. https://doi.org/10.1109/TPAMI.2005.113.

Gatnar, E. (2008). *Podejście wielomodelowe w zagadnieniach dyskryminacji i regresji*. Wydawnictwo Naukowe PWN.

Gioia, F., & Lauro, C. N. (2006). Principal component analysis on interval data. *Computational Statistics*, *21*(2), 343–363. https://doi.org/10.1007/s00180-006-0267-6.

Gnanadesikan, R., Kettenring, J. R., & Tsao, S. L. (1995). Weighting and selection of variables for cluster analysis. *Journal of Classification*, *12*, 113–136. https://doi.org/10.1007/BF01202271.

Guyon, I., & Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*, *3*, 1157–1182. https://doi.org/10.1162/153244303322753616.

Hair, J. E., Black, W. C., Babin, J. B., & Anderson, R. E. (2010). *Multivariate Data Analysis* (17th edition). Prentice Hall.

Hornik, K. (2005). A CLUE for CLUster Ensembles. *Journal of Statistical Software*, *14*(12), 65–72. https://doi.org/10.18637/jss.v014.i12.

Ichino, M. (2011). The quantile method for symbolic principal component analysis. *Statistical Analysis and Data Mining*, *4*(2), 184–198. https://doi.org/10.1002/sam.10111.

Karatzoglou, A. (2006). *Kernel Methods. Software, Algorithms and Applications* [Doctoral dissertation, Vienna University of Technology]. https://resolver.obvsg.at/urn:nbn:at:at-ubtuw:1-14467.

Krzanowski, W. J. (2000). *Principles of Multivariate Analysis. A User's Perspective*. Oxford University Press. https://doi.org/10.1093/oso/9780198507086.001.0001.

Kuncheva, L. I. (2014). *Combining Pattern Classifiers. Methods and Algorithms* (2nd edition). John Wiley and Sons.

Lauro, N. C., Verde, R., & Palumbo, F. (2000). Factorial Methods with Cohesion Constraints on Symbolic Objects. In H. A. L. Kiers, J.-P. Rasson, P. J. F. Groenen & M. Schader (Eds.), *Data Analysis, Classification, and Related Methods* (pp. 381–386). Springer. https://doi.org/10.1007/978-3-642-59789-3_60.

Leisch, F. (1999). *Bagged clustering* (SFB Working Papers No. 51). https://doi.org/10.57938/9b129f95-b53b-44ce-a129-5b7a1168d832.

Le-Rademacher, J., & Billard, L. (2012). Symbolic covariance principal component analysis and visualization for interval-valued data. *Journal of Computational and Graphical Statistics*, *21*(2), 413–432. https://doi.org/10.1080/10618600.2012.679895.

Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, *17*(4), 395–416. https://doi.org/10.1007/s11222-007-9033-z.

Meyer, D., Leisch, F., & Hornik, K. (2003). The support vector machine under test. *Neurocomputing*, *55*(1–2), 169–186. https://doi.org/10.1016/S0925-2312(03)00431-4.

Nagabhushan, P., Chidananda Gowda, K., & Diday, E. (1995). Dimensionality reduction of symbolic data. *Pattern Recognition Letters*, *6*(2), 219–223. https://doi.org/10.1016/0167-8655(94)00085-H.

Ng, A. Y., Jordan, M. I., & Weiss, Y. (2002). On Spectral Clustering: Analysis and an algorithm. In T. Dietterich, S. Becker & Z. Ghahramani (Eds.), *Advances in Neural Information Processing Systems* (pp. 849–856). MIT Press.

Oliveira, M. R., Vilela, M., Pacheco, A., Valadas, R., & Salvador, P. (2017). Extracting Information from Interval Data Using Symbolic Principal Component Analysis. *Austrian Journal of Statistics*, *46*(3–4), 79–87. https://doi.org/10.17713/ajs.v46i3-4.673.

Palumbo, F., & Lauro, C. N. (2003). A PCA for interval-valued data based on midpoints and radii. In H. Yani, A. Okada, K. Shigemasu, Y. Kano & J. J. Meulman (Eds.), *New Developments in Psychometrics* (pp. 641–648). Springer. https://doi.org/10.1007/978-4-431-66996-8_74.

Polikar, R. (2012). Ensemble learning. In C. Zhang & Y. Ma (Eds.), *Ensemble Machine Learning* (pp. 1–34). Springer. https://doi.org/10.1007/978-1-4419-9326-7_1.

Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, *66*(336), 846–850. https://doi.org/10.1080/01621459.1971.10482356.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7.

Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews*: *Data Mining and Knowledge Discovery*, *8*(4). https://doi.org/10.1002/widm.1249.

Shi, J., & Malik, J. (2000). Normalized Cuts and Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(8), 888–905. https://doi.org/10.1109/34.868688.

Su, E. C.-Y., & Wu, H.-M. (2024). Dimension reduction and visualization of multiple time series data: a symbolic data analysis approach. *Computational Statistics*, *39*(4), 1937–1969. https://doi.org/10.1007/s00180-023-01440-7.

Tomašev, N., & Radovanović, M. (2016). Clustering Evaluation in High-Dimensional Data. In M. Celebi & K. Aydin (Eds.), *Unsupervised Learning Algorithms* (pp. 71–107). Springer. https://doi.org/10.1007/978-3-319-24211-8_4.

Tsai, C. F., & Chen, M. L. (2010). Credit rating by hybrid machine learning techniques. *Applied Soft Computing*, *10*(2), 374–380. https://doi.org/10.1016/j.asoc.2009.08.003.

Walesiak, M., & Dudek, A. (2024). *Package 'clusterSim': Searching for Optimal Clustering Procedure for a Data Set*. R package version 0.51-5. https://cran.r-project.org/web/packages/clusterSim/clusterSim.pdf.

Wang, H., Guan, R., & Wu, J. (2012). CIPCA: complete-information-based principal component analysis for interval-valued data. *Neurocomputing*, *86*, 158–169. https://doi.org/10.1016/j.neucom.2012.01.018.

Wolpert, D. H., & Macready, W. G. (1997). No Free Lunch Theorems for Optimization. *IEEE Transactions on Evolutionary Computation*, *1*(1), 67–82. https://doi.org/10.1109/4235.585893.

Zhou, Z. H. (2012). *Ensemble Methods. Foundations and Algorithms*. CRC Press. https://tjzhifei.github.io/links/EMFA.pdf.

Zhou, Z. H. (2021). *Machine Learning*. Springer. https://doi.org/10.1007/978-981-15-1967-3.

Zuccolotto, P. (2006). Principal components of sample estimates: an approach through symbolic data analysis. *Statistical Methods & Applications*, *16*(2), 173–192. https://doi.org/10.1007/s10260-006-0024-6.

Przegląd Statystyczny. Statistical Review, 2024, vol. 71, 4, 17–36 https://doi.org/10.59139/ps.2024.04.2
Received: 20.03.2025, revised: 15.06.2025, accepted: 23.06.2025

When public policy matters: public intervention and satisfaction with the government

Sebastian A. Roy[a]

Abstract. This paper contributes to the economic voting literature by evaluating the effect of a policy-determined income shock on individual-level support for the government. It examines the results of a natural experiment relating to the 2016 Polish job market reform that raised the minimal hourly wage to PLN 12 per hour for all workers, regardless of their contract type, thus asymmetrically affecting low-earning individuals engaged in precarious work. The study employs a detailed microeconomic dataset from the European Social Survey to perform a difference-in-differences analysis. It concludes that the reform had a significant, positive impact on the levels of support received by the government, notably higher than any macroeconomic variable.

Keywords: economic voting, natural experiment, difference-in-differences
JEL: C31, C33, H53, J38

1. Introduction

The theoretical literature on voting highlights two pivotal dimensions of the election design: moral hazard, associated with incentivising politicians to implement policies preferred by the public (Barro, 1973; Feltovich & Giovannoni, 2015), and selection, i.e. allowing people to elect candidates with high individual talents (Duch & Stevenson, 2008; Fearon, 1999; List & Sturm, 2006). A necessary condition for the former concept to be effective is the voters' ability to correctly assess the government's decisions. Such an assessment may cover a broad range of issues, out of which economics plays a vital role (Lewis-Beck & Stegmaier, 2000). Thus, there is a seizable strand of empirical literature evaluating the impact of macroeconomic variables on election output (for summary see ibid.; Brug et al., 2007; Duch & Stevenson, 2008).

Purely macroeconomic studies of economic voting based on either aggregate or individual-level data have been criticised for omitting certain political factors (Powell & Whitten, 1993) or, more recently, for neglecting subjective wellbeing measures (Ward, 2019). Furthermore, the existing body of literature does not offer insights into the effects of singular economic policy interventions which directly impact particular voters' financial situation. Arguably, this might lead to systematic underestimation of the impact of the public policies: intuitively, an individual wage increase by 5% is likely to expand the voter's support for the government more considerably than a 5% GDP growth in the national economy. This article provides an empirical

[a] Doctoral School, SGH Warsaw School of Economics, Poland, e-mail: sr68731@doktorant.sgh.waw.pl, ORCID: https://orcid.org/0000-0002-5537-7978.

measure of the impact of a policy-induced income shock on individual-level support for the government, using Poland's 2016 labour market reform as a natural experiment. Difference-in-differences (DiD) regression is applied as an evaluation tool of the treatment effect, accurately capturing the causal relation between the minimal hourly wage increase and the response in terms of the political support shown by those treated.

The labour market intervention in question offers a unique setting to study the relation between public policy and individual support for the government due to the intervention's specific construction and certain features of the Polish labour market. The key element of the 2016 reform involved raising the minimal hourly wage to PLN 12 per hour for all workers, regardless of their contract type, thus equalling the minimal earnings applicable to employment and Civil Code contract[1] holders (Sejm Rzeczypospolitej Polskiej, 2016). These characteristics allow the identification of the treatment group. Thus, its members had to simultaneously satisfy both of the following conditions in 2016:

1. work under a Civil Code-based contract (non-employment, usually service);
2. earn less than PLN 12 per hour.

Obviously, the treatment group is not random. Therefore, the conclusions drawn from the treatment analysis can be generalised if and only if the selection for the treatment satisfies conditional exogeneity. A number of control variables and population weights were employed to make the generalisation possible.

This paper contributes to the economic voting literature by studying the effect that a policy-induced, individual positive income shock has on the level of satisfaction with the government. Using treatment assessment methodology (DiD) allows the identification of a causal relationship, while accounting for local-specific political and policy control measures provides an opportunity for robust inference.

The article is organised in the following way: Section 2 presents a review of literature on economic voting and discusses this paper's contribution into it. Section 3 focuses on the Polish minimal wage reform of 2016. Section 4 provides an overview of the dataset along with detailed definitions and assumptions used for the creation of the variables. The empirical DiD model and robustness checks are discussed in Section 5. Section 6 consists of final remarks.

2. Literature review

2.1. Benchmark literature

The literature on economic voting can be roughly divided into two parts: the literature on popularity function and vote function analysis (Nannestad & Paldam, 1994; Stegmaier et al., 2017). The former deals with government support rates derived from public opinion surveys, while the latter explores the impact of economic factors on

[1] Polish Civil Code includes two benchmark type of contracts: Service Contract (Pol. umowa zlecenia) and Mandate Contract (Pol. umowa o dzieło), each of which may be used to evade the Labour Code provisions.

voting choices. The popularity function research dates back to the 1970s (Kenski, 1977; Monroe, 1978; Mueller, 1970). The literature concerning the vote function focuses either on the explanation of (Erikson, 1989; Tufte, 1978) or forecasting (Campbell & Garand, 2000) election results. These early analyses usually employ scarce economic regressors, often at macro-aggregate level only. Certain specifications of this type of economic vote models are presented in Table 1.

Table 1. Selected specifications of early economic vote models

Study	Dependent variable	Economic independent variables	Political controls (if included)
Mueller, 1970	US President popularity	Economic slump (% unemployment)	Yes
Monroe, 1978	US President popularity	Inflation and defence spending (with lag distribution)	No
Tufte, 1978	US President vote function	Income	Yes
Erikson, 1989	US President vote function	Income	Yes

Source: author's work.

The papers above set a benchmark for all studies to follow. However, due to aggregation, they neither control for particular voter group characteristics nor allow the study of the impact that specific policy interventions have on voter political attitudes. An interesting attempt to include individual-level, microeconomic factors into economic voting analysis was proposed by Fiorina (1978). He found that individual financial conditions affect decisions at the ballot box, but they seem uncorrelated with the decision whether to vote or not.

2.2. Refinements

Historically, efforts were made to embed voter heterogeneity, especially with respect to asset holding, into economic voting analyses. These efforts produced literature on what is called patrimonial economic voting. Its standard result is that asset-rich voters tend to support right-wing parties, while those less affluent the left-wing. More recently, the patrimonial voting theory has been reconciled with political factor analysis by Hellwig and McAllister (2019). They claim that patrimonial model predictions are accurate as long as the candidates do not differ in the proposed economic policies.

Another important refinement takes into account the role of welfare as a specific economic factor likely to impact voting decisions. Pacek and Radcliff (1995) argue that as welfare spending expands, the effect of the economy on voting decisions diminishes. Their results are supported by Park and Shin (2017), who further control for clarity of the government's political responsibility. However, contradictory evidence, although scarce, exists (Palmer & Whitten, 2003) alongside some inconclusive studies (Brug et al., 2007).

Finally, a variety of literature is emerging on subjective wellbeing measures rather than macroeconomic variables as ballot box outcome determinants. The happiness-voting relation has been studied e.g. by Ward (2019) or Liberini et al. (2017). The former argues that the subjective wellbeing indicator (SWB) explains a higher fraction of incumbent vote share than any standard macroeconomic factor. The latter authors confirm this finding with respect to voting intention. They also contribute to the better understanding of the causal relationship between subjective happiness shocks and individual voting patterns, studying the electoral impact of a spouse's death with the DiD analysis.

3. The Polish minimal wage reform of 2016

3.1. Overview

The duality of the Polish labour market remains a persistent social and economic problem (Kamińska et al., 2014). Until 2016, Civil Code contracts in Poland had not needed to comply with the minimal wage regulations[2]. Moreover, holders of multiple Civil Code contracts were allowed to deduct social insurance contributions from only one of them, thus largely increasing the net-to-gross income ratio. Finally, compared to employment, Civil Code contracts offered significantly more flexibility, e.g. no fixed period of notice, no paid sick leave and no maternity leave.

Hence, strong incentives were convincing enough for employers to shift workers from employment into Civil Code contract-based work. The practice is estimated to have affected from almost 200,000 to over 800,000 workers in 2016 (Chrostek et al., 2019). The estimates are based on the following assumption: individuals who repeatedly sign non-employment contracts with the same entity for similar remunerations are suspected to be actual employees evading the Labour Code. A detailed summary of such a practice is presented in Table 2[3].

Table 2. Number of workers with suspected employment substitution by Civil Code contracts; estimates by total contract duration and wages deviation

Contract duration (no. of months)	Wages deviation			
	1%	5%	10%	20%
6	347,000	469,000	622,000	847,000
8	276,000	361,000	462,000	606,000
10	225,000	265,000	325,000	408,000
12	172,000	184,000	202,000	235,000

Source: Chrostek et al. (2019).

[2] In 2016, the minimal salary for employment contract holders was set at PLN 1,850 per month, equivalent to approx. PLN 12 per hour in full-time employment (Sejm Rzeczypospolitej Polskiej, 2016).

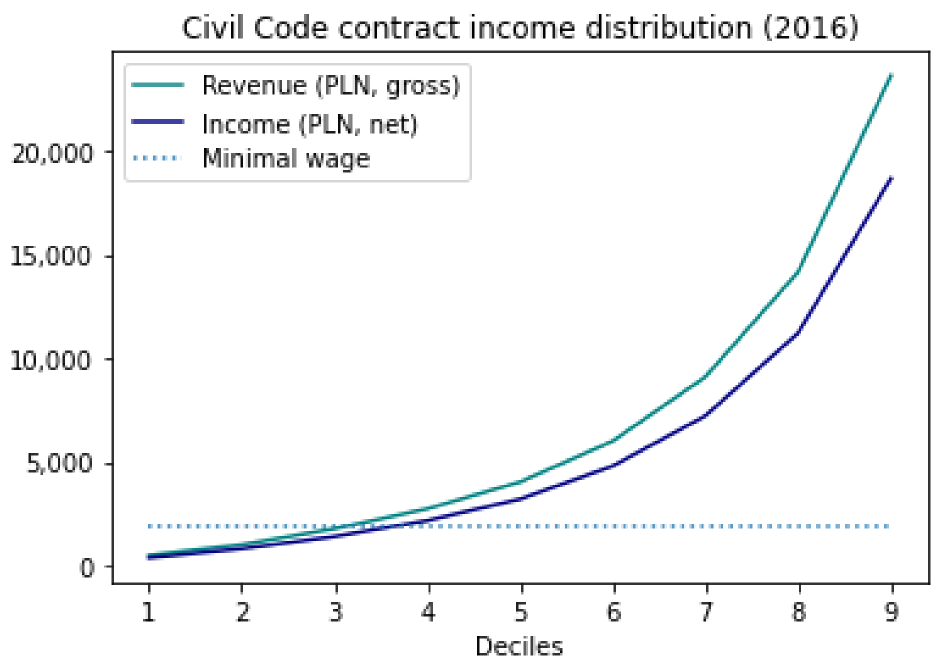
[3] As discussed by Chrostek et al. (2019), it is not straightforward to tell from the official data what the number of fictitious Labour Code-evading Civil Code-based contracts is. The figures in Table 2 report on the number of individuals who for a given number of months in one year signed contracts with wages that did not differ from the yearly mean more than the given percentage.

As noted by the Legislator, some of the Civil Code contract-based workers had been offered significantly lower salaries than the minimal wage guaranteed by the Labour Code (Sejm Rzeczypospolitej Polskiej, 2016, Uzasadnienie). The substitution of employment contracts with Civil Code ones became a malpractice used to evade the Labour Code and limit the costs of employment. Thus, in 2016, the Sejm of the Republic of Poland passed a bill which banned wages lower than PLN 12 per hour of work, regardless of the legal form of the contract between the worker and employer (ibid.). The reform was designed to equalise the minimal wage protection between different classes of contract holders, including occasional, part-time and temporary workers. However, it did not concern non-pecuniary provisions of contracts not specified in the Labour Code. Thus, the reform increased the wages of only those workers who earned less than PLN 12 per hour and did not hold a standard employment contract.

3.2. Treatment group characteristics

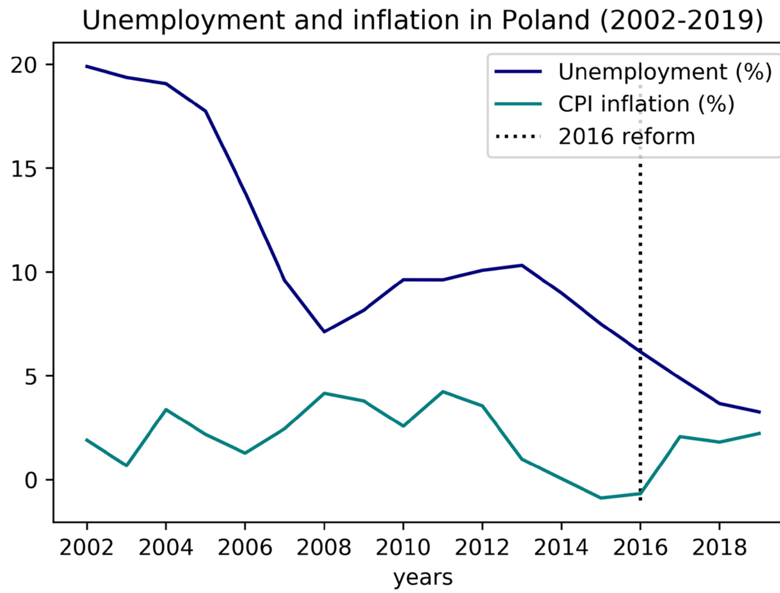
The 2016 minimal wage reform covered individuals forming the lowest-earning subset of the Civil Code contract holders. Out of all, only those whose earnings fell below the minimal wage in 2016 (i.e. PLN 1,850 per month or, equivalently, PLN 12 per hour) were subject to the reform-induced pay rise. As shown by Chrostek et al. (2019), they might have represented up to 30% of the total number of Civil Code contract holders (see Figure 1).

Figure 1. Income distribution in Poland (2016) from Civil Code contracts



Source: Chrostek et al. (2019).

Figure 2. Rates of unemployment and inflation in Poland (2002–2019)



Source: World Bank.

The assumption of a random, exogenous selection into the 2016 reform treatment group is, arguably, not satisfied. Given the outstandingly low unemployment in Poland in 2016 (see Figure 2) and the fact that, as shown by Statistics Poland (Główny Urząd Statystyczny [GUS], 2018b), median monthly salaries in the elementary occupations sector reached almost PLN 2,500 (in comparison to the PLN 1,850 threshold of the reform), it should be expected that the treatment group consisted of the most vulnerable individuals with the weakest bargaining position in the labour market. Hence, the selection into this treatment group is likely to have been correlated with low education, low skills and particular occupations (menial, repetitive, physical jobs).

Endogenous selection into the treatment group demands particular attention to the risk of the omitted variable bias. More specifically, some factors could be correlated both with the subjective satisfaction with the government and the probability of being in the treatment group. For example, people with high education might be in general more critical towards the government, while also significantly less exposed to minimal wage-earning jobs. Not accounting for that would lead to the overestimation of the reform's effect on the satisfaction with the government.

Another issue impeding empirical identification is simultaneity. Any factor influencing the government's popularity among voters, if introduced simultaneously to the treatment of interest, may be erroneously interpreted as an effect of the minimal wage reform (if it is not specifically accounted for). Actually, in 2016 another welfare transfer (a child benefit of PLN 500 monthly per child[4]) was implemented (Ustawa z dnia 11 lutego 2016 r. o pomocy państwa w wychowywaniu dzieci), whose effect must be included in the empirical analysis.

However, the 2016 reform provides a favourable opportunity to study the effect of an individual-level income shock on satisfaction with the government for two reasons. Firstly, once the observable characteristics of workers earning less than PLN 12 per hour under a Civil Code-based contract are accounted for, the treatment remains uncorrelated with individual support for government or any unobservable psychological traits. It is impossible to raise minimal wages of only those who have a particularly positive attitude towards the government. Thus, the reform may be considered a conditionally exogenous shock on a treatment group with certain observable characteristics. Secondly, remaining in or exiting the treatment group does not depend on the individual decision at the ballot box after the introduction of the reform. In other words, exposure to the effects of the reform in the years following 2016 did not depend on one's political predilections. As discussed in Farré et al. (2015)[5], widespread events such as sector collapses or factory closures, due to their exogenous character and independence from individual choices, might prove useful in retrieving particular causal effects of interest.

It is also worth mentioning that the general class of Civil Code contract holders obviously does not constitute a homogenous group partly due to the scale of Labour Code evasion: estimates show that 1 out of 5 Civil Code contracts may be in fact fraudulent (Arak et al., 2014). Low social security contributions attract workers of different occupations and education levels, including high-skilled experts such as researchers, journalists or IT developers. According to a survey by Sedlak&Sedlak, 70% of respondents would willingly give up a Labour Code employment contract in exchange for a higher salary.

4. The dataset

4.1. European Social Survey

This study uses individual-level data from the European Social Survey (ESS; rounds 3 to 9). The survey is conducted by the European Commission's ESS European Research Infrastructure every two years. Altogether, the sample consists of 15,624 individual

[4] For details see Section 4.3.2.

[5] Farré et al. (2015) use the Spanish 2007 construction sector collapse as an instrument for unemployment in their study of the joblessness in connection with mental health issues. They argue for their instrument validity on the grounds that it 'pushes individuals into unemployment irrespective of their unobserved [mental health, note by the Author] characteristics'.

records for Poland and is representative[6] with respect to age, gender, education and region.

The sample used has been restricted to the years 2006–2018, because since 2005, the political system in Poland shows strong patterns of bipartisanship (for details see Table 3). This allows a clear identification of the polarised voters' sentiments. On the other hand, since 2020, the COVID-19 pandemic has had a tremendous impact on government popularity surveys. Such a confounding factor could adversely affect the reliability of statistical inference.

The ESS dataset has been expanded to include yearly macroeconomic variables, i.e. GDP growth rate, unemployment rate and CPI inflation. The macroeconomic data were sourced from the World Bank Economic Outlook.

Due to the apparently high rotation, the ESS does not effectively constitute a panel. Approximately a third of the respondents appear in the dataset only once. Furthermore, rotation significantly increases in the final rounds (corresponding to years 2014–2018, see Figure 3).

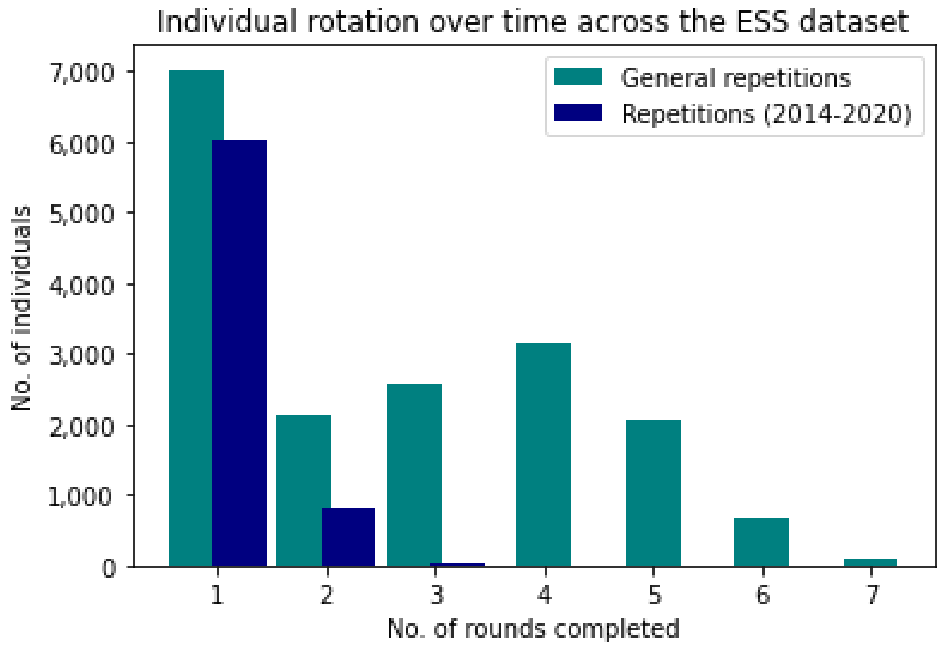
Table 3. Governing parties in Poland (2005–2019)

Parliamentary term	Years	Govt. leading party	Allies
V	2005–2007	Law and Justice (PiS)	Self-Defence (Samoobrona), League of Polish Families (LPR)
VI	2007–2011	Civic Platform (PO)	Polish People's Party (PSL)
VII	2011–2015	Civic Platform (PO)	Polish People's Party (PSL)
VIII	2015–2019	Law and Justice (PiS)	Agreement (Porozumienie), Solidarity Poland (SP)

Source: author's work.

[6] Representative, statistically robust results are obtained after weighting the data. Respective weights are provided by the ESS.

Figure 3. Rates of unemployment and inflation in Poland (2002–2019)



Source: World Bank.

4.2. Treatment group identification

The proper identification of the treatment group is the key empirical issue of this paper. As discussed in Sections 1 and 3, the minimal wage reform affected individuals only if two conditions were satisfied simultaneously: holding a Civil Code-based contract and earning less than PLN 12 per hour. However, neither of those conditions is reported directly in the ESS data, which necessitates an indirect, yet robust identification strategy.

The participants of the ESS are not asked about the legal form of their employment. However, the ESS Panel F provides sufficient information to develop an approximate non-Labour Code contract dummy. In particular, the respondents are asked the following questions:

1. **(F21)** In your main job are/were you:
   a) an employee
   b) self-employed
   c) working for your own family business
   d) refusal/don't know;
2. **(F22)** *(If self-employed in F21)* How many employees do you/did you have?;
3. **(F23)** *(If **not** self-employed in F21)* Do/did you have a work contract for:
   e) unlimited duration
   f) limited duration

g) no contract
h) refusal/don't know.

Based on the responses, Civil Code-based contract holders may be identified as those who either are self-employed with no employees or who are employees with a contract of limited duration[7] or no contract. Individuals with no contract are included due to the fact that, as pointed out by Chrostek et al. (2019), employers often offered their workers Civil Code contracts with low fixed salaries, while the majority of the actual wages was paid without any notice in the books. Thus, their income was directly increased by the reform, even if its main part did not come from a Civil Code-based contract.

Table 4. Deciles of monthly net and gross income distribution in Poland in 2016

Decile/Decile group	Net		Gross given by Statistics Poland
	ESS	HBS[a]	
	in PLN		
1	<= 1,400	1,375.88	1,890.32
2	1,401–1,900	1,840.00	2,279.31
3	1,901–2,400	2,320.00	2,673.49
4	2,401–2,900	2,812.69	3,070.00
5	2,901–3,400	3,301.50	3,510.67
6	3,401–4,000	3,873.90	4,000.00
7	4,001–4,700	4,521.50	4,583.33
8	4,701–5,600	5,414.00	5,444.88
9	5,601–7,100	6,880.00	7,200.00
10	>7,100		

a Household Budget Survey.

Source: author's work based on the ESS (2018), Household Budget Survey and Statistics Poland data (GUS, 2018a).

As mentioned in Section 3, holding a Civil Code contract itself is not enough to be classified into treatment. Identification must take into consideration not only the legal form of the held contract, but also the earnings. ESS Panel G provides the most accurate information on that (G9a, *What is your usual weekly/monthly/annual gross pay before tax and compulsory deductions?*); unfortunately, Panel G is a rotating panel included in rounds 2, 5 and 9 only. Hence, two alternative measures of income are proposed: a subjective poverty dummy and a decile poverty dummy.

1. **(Subjective poverty)** This dummy takes the value of 1 if the respondent finds it difficult or very difficult to cope on their present income. The dummy is coded on the basis of **(F42)**: *Which of the descriptions on this card comes closest to how you feel about your household's income nowadays?*

[7] Polish Labour Code generally does not allow an employment relation to last longer than 33 months based on limited duration contracts. A limited duration contract cannot be renewed more than twice. See Ustawa z dnia 26 czerwca 1974 r. – Kodeks pracy, art. 25.1.

a) *living comfortably on present income*
b) *coping on present income*
c) *finding it difficult to cope on present income*
d) *finding it very difficult to cope on present income*;

2. **(Decile poverty)** This dummy takes the value of 1 if the respondent falls into a given income decile group. The dummy is coded on the basis of **(F41)**: *Using this card, please tell me which letter describes your household's total income, after tax and compulsory deductions, from all sources?* The response card gives income decile ranges (see Table 4). The income brackets applied by the ESS have been sourced from the Household Budget Survey (HBS; Pol. Badanie Budżetów Gospodarstw Domowych, BBGD). Table 4 compares those net figures with the gross estimates provided by Statistics Poland.

4.3. Timing issues

The ESS dataset does not include information about the exact date of the respondent interview. The only data available are about the ESS round and equivalent year. The following question thus arises: should the year 2016 be treated as pre– or post–treatment? The reformed minimal salary regulation came into effect on 17th August 2016[8] (Ustawa z dnia 22 lipca 2016 r. o zmianie ustawy o minimalnym wynagrodzeniu za pracę oraz niektórych innych ustaw), which means that the last five months of 2016 were affected by the reform, while the first seven were not. Eventually, the decision was to interpret 2016 as post–reform, as the reform had been heavily discussed even before[9] its full implementation, thus affecting satisfaction with the government.

Finally, one must deal with the simultaneity problem posed by the fact that in 2016 – simultaneously with the minimal wage reform – the Family 500+ (Pol. Rodzina 500 Plus) transfer package was introduced. It offered a monthly transfer of PLN 500 per each but the oldest underage child in families having two or more children, and for the first (or only) child under certain income criteria (families could apply for the benefit for the first (or only) child if the net income per household member did not exceed 800 PLN) and it came into effect on 1st April 2016. Not accounting for that particular programme would lead to the overestimation of the impact of the minimal wage on government popularity due to the simultaneous timing.

ESS does not offer direct information about the amount of transfers each respondent receives from that programme. Thus, its effect is measured indirectly, as the number

[8] The official date of the reform implementation was 1st January 2017, but particular articles of the bill (increasing the minimal hourly wage) came into effect earlier, on 17th August 2016 (Ustawa z dnia 22 lipca 2016 r. o zmianie ustawy o minimalnym wynagrodzeniu...).

[9] The reform was promised in the 2015 expose of PM Beata Szydło, and the draft of the bill was passed to the Parliament as early as 6th June 2016.

of children entitled to the benefit. For the years 2016 and 2018 (ESS rounds 8 and 9), for every respondent, the number of children aged 18 or less has been calculated. The number of children entitled to the benefit is the total number of children minus 1 (unless the respondent reports being in the first decile of income distribution).

5. Regression analysis

5.1. Experiment design

The benchmark model used to estimate the extent to which the 2016 reform affected the satisfaction with the government is a DiD linear probability model (LPM) in the following form:

$$satgovt_{i,t} = \alpha + \beta_{treat}\mathrm{I}_{treat,i} + \beta_{post}\mathrm{I}_{post,t} + \beta_{DiD}\mathrm{I}_{treat,i}\mathrm{I}_{post,t} + \gamma P_t + \\ +\delta X_{i,t} + \varsigma E_t + \varepsilon_{i,t}.$$

The dependent variable, $satgovt_{i,t}$, is founded on the **B29** ESS question (*Now thinking about the [country] government, how satisfied are you with the way it is doing its job?*). The survey participants' answers were based on a 0–10 scale, where 0 meant 'extremely dissatisfied' and 1 meant 'extremely satisfied'. Thus, $satgovt_{i,t} = 1$ if the respondent selects 6 or higher.

Treatment indicator $\mathrm{I}_{treat,i}$ and post-reform indicator $\mathrm{I}_{post,t}$ are defined as shown in the previous section.

P_t represents political controls and responds to the bipartisan character of the Polish politics after 2005. It takes the value of 1 in the years when PiS was in power and 0 when PO was the governing party. Some sources in the literature suggest including not only the political, but also cultural factors (e.g. individual satisfaction with public services such as healthcare or education) in the models of satisfaction with the government (see Christensen & Lægreid, 2005). This approach comes from the distinction between the concepts of a specific and diffuse political support by Easton (1965). He argues that voters distinguish between general support for the public institutions and specific support rooting from their individual interactions and experiences with public services (see Christensen et al., 2020). However, Christensen and Lægreid (2005) show that those concepts are difficult to disentangle, as individuals who generally trust the government tend to also be satisfied with particular services provided by the authorities. Arguably, this effect could be further strengthened by the growing political polarisation in Poland. Consequently, variables measuring individual satisfaction with public services have been excluded from the list of the potential political controls.

$X_{i,t}$ represents the vector of individual controls which should guarantee conditional exogeneity of the selection into the treatment group. It was assembled in line with the literature, especially with the assertions of Liberini et al. (2017), who also use the DiD method. It includes age (linear and squared), gender, highest education attainment indicator and unemployment dummies (differentiated for those actively looking for a job and those not). Unlike in Liberini et al., however, our vector of individual controls does not include log of income due to the collinearity with the treatment effect.

Finally, E_t is an array of macroeconomic variables: aggregate output growth, unemployment rate and CPI inflation. All the explanatory variables are summarised in Table 5. Altogether, the variables used in this study represent a notion – well-rooted in the economic voting literature – that individual satisfaction with the government depends both on the soundness of the general economy and on individual welfare measures. Furthermore, variable selection stems from the expectation that non-economic factors, either socio-demographic or cultural, constitute an insightful complement to the purely economic ones.

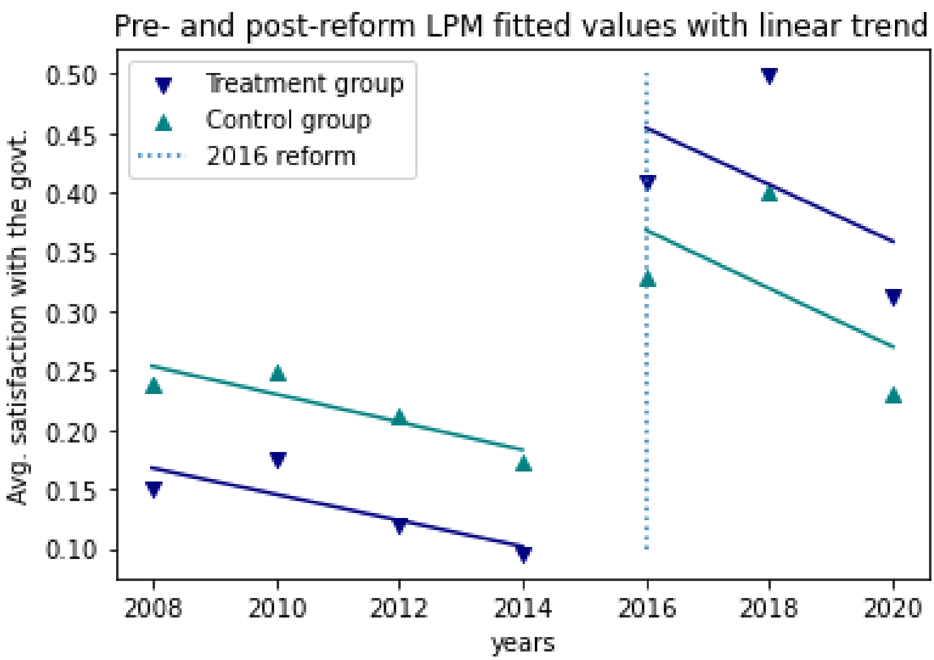
Table 5. Explanatory variables used in the study

Type	Variable	Description
Diff-in-diff operator	$\mathrm{I}_{treat,i}$	Treatment group indicator (=1 if treated)
Diff-in-diff operator	$\mathrm{I}_{post,t}$	Time indicator (=1 in 2016 or later)
Political controls	P_t	Governing party indicator (=1 if the PiS party is in power)
Individual socio-economic controls	$i.childbenef$	Number of children in the household who are entitled to the Family 500+ benefit
Individual socio-economic controls	$i.edulvla$	Highest educational attainment
Individual socio-economic controls	$uempla$	Unemployment indicator (=1 if unemployed)
Individual socio-economic controls	$uempli$	Economic activity indicator (=1 if unemployed and not actively looking for a job)
Individual socio-economic controls	$age, agesq$	Age (linear and squared)
Individual socio-economic controls	$gndr$	Gender indicator
Macroeconomic controls	$gdpgwth$	Real GDP *per capita* annual growth rate
Macroeconomic controls	$unempPL$	Unemployment rate
Macroeconomic controls	cpi	CPI inflation annual rate

Source: author's work.

The impact of the reform is studied with DiD, which is a standard policy assessment method in applied econometrics (see Angrist & Pischke, 2009). The core assumption underlying DiD applicability is the common trends assumption. DiD requires the control group to be a good counterfactual analogue of the treatment group, i.e. both groups – once controlled for socio-economic confounders – must follow parallel trends. Some authors, such as Kahn-Lang and Lang (2018) argue that DiD not only requires parallel trends between control and treatment groups, but also comparable levels. Nevertheless, both the assumptions are satisfied as presented in Figure 4, which clearly shows that both studied groups exhibit similar levels of satisfaction with the government and that their respective satisfaction metrics follow parallel trends. Arguably, the control group may be considered an accurate, counterfactual analogue of the treatment group.

Figure 4. Pre- and post-reform fitted values from LPM with controls



Note. The vertical axis represents the predicted mean probability of being satisfied with the government.
Source: author's work.

Under the conditional independence assumption, the LPM delivers unbiased, consistent estimates of the regressors' marginal effects. Linear probability regressions are estimated with the weighted least squares method with ESS-provided post-stratification weights (ESS, 2014).

5.2. Robustness checks

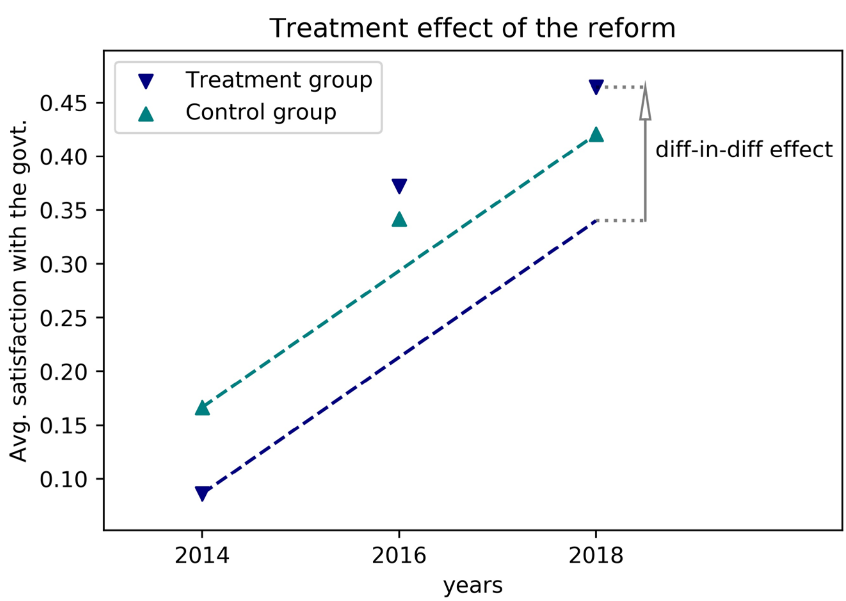
As a means of robustness checks, after the baseline DiD is estimated, controls and time-fixed effects are added to the benchmark model specification. In addition, binomial and multinomial discrete choice logistic models are fitted, and pooled and fixed effects panel models are introduced. Finally, alternative treatment identification techniques (subjective poverty indicator instead of decile poverty) are employed.

5.3. Results

Selected results are presented in Table 6. Statistical significance is assessed with respect to robust standard errors.

The DiD assessment of the reform's impact on individual, subjective satisfaction with the government is positive and statistically significant (see also Figures 4 and 5). As shown in Figure 4, in the pre-reform years, individuals in the treatment group had been on average less likely to be satisfied with the government; from 2016, however, the situation changed, and consequently those treated displayed a much higher probability of being satisfied with the government. This remains true for both the linear probability and the logit model. Only in the panel regression with individual fixed effects and year dummies, the reform's effect – despite the highest point estimate of all models – loses statistical significance; nevertheless, the *p*-value of almost 0.2 does not allow this impact to be altogether ignored. Furthermore, panel regression credibility might be questioned due to high in-sample rotation over time.

Figure 5. Graphical assessment of the treatment effect



Note. The vertical axis represents the predicted mean probability of being satisfied with the government.
Source: author's work.

The LPM, which gives unbiased and consistent estimates of the marginal effects *ceteris paribus*, concludes that the reform is related to an almost 15% increase in the probability that a given individual will be satisfied with the government. Thus, the effect of the reform on the popularity of the government is stronger than of any macro-economic variable considered. The only factors that seem to have comparable effects to the reform is the Family 500+ benefit in the case of three or more children. Significant errors related to some of the high-order Family 500+ coefficients might be related to the low number of the respective observations: there are only 90 recipients of the benefit for three children, 20 for four children and only five for five or more children.

Table 6. Estimation results (the *p*-values are reported in brackets)

Method dep. var	WLS (coef. reported)			WLogit (OR reported)			FE (coef. reported)
	$satgovt_{i,t}$						
I_{treat}	−0.0743 (0.002)	−0.088 (0)	−0.088 (0)	0.581 (0.010)	0.518 (0.002)	0.518 (0.002)	−0.060 (0.038)
I_{post}	0.120 (0)	0.643 (0)	0.167 (0)	1.871 (0)	44.691 (0)	2.628 (0)	0.266 (0)
I_{DiD}	0.142 (.043)	0.147 (0.031)	0.148 (0.031)	2.308 (0.016)	2.52 (0.002)	2.52 (0.008)	0.171 (0.199)
500+ (1)		−0.059 (0.053)	−0.059 (0.053)		0.763 (0.100)	0.764 (0.100)	
500+ (2)		0.024 (0.547)	0.024 (0.547)		1.161 (0.417)	1.161 (0.417)	
500+ (3)		0.143 (0.073)	0.143 (0.073)		1.915 (0.047)	1.916 (0.047)	
500+ (4)		0.294 (0.192)	0.294 (0.192)		3.546 (0.178)	3.546 (0.178)	
500+ (5)		0.699 (0)	0.699 (0)				
U_{act}		−0.056 (0.001)	−0.056 (0.001)		0.65 (0.002)	0.65 (0.002)	
U_{inact}		0.028 (0.321)	0.028 (0.321)		1.18 (0.316)	1.18 (0.316)	
GDP_t		0.049 (0)			1.34 (0)		
U_t		0.036 (0)			1.24 (0)		
CPI_t		0.021 (0)			1.145 (0)		
$X_{i,t}$		yes	yes		yes	yes	
T_t			yes			yes	yes
FE							yes
R^2	0.013	0.04	0.04	0.012	0.036	0.036	
obs. no	10,298	10,274	10,274	10,298	10,273	10,273	11,798

Source: author's work.

Interestingly, the variable responsible for the bipartisan system characteristics takes a negative value when PiS is in power (−36% with null *p*-value), which implies that the reform-induced support does not align with the general popularity of the governing party (in this case with the PiS party).

The positive coefficients on inflation and the general unemployment, however, are theoretically compelling. The positive impact of inflation on the satisfaction with the government could be explained with the CPI procyclical behaviour, which is a standard result in New Keynesian economics (for theory on the NK Phillips Curve, see Roberts (1995)). Justifying the positive coefficient on the general unemployment is, however, more difficult. Possibly, it can be explained by the relatively short sample in the model (seven periods only), which does not provide enough variability to correctly assess the impact of the unemployment.[10]

The logistic regression results confirm those of the LPM. The reported odds ratios show that there are approximately 2.5 times higher odds that an individual treated by the 2016 reform is satisfied with the government.

Finally, the panel regression with individual and time-fixed effects provides the most robust matching, as it is able to deal with all the unobserved, individual-level heterogeneity through fixed effects. The panel model also reports positive effects of the treatment on the satisfaction with the government. However, precise matching leads to higher error associated with the coefficient of interest, which might imply that the significance of the reform could have been overestimated in the previous methods.

6. Conclusions

The difference-in-differences analysis of the Polish 2016 minimal hourly wage reform effect on satisfaction with the government provides robust evidence that those treated are on average more satisfied with the government *ceteris paribus*. The effect of the reform-induced individual-level positive income shock is stronger than the impact of any other macroeconomic variable. This result supports the findings of Ward (2019), who claims that individual subjective happiness measures determine a higher fraction of the share of government vote than the macroeconomic factors do.

In line with the study by Liberini et al. (2017), this paper provides evidence for treatment analysis usefulness in the individual-level economic voting research. Finally, this paper develops a detailed insight into the recent labour market policies adopted in Poland, along with precise methods of identifying the affected groups. The ESS-based identification methods discussed above may be employed in further research of the public policy in Poland. In particular, investigating specific political preferences instead of general satisfaction with the government could shed light on the intricacies of the electoral dynamics and motivations. Furthermore, taking into account the role of satisfaction with particular public services (such as healthcare or public education, both reported in the ESS) could resolve whether improvement in the state-provided services is a necessary condition for increasing satisfaction with the government or if individual welfare growth suffices.

[10] Since general unemployment is individually-invariant, in absence of the explicit time effects, it might capture some residual components not captured by the other variables.

The identification method proposed in this study is deeply rooted in the realm of the Polish economy and politics. Nevertheless, one should expect its findings on the patterns of economic voting to be readily generalisable to other democratic states due to the method used as shown in the vast literature describing natural experiments and the DiD approach used to draw general inference from particular cases (as in the famous study of the impact of minimum wage increases on unemployment by Card and Krueger (1994). Arguably, during dynamic macroeconomic growth, targeted reforms tackling wages of low-income workers cause them to be satisfied with the government even more than those earning high wages, who can participate more actively in the general economic growth.

Reference

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics. An Empiricist's Companion*. Princeton University Press. https://doi.org/10.2307/j.ctvcm4j72.

Arak, P., Lewandowski, P., & Żakowiecki, P. (2014). *Fikcja zatrudniania w Polsce – propozycje wyjścia z impasu*. Instytut Badań Strukturalnych. https://ibs.org.pl/wp-content/uploads/2014/05/IBS_Policy_Paper_01-2014_streszczenie.pdf.

Barro, R. J. (1973). The Control of Politicians: An Economic Model. *Public Choice*, *14*(1), 19–42. https://doi.org/10.1007/BF01718440.

van der Brug, W., van der Eijk, C., & Franklin, M. (2007). *The Economy and the Vote. Economic Conditions and Elections in Fifteen Countries*. Cambridge University Press. https://doi.org/10.1017/CBO9780511618857.

Campbell, J. E., & Garand, J. C. (Eds.). (2000). *Before the Vote. Forecasting American National Elections*. SAGE Publications.

Card, D., & Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review*, *84*(4), 772–793. https://davidcard.berkeley.edu/papers/njmin-aer.pdf.

Christensen, T., & Lægreid, P. (2005). Trust in Government: The Relative Importance of Service Satisfaction, Political Factors, and Demography. *Public Performance & Management Review*, *28*(4), 487–511. https://doi.org/10.1080/15309576.2005.11051848.

Christensen, T., Yamamoto, K., & Aoyagi, S. (2020). Trust in local government: Service satisfaction, culture, and demography. *Administration & Society*, *52*(8), 1268–1296. https://doi.org/10.1177/0095399719897392.

Chrostek, P., Klejdysz, J., Korniluk, D., & Skawiński, M. (2019). *Wybrane aspekty systemu podatkowo-składkowego na podstawie danych PIT i ZUS 2016*. Ministerstwo Finansów RP. https://www.gov.pl/web/finanse/wybrane-aspekty-systemu-podatkowo-skladkowego-na-podstawie-danych-pit-i-zus-2016.

Duch, R. M., & Stevenson, R. T. (2008). *The Economic Vote. How Political and Economic Institutions Condition Election Results*. Cambridge University Press. https://doi.org/10.1017/CBO9780511755934.

Easton, D. (1965). *The political system: An inquiry into the state of political science*. Knopf.

Erikson, R. S. (1989). Economic Conditions and The Presidential Vote. *The American Political Science Review*, *83*(2), 567–573. https://doi.org/10.2307/1962406.

European Social Survey. (2014). *Documentation of ESS Post-Stratification Weights*.

European Social Survey. (2018). *ESS8 Appendix A2. Income*. https://stessrelpubprodwe.blob.core.windows.net/data/round8/survey/ESS8_appendix_a2_e02_1.pdf.

Farré, L., Fasani, F., & Mueller, H. (2015). *Feeling Useless: The Effect of Unemployment on Mental Health in the Great Recession* (Barcelona GSE Working Paper No. 838). https://bse.eu/research/working-papers/feeling-useless-effect-unemployment-mental-health-great-recession.

Fearon, J. D. (1999). Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In A. Przeworski, S. C. Stokes & B. Manin (Eds.), *Democracy, Accountability, and Representation* (pp. 55–97). Cambridge University Press. https://doi.org/10.1017/CBO9781139175104.

Feltovich, N., & Giovannoni, F. (2015). Selection vs. accountability: An experimental investigation of campaign promises in a moral-hazard environment. *Journal of Public Economics*, *126*, 39–51. https://doi.org/10.1016/j.jpubeco.2015.03.008.

Fiorina, M. P. (1978). Economic Retrospective Voting in American National Elections: A Micro-Analysis. *American Journal of Political Science*, *22*(2), 426–443. https://doi.org/10.2307/2110623.

Główny Urząd Statystyczny. (2018a). *Rocznik Statystyczny Pracy 2017*. https://stat.gov.pl/obszary-tematyczne/roczniki-statystyczne/roczniki-statystyczne/rocznik-statystyczny-pracy-2017,7,5.html.

Główny Urząd Statystyczny. (2018b). *Struktura wynagrodzeń według zawodów w październiku 2016 r.*

Hellwig, T., & McAllister, I. (2019). Party Positions, Asset Ownership, and Economic Voting. *Political Studies*, *67*(4), 912–931. https://doi.org/10.1177/0032321718815781.

Kahn-Lang, A., & Lang, K. (2018). The Promise and Pitfalls of Differences-in-Differences: Reflections on *16 and Pregnant* and Other Applications. *Journal of Business & Economic Statistics*, *38*(3), 613–520. https://doi.org/10.1080/07350015.2018.1546591.

Kamińska, A., Lewandowski, P., & Pogorzelski, K. (2014). *Dualny rynek pracy, płaca minimalna i nierówności* (IBS Policy Paper No. 02). https://ibs.org.pl/wp-content/uploads/2022/12/IBS_Policy_Paper_02_2014_pl.pdf.

Kenski, H. C. (1977). The Impact of Economic Conditions on Presidential Popularity. *The Journal of Politics*, *39*(3), 764–773. https://doi.org/10.2307/2129651.

Lewis-Beck, M. S., & Stegmaier, M. (2000). Economic Determinants of Electoral Outcomes. *Annual Review of Political Science*, *3*, 183–219. https://doi.org/10.1146/annurev.polisci.3.1.183.

Liberini, F., Redoano, M., & Proto, E. (2017). Happy Voters. *Journal of Public Economics*, *146*, 41–57. https://doi.org/10.1016/j.jpubeco.2016.11.013.

List, J. A., & Sturm, D. M. (2006). How Elections Matter: Theory and Evidence from Environmental Policy. *The Quarterly Journal of Economics*, *121*(4), 1249–1281. https://doi.org/10.1093/qje/121.4.1249.

Monroe, K. R. (1978). Economic Influences on Presidential Popularity. *Public Opinion Quarterly*, *42*(3), 360–369. https://doi.org/10.1086/268459.

Mueller, J. E. (1970). Presidential Popularity from Truman to Johnson. *American Political Science Review*, *64*(1), 18–34. https://doi.org/10.2307/1955610.

Nannestad, P., & Paldam, M. (1994). The VP-function: A survey of the literature on vote and popularity functions after 25 years. *Public Choice*, *79*(3–4), 213–245. https://doi.org/10.1007/BF01047771.

Pacek, A. C., & Radcliff, B. (1995). Economic Voting and the Welfare State: A Cross-National Analysis. *The Journal of Politics*, *57*(1), 44–61. https://doi.org/10.2307/2960270.

Palmer, H. D., & Whitten, G. D. (2003). Economics, politics, and the cost of ruling in advanced industrial democracies: How much does context matter?. In H. Dorussen & M. Taylor (Eds.), *Economic Voting* (pp. 92–117). Routlege.

Park, B. B., & Shin, J. (2017). Do the welfare benefits weaken the economic vote? A cross-national analysis of the welfare state and economic voting. *International Political Science Review*, *40*(1), 108–125. https://doi.org/10.1177/0192512117716169.

Powell, G. B. Jr., & Whitten, G. D. (1993). A Cross-National Analysis of Economic Voting: Taking Account of the Political Context. *American Journal of Political Science*, *37*(2), 391–414. https://doi.org/10.2307/2111378.

Roberts, J. M. (1995). New Keynesian Economics and the Phillips Curve. *Journal of Money, Credit and Banking*, *24*(4), 975–984. https://doi.org/10.2307/2077783.

Sejm Rzeczypospolitej Polskiej. (2016). *Projekt ustawy o zmianie ustawy o minimalnym wynagrodzeniu za pracę oraz niektórych innych ustaw*. https://orka.sejm.gov.pl/Druki8ka.nsf/0/90371E5981A7C5FEC1257FCD0033878F/%24File/600.pdf.

Stegmaier, M., Lewis-Beck, M. S., & Park, B. (2017). The VP-Function: A Review. In K. Arzheimer, J. Evans & M. S. Lewis-Beck (Eds.), *The SAGE Handbook of Electoral Behaviour* (pp. 584–605). SAGE Publications.

Tufte, E. R. (1978). *Political control of the economy*. Princeton University Press. https://doi.org/10.2307/j.ctv15r5816.

Ustawa z dnia 26 czerwca 1974 r. – Kodeks pracy (Dz.U. 1974 nr 24 poz. 141).

Ustawa z dnia 11 lutego 2016 r. o pomocy państwa w wychowywaniu dzieci (Dz.U. 2016 poz. 195).

Ustawa z dnia 22 lipca 2016 r. o zmianie ustawy o minimalnym wynagrodzeniu za pracę oraz niektórych innych ustaw (Dz.U. 2016 poz. 1265).

Ward, G. (2019). Happiness and Voting: Evidence from Four Decades of Elections in Europe. *American Journal of Political Science*, *64*(3), 504–518. https://doi.org/10.1111/ajps.12492.

Przegląd Statystyczny. Statistical Review, 2024, vol. 71, 4, 37–42 https://doi.org/10.59139/ps.2024.04.3

Topicality of the scientific achievements of Jerzy Spława-Neyman (1894–1981)[1]

Mirosław K. Krzyśko[a]

The year 2024 marks the 130th anniversary of Jerzy Spława-Neyman's birth. This is an opportunity to recall the life and achievements of this outstanding Polish statistician, one of the founders of modern mathematical statistics. His concepts and the procedures he initiated relating to interval estimation, statistical hypothesis testing, sampling methods and planning experiments have been used in research work in virtually every scientific discipline until today.

In the early 1930s, Spława-Neyman spotted a new form of the estimation problem, i.e. the problem of the estimation by 'an interval' or, more generally, 'by a set'. The inspiration to define confidence intervals came from a 1932 paper on the precision of the estimation of the regression coefficient by Wacław Pytkowski (Pytkowski, 1932), an assistant in the Department of Mathematical Statistics (headed by Spława-Neyman) at the Warsaw University of Life Sciences. The final version of the theory of confidence intervals was published in Spława-Neyman's two papers (1935, 1937a). The latter work was presented at a meeting of the Royal Statistical Society by Harold Jeffreys.

Interval estimation functions today as one of the standards in statistics. Although Spława-Neyman was not the only creator of this method, it is to him that we owe its modern form expressed in terms of frequency probability. The scholar's results in the field of confidence intervals have been continued – the examples include Efron's (1987) bootstrap method of confidence interval construction, Newcombe's (1998) confidence interval for the difference of two fractions, and, on the Polish ground, the results of the late Professor Ryszard Zieliński (2009, 2011) from the Mathematical Institute of the Polish Academy of Sciences and his son Professor Wojciech Zielinski (Jędrzejczak et al., 2021; Zieliński, 2017, 2022) from the Warsaw University of Life Sciences.

The second pillar of statistical inference, invented by Spława-Neyman with Egon Pearson, is the theory of statistical hypothesis testing. In 1928, they published their first joint paper, focused mainly on the test of the quotient of confidence. It appeared in two parts in the *Biometrika* journal (Neyman & Pearson, 1928a, 1928b). As a result

[1] The article is based on the paper delivered at the 42nd Conference Multivariate Statistical Analysis, held on 4th–6th November 2024 in Łódź.

[a] University of Kalisz, Inter-Faculty Department of Mathematics and Statistics, pl. Wojciecha Bogusławskiego 2, 62–800 Kalisz, Poland, e-mail: mkrzysko@amu.edu.pl, ORCID: https://orcid.org/0000-0001-8075-4432.

of further discussions with Pearson, Spława-Neyman formulated the testing problem in the language of optimisation, and in 1930 proved the basic lemma, today referred to as the Neyman-Pearson lemma. This is what Spława-Neyman said about the circumstances in which this discovery was made in his biographical note (Klonecki, 1970, pp. 21–23):

'I am able to point to a specific moment in which I understood how to formulate non-dogmatically the problem of the strongest test of a simple statistical hypothesis given a fixed alternative. Today this problem seems completely trivial, within the capacity of a novice doctoral student, but then, I must confess, it took Egon Pearson and me several years of work. The solution to this particular problem came to me one evening as I sat alone in the Statistical Laboratory of the University of Life Sciences in Warsaw, thinking persistently about something that could have been solved much earlier. The building was locked. Around 8 p.m. I heard voices calling out for me. It was my wife with some friends, who announced that it was time to go to the cinema. At first I felt irritated. Getting up from my desk to answer the call, I suddenly understood that for a given critical area and a given alternative hypothesis, the probability of making an error of the second kind could be calculated – it is represented by a specific integral. If so, the optimal critical area would be one that minimises this integral while at the same time the boundary condition would be related to the probability of the first-kind error. So we are dealing here with a specific problem of a calculus of variations, probably a very simple one. These thoughts came to me in a split second before I looked out the window to reassure my wife. They are still very clear in my memory, but I don't quite remember what film we were watching that evening. Probably one with Buster Keaton'.

The results obtained by Spława-Neyman were incorporated into the paper he and Pearson wrote about uniformly strongest and uniformly best tests in a class of similar tests, presented in November 1932 by Pearson at a meeting of the Royal Statistical Society. After being accepted by the Society, the article was published in the *Philosophical Transactions of the Royal Society* in 1933 (Neyman & Pearson, 1933). This work has been of fundamental importance for the theory of hypothesis testing with a fixed sample size. Spława-Neyman and Pearson provided a template for mathematical statistics in general and for a general decision theory, later developed by Abraham Wald (1950). In 1992, the above-mentioned work was selected for the first volume (Volume I) of a series describing the most important achievements in the foundations of statistics in the 20th century, published by the German publishing house Springer.

The theories of statistical hypothesis testing have been widely accepted and are being further developed. On the Polish ground, examples are the smooth and adaptive tests developed by Professor Teresa Ledwina of the University of Wroclaw and her students, and by Professor Tadeusz Inglot of the Wroclaw University of Technology.

The results of the aforementioned studies were presented by Inglot in his 2021 book entitled *Lectures on Hypothesis Testing Theory* (Inglot, 2021).

In 1937, Spława-Neyman published a paper in a Scandinavian actuarial journal in which he gave an asymptotically-optimal solution to the problem of testing the consistency of an observation distribution with an assigned, completely known, continuous distribution (Neyman, 1937b). He also introduced sequences of local distributions – 'contiguous' distributions, which later became a standard tool of asymptotic statistics. This work was another milestone in the development of statistics.

In 1933, Spława-Neyman's monographic study, *The outline of the theory and practice of surveying the structure of the population by means of the representative method* (Neyman, 1933), was published by the Institute of Social Affairs, with which the scholar collaborated. This monograph pioneered optimal sampling methods – prior to Spława-Neyman's work, no such idea was known. It has since stimulated new research into the sampling theory. On the basis of the aforementioned monograph, a landmark paper entitled *On two different aspects of the representative method: the stratified sampling method and the purposive selection method* was written. It was presented at the Royal Statistical Society on 19th June 1934, and published in the *Journal of the Royal Statistical Society* that same year (Neyman, 1934). The paper contributed to the development of modern scientific sampling. In 1992, it was ranked among the greatest 20th-century achievements in statistical methodology and published in Springer's Volume II. Spława-Neyman's legacy has been continued, among other studies, in the works of Professor Janusz Wywiał (2020, 2023) of the University of Economics in Katowice and in the research of the latter's students.

Two papers on agricultural experimentation written by Spława-Neyman at the State Scientific and Agricultural Institute in Bydgoszcz (Neyman, 1923a, 1923b) mark the author's another important scientific achievement. In 1924, on the basis of the former paper, Spława-Neyman received a PhD in mathematics from Warsaw University. The eminent mathematician Wacław Sierpiński was his promoter. Some part of his doctoral thesis was translated into English in 1990 and published with comprehensive comments in the *Statistical Science* (Splawa-Neyman, 1990). Thanks to the latter paper (Neyman, 1923b), the scholar received his habilitation in 1928 from the University of Life Sciences in Warsaw.

In 1935, at a meeting of the Industrial and Agricultural Section of the Royal Statistical Society, Spława-Neyman presented a paper on orthogonal systems and randomised blocks, which he wrote together with his students Karolina Iwaszkiewicz and Stanislaw Kołodziejczyk (Neyman et al., 1935). The 70-page-long paper was printed in 1935 in the supplement to the *Journal of the Royal Statistical Society*. This publication, along with the work of Ronald Fisher, was important in the development of experience planning.

The research directions initiated by Spława-Neyman have been extensively developed in Poland. A good example here is the study carried out by the team centered around Professor Tadeusz Caliński (Caliński, 2007; Caliński et al., 2017; Caliński & Lejeune, 1998) of the University of Life Sciences in Poznań. Spława-Neyman's legacy has also been continued in the work of Professors Tomasz Górecki and Łukasz Smaga (2017, 2019) from the Adam Mickiewicz University in Poznań, who investigated the methods of variance analysis for functional data, i.e. continuous functions in Hilbert space. The subject of functional data has been comprehensively explored at the Adam Mickiewicz University in Poznań for many years. It is one of the main directions of development of modern mathematical statistics. Another example is the study of linear models conducted by Professor Roman Zmyślony (Fonseca et al., 2018; Zmyślony & Kozioł, 2021) of the University of Zielona Góra and his students and collaborators. As regards foreign continuators of Spława-Neyman's legacy, a number of significant scholars from many scientific centers from around the world could be mentioned, including a distinguished statistician Professor C. R. Rao (Baksalary et al., 1995; Baksalary et al., 1992).

One more field of Spława-Neyman's scientific activity, relatively rarely recalled, is his development of the optimisation models for the American Army. The project, which started in February 1942, was carried out at the Berkely Statistical Laboratory with varying intensity until the end of the war.

Spława-Neyman saw almost unlimited applications of his methods and mathematical statistics in general in different areas of life and science, including medicine (the mechanism of the origin of cancer), astronomy (a model of the distribution of galaxies in interstellar matter), psychology (factor analysis), demography, or studies on the environmental pollution.

It is no exaggeration to say that Spława-Neyman was one of the most distinguished intellectuals of the 20th century. His scientific achievements paved the way for decades-long development of mathematical statistics, and are still relevant and creatively developed today.

References

Baksalary, J., Markiewicz, A., & Rao, C. R. (1995). Admissible linear estimation in the general Gauss-Markov model with respect to an arbitrary quadratic risk function. *Journal of Statistical Planning and Inference*, *44*(3), 341–347. https://doi.org/10.1016/0378-3758(94)00081-6.

Baksalary, J. K., Rao, C. R., & Markiewicz, A. (1992). A study of the influence of the 'natural restrictions' on estimation problems in the singular Gauss-Markov model. *Journal of Statistical Planning and Inference*, *31*(3), 335–351. https://doi.org/10.1016/0378-3758(92)90141-E.

Caliński, T. (2007). On some results of C. Radhakrishna Rao applicable to the analysis of multi-environment variety trials. *Journal of Statistical Theory and Practice*, *1*(3–4), 347–356. https://doi.org/10.1080/15598608.2007.10411845.

Caliński, T., Czajka, S., Kaczmarek, Z., Krajewski, P., Pilarczyk, W., Siatkowski, I., & Siatkowski, M. (2017). On a mixed model analysis of multi-environment variety trials: a reconsideration of the one-stage and two-stage models and analyses. *Statistical Papers*, *58*(2), 433–465. https://doi.org/10.1007/s00362-015-0706-y.

Caliński, T., & Lejeune, M. (1998). Dimensionality in Manova Tested by a Closed Testing Procedure. *Journal of Multivariate Analysis*, *65*(2), 181–194. https://doi.org/10.1006/jmva.1997.1722.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*(397), 171–185. https://doi.org/10.2307/2289144.

Fonseca, M., Kozioł, A., & Zmyślony, R. (2018). Testing hypotheses of covariance structure in multivariate data. *Electronic Journal of Linear Algebra*, *33*(1), 53–62. https://journals.uwyo.edu/index.php/ela/article/view/1823/1823.

Górecki, T., & Smaga, Ł. (2017). Multivariate analysis of variance for functional data. *Journal of Applied Statistics*, *44*(12), 2172–2189. https://doi.org/10.1080/02664763.2016.1247791.

Górecki, T., & Smaga, Ł. (2019). fdANOVA: an R software package for analysis of variance for univariate and multivariate functional data. *Computational Statistics*, *34*(2), 571–597. https://doi.org/10.1007/s00180-018-0842-7.

Inglot, T. (2021). *Wykłady z teorii testowania hipotez*. Oficyna Wydawnicza GiS. https://www.gis.wroc.pl/pdf/tth_www.pdf.

Jędrzejczak, A., Pekasiewicz, D., & Zieliński, W. (2021). Confidence interval for quantile ratio of the Dagum distribution. *REVSTAT. Statistical Journal*, *19*(1), 87–97. https://doi.org/10.57805/revstat.v19i1.333.

Klonecki, W. (1970). [unpublished manuscript]. Colloquium Metodologiczne z Agrobiometrii.

Newcombe, R. G. (1998). Interval estimation for the difference between independent proportions: Comparison of eleven methods. *Statistics in Medicine*, *17*(8), 873–890. https://doi.org/10.1002/(SICI)1097-0258(19980430)17:8%3C873::AID-SIM779%3E3.0.CO;2-I.

Neyman, J. (1923a). Próba uzasadnienia zastosowań rachunku prawdopodobieństwa do doświadczeń polowych. *Roczniki Nauk Rolniczych i Leśnych*, *10*, 1–51.

Neyman, J. (1923b). Próba rozwiązania niektórych zagadnień doświadczalnictwa rolniczego za pomocą statystyki matematycznej. *Miesięcznik Statystyczny*, *6*(9/12), 241–262.

Neyman, J. (1933). *Zarys teorji i praktyki badania struktury ludności metodą reprezentacyjną*. Instytut Spraw Społecznych.

Neyman, J. (1934). On the Two Different Aspects of the Representative Metod: the Metod of Stratified Sampling and the Metod of Purposive Selection. *Journal of the Royal Statistical Society*, *97*(4), 558–606. https://doi.org/10.2307/2342192.

Neyman, J. (1935). On the problem of confidence intervals. *Annals of Mathematical Statistics*, *6*(3), 111–116. https://doi.org/10.1214/aoms/1177732585.

Neyman, J. (1937a). Outline of a theory of statistical estimation based on the classical theory of probability. *Philosophical Transactions of the Royal Society A*, *236*, 333–380.

Neyman, J. (1937b). Smooth test for goodness of fit. *Skandinavisk Aktuarietidskrift*, *20*, 149–199.

Neyman, J., Iwaszkiewicz, K., & Kołodziejczyk, S. (1935). Statistical problems in agricultural experimentation (with discussion). *Supplement to the Journal of the Royal Statistical Society*. *Series B*, *2*(2), 107–180. https://doi.org/10.2307/2983637.

Neyman, J., & Pearson, E. S. (1928a). On the use and interpretation of certain test criteria for purpose of statistical inference. Part I. *Biometrika*, *20A*(1/2), 175–240. https://doi.org/10.1093/biomet/20A.1-2.175.

Neyman, J., & Pearson, E. S. (1928b). On the use and interpretation of certain test criteria for purpose of statistical inference. Part II. *Biometrika*, *20A*(3/4), 263–294. https://doi.org/10.1093/biomet/20A.3-4.263.

Neyman, J., & Pearson, E. S. (1933). On the Problem of the most Efficient Tests of Statistical Hypotheses. *Philosophical Transactions of the Royal Society A*, *231*, 289–337. https://doi.org/10.1098/rsta.1933.0009.

Pytkowski, W. (1932). *Zależność dochodów w małych gospodarstwach od ich powierzchni, nakładów i kapitału zainwestowanego w krowy*. Państwowy Instytut Naukowy Gospodarstwa Wiejskiego.

Splawa-Neyman, J. (1990). On the Application of Probability Theory to Agricultural Experiments. Essay on Principes. Section 9 (D. M. Dąbrowska & T. P. Speed, Trans.). *Statistical Science*, *5*(4), 465–472.

Wald, A. (1950). *Statistical Decision Functions*. Wiley.

Wywiał, J. L. (2020). Estimating the population mean using a continuous sampling design dependent on an auxiliary variable. *Statistics in Transition new series*, *21*(5), 1–16. https://doi.org/10.21307/stattrans-2020-052.

Wywiał, J. L. (2023). On the Maximum Likelihood Estimation of Population and Domain Means. *Journal of Statistical Theory and Practice*, *17*(40), 1–19. https://doi.org/10.1007/s42519-023-00337-4.

Zieliński, R. (2009). Przedział ufności dla frakcji. *Matematyka Stosowana*, *37*(10), 1–17. https://doi.org/10.14708/ma.v37i51/10.265.

Zieliński, R. (2011). *Statystyka matematyczna stosowana. Elementy*. Centrum Studiów Zaawansowanych Politechniki Warszawskiej.

Zieliński, W. (2017). The shortest Clopper-Pearson randomized confidence interval for binomial probability. *REVSTAT. Statistical Journal*, *15*(1), 141–153. https://doi.org/10.57805/revstat.v15i1.207.

Zieliński, W. (2022). The shortest confidence interval for Poisson mean. *Statistical Papers*, *63*(6), 2065–2072. https://doi.org/10.1007/s00362-022-01305-z.

Zmyślony, R., & Kozioł, A. (2021). Ratio F test for testing simultaneous hypotheses in models with blocked compound symmetric covariance structure. *Statistical Papers*, *62*(5), 2109–2118. https://doi.org/10.1007/s00362-020-01182-4.