

## **Using a Two-Step Clustering Approach to Examine Judiciary Efficiency in European Countries**

**Maria Stachova**

Matej Bel University, Faculty of Economics, Department of Quantitative Methods and Information Systems, Tajovského 10, 975 90, Banská Bystrica, Slovakia

E-mail: [maria.stachova@umb.sk](mailto:maria.stachova@umb.sk)

**Jan Hunady**

Matej Bel University, Faculty of Economics, Department Finance and Accounting, Tajovského 10, 975 90, Banská Bystrica, Slovakia

E-mail: [jan.hunady@umb.sk](mailto:jan.hunady@umb.sk)

### **Abstract**

*Panel data, also known as longitudinal data, is collected and analysed across various research areas. This type of data consists of statistical objects that are periodically observed over time. In comparison to cross-sectional data, there are significantly fewer clustering techniques available for panel data. Therefore, the main objective of the contribution is to present a two-step clustering approach. In the first step, the panel data are transformed into a static form using a set of proposed characteristics that capture time dynamics. In the second step, the objects are clustered using conventional spatial clustering algorithms, such as K-means clustering or hierarchical partitioning. The clustering performance of this approach is then compared to that of the well-known KML method using real panel data sets. These datasets include indicators that assess the effectiveness of courts at the first instance level. Factors like digitalisation in the public sector are affecting the judiciary's efficiency during this period. The methodology implemented allowed us to categorise European countries based on the efficiency of their courts while capturing the dynamic trends. This approach is generally helpful in assessing and comparing the efficiency of public spending and evaluating the quality of public institutions, including courts.*

**Keywords:** Panel data, Clustering, CluMP algorithm, KML algorithm, courts' effectiveness

**JEL classification:** C33, H59, H83

Received: 10 June 2025; Received in revised form: 14 July 2025; Accepted: 15 July 2025

### **1. Introduction**

As public budgets face increasing pressure, it has become crucial to assess performance and economic efficiency in the public sector. In various areas, such as education, healthcare, and infrastructure development, evaluating effectiveness has become a standard practice. However, significantly less attention has been given to this type of assessment in the field of justice. Despite the economic consequences of judicial efficiency, such as its impact on the growth potential of companies (Demirgüç-Kunt and Maksimovic, 1998), financial development (Beck & Levine, 2005), and stock markets (Lombardo and Pagano et al., 1999), this issue has often been overlooked in empirical studies.

This situation may stem from its specificity and the lack of data needed to identify and assess key indicators. However, in recent years, the situation has improved significantly. The Council of Europe's European Commission for the Efficiency of Justice (CEPEJ) has made significant progress in collecting internationally comparable data, which enables the comparison and assessment of courts in Europe. However, the potential of these data for quantitative studies was almost completely untapped until recently. The number of studies examining these data using quantitative techniques has only begun to grow in very recent years. Our study is using CEPEJ data to examine trends in the efficiency of courts at first instance in European countries. Our primary objective was to identify distinct groups of countries based on the efficiency of their judicial systems using the panel data set. Two-step clustering has been employed to achieve this aim, a novel technique developed by the authors. To the best of our knowledge, the proposed approach is the first to perform this type of clustering using these or similar datasets. The judiciary's efficiency is changing relatively dynamically over time and is influenced by several factors, among which digitisation and the use of emerging technologies appear to be crucial (Ahmed, 2021). Digital technologies and AI, in particular, can be utilised to enhance efficiency and expedite judicial processes (Emelianova, 2021). Hence, it is helpful to compare not only cross-sectional data but also to examine the dynamic changes that occur over time. The technique enables us to compare trajectories of development in efficiency, which is crucial for understanding the effects of policy changes and their impact, as well as cluster countries with similar results and development outcomes. The paper presents essential findings for establishing effective practices and providing recommendations for public policy in the justice sector.

Despite the growing interest of economists and statisticians in the judiciary in general, there is a lack of studies focused on the productivity or efficiency of courts in Europe. To the best of our knowledge, no other study has focused on clustering countries based on panel data or a similar dataset. However, machine learning techniques have already been applied to some extent in examining similar problems. A few studies have examined the typology of judicial decisions using statistical and machine-learning techniques, such as cluster analysis (Boyd et al., 2013) and decision trees (Kastellec, 2010). Giacalone et al. (2020) used K-means cluster analysis to identify clusters of Italian regions based on court efficiency and technology change. Vasconcelos et al. (2023) examine the productivity of the judiciary based on national data from Brazil. A summary of selected studies utilising machine learning tools in the field of judiciary is presented in Table 1.

Table 1 - Studies applying machine learning techniques in the field of judiciary

| Examined problem   | Authors/Study             | Methodology   | Data/Subjects   |
|--|---------------------------|---|---|
| Examine the productivity of the judiciary                              | Vasconcelos et al. (2023) | K-means clustering and neural networks (to identify the main variables within each cluster)             | State Courts of Justice in 27 Brazilian state (variables related to productivity) |
| Classify civil branches of a trial court in Iran based on performance  | Farzammehr, M. A. (2021). | PCA and hierarchical cluster analysis   | 18 civil branches of a trial court in Iran based on 20 indicators                 |
| Latent ideological traits allegedly affecting judicial voting behavior | Medina et al. (2022)      | Explorative cluster analysis used as applied to construct prior in the context of Item Response Theory. | Brazilian Supreme court data  |
| Cluster Italian regions  | Giacalone et al. (2020)   | K-means cluster   | Italian ordinary courts   |

|  |  |   |   |
|--|--|---|---|
| based on courts efficiency and technology change<br>Examine causes of action in complaints to illustrate the relationship of legal claims to one another<br>Judicial decision: relationships between case facts and case outcomes. | Boyd et al. (2013)   | analysis with application of PCA<br>Spectral cluster analysis   | data (Incoming cases, pending cases and resolved cases)<br>Newly compiled federal district court data in US |
|  | Kastellec (2010)   | Classification trees  | U.S. Supreme Court data   |
| Prediction the verdict of a court (predictive justice)   | Rani et al. (2023);<br>Alghazzawi,et al. (2022);<br>Sert et al. (2022);<br>Sil (2022);<br>Mumcuoğlu et al. (2021). | Combination of support vector machine and k-means cluster analysis;<br>Classification trees;<br>Random forests; neural networks; deep learning and other techniques | Different historic judicial datasets.   |

Multiple studies have utilised data envelopment analysis and the Malmquist Productivity Index to evaluate the relative efficiency of courts within a single jurisdiction (Giacalone, 2020; Falavigna, 2018). Despite the relatively widespread use of the DEA models in this field, the interpretation of the results concerning judicial efficiency is limited on an international scale. DEA serves as a more descriptive rather than an analytical tool. It shows the size of efficiency differences among courts but does not reveal the reasons for these differences (Voigt, 2016). Moreover, the inputs and outputs used in the models are determined by the supply and demand sides. The demand for court services is an important factor in determining judicial efficiency (Voigt, 2016). Therefore, while DEA models have some advantages, they also come with significant limitations compared to, for example, machine learning techniques. Other previous studies have applied machine learning to predict final verdicts. We have decided to analyse the performance of the courts as a multivariate problem, relying on a set of indicators. Moreover, our approach is based on panel data, which allows us to capture dynamic changes in the development of indicators.

Our research is one of a few studies examining judicial performance from a cross-country perspective, and there are at least three ways it contributes to existing literature:

- we introduce a novel two-step clustering algorithm (Clustering Micro-Panel trajectories), designed to handle panel data and capture time dynamics and trajectories.
- we critically compare the performance of our novel CluMP algorithm with the KML method.
- unlike previous studies that rely on cross-sectional data or time series for individual countries or regions, our research utilises panel data to analyse the efficiency of courts. The used methodology allows us to categorise European countries based on their court efficiency, considering dynamic trends, and offers important findings and potential recommendations for public policy.

In the next section, we describe the methodology and data, followed by the results and conclusions.

## 2. Methodology

In this section, we explain the novel methodology and data used in our paper. We employ our two-step algorithm for clustering micro-panel trajectories, CluMP (Sobíšek et al., 2018). In the first step, the panel data are transformed into static data with lower dimensions using a set of proposed dynamic characteristics that represent different features of the time course of the observed variables. In the second step, the elements are clustered by clustering techniques designed for static data.

Briefly, we can describe the algorithm as follows:

Initially, we identify a set of seven features that describe the dynamics of individual trajectories, providing a specific non-parametric description of the trend and variation in repeated measurements. The suggested characteristics include:

*Average triangular difference between the two consecutive measurement values*

For the  $i$ -th object, the average value is denoted  $\overline{diff}_i$ . It is calculated as the mean of all triangular differences between the two consecutive times points  $diff_{it}$  as follows:

$$\overline{diff}_i = \frac{\sum_{t=2}^T diff_{it}}{T-1} = \frac{\sum_{t=2}^T 1/2 \left( \frac{y_{it} - y_{i(t-1)}}{D_{it} - D_{i(t-1)}} \right)}{T-1}, \quad (1)$$

where  $y_{it}$  is the value of the monitored variable for the  $i$ -th object in time  $t$ ,  $D_{it}$  denotes the time difference of the time point  $t$  from the beginning ( $t = 1$ ) for the  $i$ -th object (e.g., in years) and  $T$  is the number of observations of the  $i$ -th object.

*Selective standard deviation of triangular differences between the two consecutive measurements*

For the  $i$ -th object, the variability characteristic, denoted  $sd(diff_i)$ , is calculated as

(2)

*Average absolute triangular difference between the two consecutive measurements*

For the  $i$ -th object, this absolute value, denoted  $|\overline{diff}_i|$ , is calculated as

$$|\overline{diff}_i| = \frac{\sum_{t=2}^T |diff_{it}|}{T-1}. \quad (3)$$

Compared to (1), absolute values of triangular differences between  $y_{it}$  and  $y_{i(t-1)}$  are averaged.

*Selective standard deviation of absolute triangular differences between the two consecutive measurements*

For the  $i$ -th object, the standard deviation, denoted  $sd(|diff_i|)$ , is calculated as

$$sd(|diff_i|) = \sqrt{\frac{\sum_{t=2}^T (|diff_{it}| - |\overline{diff}_i|)^2}{T-2}} \quad (4)$$

*Average growth coefficient  $\overline{k}_i$*

The mean growth rate indicates the average relative change of the monitored variable  $Y$  and is calculated for the  $i$ -th object as

$$\bar{k}_i = n_i^{-1} \sqrt{\frac{y_{iT}}{y_{i1}}} - 1 = n_i^{-1} \sqrt{k_{i2} k_{i3} \dots k_{iT}} - 1, \quad (5)$$

where  $k_{it} = y_{it}/y_{i(t-1)}$  for  $t = 2, \dots, T$  denotes the coefficient of growth between periods  $t$  and  $t - 1$  of the variable  $Y$ .

*The ratio of positive to negative changes (%pos)*

$$\% \text{ pos} = \frac{\text{number}_{k_{it} \geq 1}}{\text{number}_{k_{it} \leq 1}}, \quad t = 2, \dots, T, \quad (6)$$

if the number of negative changes (denominator) is zero, the value in the denominator is replaced with a value of 0.1.

*The value of maximum angle between the line connecting peripheral measurements and the one between the inner point and the first measurement (in radians)*

The maximum angle is given in radians as an angle between the line connecting peripheral measurements  $[t = 1, y_{i1}]$  and  $[t = T, y_{iT}]$  and the line connecting the inner point  $[t = t, y_{it}]$ , for which it holds that  $t > 1$  and  $t < T$ , to the first measurement  $[t = 1, y_{i1}]$ . For the  $i$ -th object, the calculation of the radian angle for the inner point  $t$ , where  $t > 1$  and  $t < T$ , can be done in the following way:

$$\angle T1t = \arccos \left( \frac{(D_{iT} - D_{i1}) \cdot (D_{it} - D_{i1}) + (y_{iT} - y_{i1}) \cdot (y_{it} - y_{i1})}{\sqrt{(D_{iT} - D_{i1})^2 + (y_{iT} - y_{i1})^2} \cdot \sqrt{(D_{it} - D_{i1})^2 + (y_{it} - y_{i1})^2}} \right), \quad (7)$$

$t = 2, \dots, T - 1$ . The maximum angle in radians is selected as the clustering variable from  $T - 2$  of calculated angles  $\angle T1t$ . The minus sign is assigned to the selected maximum angle if the slope of the tangent line passing through peripheral measurements is larger than that running through the selected inner  $t$  and first measurement. Otherwise, the maximum angle remains positive, i.e. the slope of the tangent line passing through peripheral measurements is smaller than that running through  $t$  and the first measurement. This clustering variable is referred to as **max  $\angle$**  hereinafter.

The features representing the objects (trajectories) extracted in the first step are clustered in the second step. Based on the simulation study (Sobišek, 2017), the Ward's hierarchical clustering method (Ward, 1963) applied to the Euclidean distance matrix (of seven extracted features) was chosen as the most appropriate, the effectiveness of this approach to micro-panel data clustering being confirmed by (Ferreira and Hitchcock, 2009).

The novel featured-based clustering method was compared with the clustering method based on standard approach. For this purpose, we selected the K-means algorithm applied in the "kml" package (Genolini et al., 2015) of statistical system R, that is modified for panel data. This algorithm is based on the original K-means clustering (MacQueen, 1967).

This method minimizes the utility function iteratively for the time  $t$ ,  $N$  objects according to an assumption of  $C$  clusters. The utility function can be expressed as follows:

$$\min \sum_{i=1}^N \sum_{c=1}^C u_{ict} d_{ict}^2, \quad (8)$$

where  $u_{ict}$  is a degree of appropriateness of the  $i$ -th object into the  $K$ -th cluster in the time  $t$  with conditions:

$$\sum_{c=1}^C u_{ict} = 1, \forall i, t, \quad (9)$$

$$\forall u_{ict} : u_{ict} = \begin{cases} 1 & \|\mathbf{x}_{it} - \mathbf{h}_{ct}\| = \arg \min_i \|\mathbf{x}_{it} - \mathbf{h}_{ct}\| \\ 0 & \text{elsewhere} \end{cases}$$

We used the Euclidean distance  $d_{ict} = \|\mathbf{x}_{it} - \mathbf{h}_{ct}\|$  between  $i$ -th vector of objects  $\mathbf{x}_{it} = (x_{it1}, \dots, x_{ijt}, \dots, x_{ijt})'$  and  $K$ -th centroid  $\mathbf{h}_{ct} = (h_{ct1}, \dots, h_{ctj}, \dots, h_{ctj})'$  in the time  $t$ . We applied the algorithm to the standardized values of variables.

### 3. Data

The research is based on the cross-country dataset from the European Commission for the Efficiency of Justice (CEPEJ) database. The CEPEJ collects and provides data on a wide range of indicators related to the efficiency, quality, and fairness of the justice system in the member states of the Council of Europe. According to Voigt and El-Bialy (2016), the CEPEJ dataset reflects the situation of relatively homogenous countries. It offers a unique opportunity to identify the determinants of judicial performance on a cross-country basis. The primary objective of CEPEJ is to analyse and enhance the efficiency and functioning of Member States' judicial systems. The data from the CEPEJ database have been utilized in multiple studies to compare the performance of judicial systems in Europe (e.g., Kapopoulos and Rizos, 2024; Cappellina, 2020; Lepore, 2018; Ippoliti, 2015). We are particularly focused on two key indicators that capture the efficiency of judicial systems: disposition time and clearance rate. The clearance rate is the ratio of the number of resolved cases over the number of incoming cases within a year. It measures the performance of the judicial system in terms of its caseload. On the one hand, a clearance rate lower than 100% implies that courts in the country are creating a backlog, as they resolve fewer cases than they receive.

On the other hand, a clearance rate over 100% means that they can handle more cases than they received. The disposition time measures the time required to resolve a pending case, taking into account the pace of each country's court caseload. It is calculated as the total number of pending cases divided by the total number of cases resolved during the same period. The ratio is then further multiplied by 365 to obtain the total number of days. Both indicators are commonly applied to measure the changes in court efficiency (Castelliano, 2023). As argued by Sarantis (2017), these two indicators should be studied together to provide a comprehensive evaluation of court efficiency. Moreover, analysing their evolution over time allows for a better understanding of the efforts to maintain or improve their efficiency (Sarantis, 2017). Both indicators (disposition time and clearance rate) have been previously used as measures of judicial system efficiency in several cross-country, quantitative-based studies (Kapopoulos and Rizos, 2024; Tsintzos and Plakandaras, 2020; Magalhães and Garoupa, 2020). Furthermore, disposition time has also been used as a separate indicator of efficiency by Lepore et al. (2018) and Lepore et al. (2017). Although these data have been used previously for international comparisons, it is necessary to note that comparing data from courts in different countries remains challenging due to differences in geographical, economic, and legal conditions.

Table 2 - Description of variables used in the analysis

| Variable  | Description  | Data type   | Source  |
|---|--|---|---|
| Disposition time - criminal cases (first instance)      | Disposition time is calculated as the number of pending cases (separately criminal and other than criminal cases) at the end of a year divided by the number of resolved cases within that year, multiplied by 365 (days in a year) at the courts of the first instance. | Panel data:<br>Cross-sections: 32 European countries<br>Periods: biannually (every two years in the period 2010-2020: 2010, 2012, 2014, 2016, 2018, 2020) | Council of Europe European Commission for the efficiency of justice (CEPEJ): Dynamic database of European judicial systems. |
| Disposition time - other than criminal (first instance) | The clearance rate is calculated as the ratio of the number of resolved cases (separately criminal and other than criminal cases) over the number of incoming cases at courts of the first instance  |   |   |
| Clearance rate - criminal cases (first instance)        |  |   |   |
| Clearance rate - other than criminal (first instance)   |  |   |   |

Due to the limited data availability and the need for more accurate international comparability, we only used data for the courts of first instance. Cases are classified as criminal and non-criminal types. Both types of cases have their own specific procedures; therefore, it is essential to consider them separately.

#### 4. Results

We will present the results obtained with CluMP portioning and compare them with those obtained by using the “kml” algorithm. Both methods were applied to values from 2010 to 2020 for each of the four variables separately. We were trying to find countries that are similar in the dynamic of variables describing the court's performance.

Initially, we focus on the clearance rate in both criminal and non-criminal cases. In the figure below, we can see the resulting segmentation of countries. The CluMP algorithm was able to identify finer patterns in the data and divide them into more evenly distributed groups with similar means and less variability within these groups. The KML algorithm divided the data in both cases into one larger cluster and two smaller ones, solely based on their median values, without considering their variability.

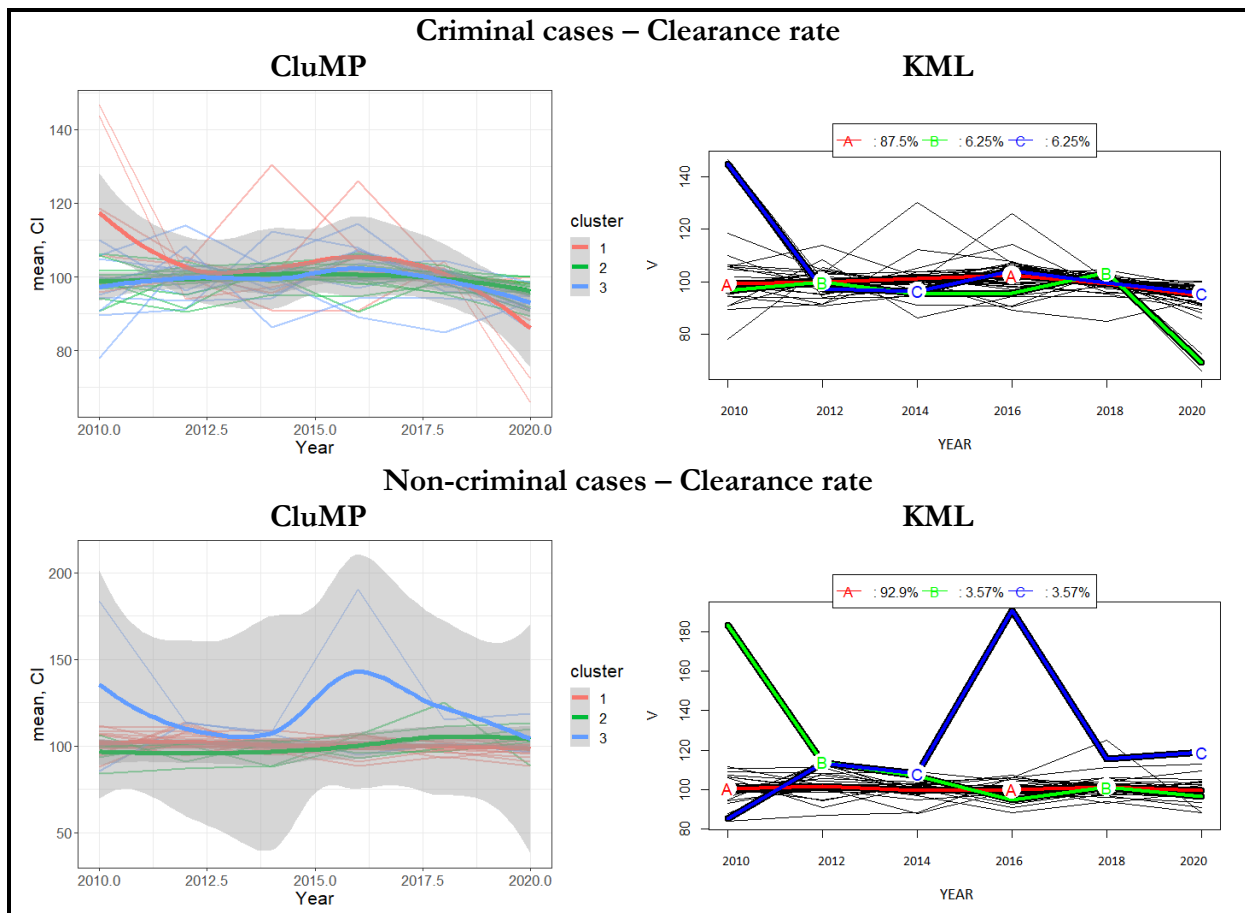


Figure 1- Resulting clusters for clearance rate: criminal cases and non-criminal cases

Table 3 presents descriptive statistics for all clusters created by both algorithms, including clearance rates for both criminal and non-criminal cases. To describe the results, we analyzed and compared the clusters based on their growth factors, means, and standard deviations (SD) across both algorithms. The CluMP algorithm produces clusters with more consistent and stable means and less variability, which suggests it is better at identifying more homogeneous groups. The KML generally results in clusters with higher means and greater variability, indicating it may capture a broader range of cases but with less consistency within the clusters.

Table 3 - Descriptive statistics of clusters for clearance rate of criminal and non-criminal cases

| Cluster | Criminal cases - Clearance Rate     |         |        |       |        |               |         |        |       |        |
|---------|-------------------------------------|---------|--------|-------|--------|---------------|---------|--------|-------|--------|
|         | CluMP                               |         |        |       |        | KML           |         |        |       |        |
|         | growth factor                       | Dim     | Mean   | SD    | Obs.   | growth factor | Dim     | Mean   | SD    | Obs.   |
| 1       | 0.94                                | overall | 102.19 | 15.51 |        | 0.99          | overall | 98.33  | 6.05  |        |
|         |                                     | between |        | 6.93  | n = 6  |               | between |        | 2.67  | n = 27 |
|         |                                     | within  |        | 14.12 |        |               | within  |        | 5.44  |        |
| 2       | 0.99                                | overall | 99.17  | 3.66  |        | 0.93          | overall | 106.91 | 16.19 |        |
|         |                                     | between |        | 2.02  | n = 15 |               | between |        | 1.068 | n = 3  |
|         |                                     | within  |        | 3.09  |        |               | within  |        | 16.16 |        |
| 3       | 0.99                                | overall | 98.54  | 7.00  |        | 1.00          | overall | 104.47 | 10.80 |        |
|         |                                     | between |        | 3.00  | n = 11 |               | between |        | 1.82  | n = 2  |
|         |                                     | within  |        | 6.38  |        |               | within  |        | 10.72 |        |
| Cluster | Non-Criminal cases - Clearance Rate |         |        |       |        |               |         |        |       |        |
|         | CluMP                               |         |        |       |        | KML           |         |        |       |        |
|         | growth factor                       | Dim     | Mean   | SD    | Obs.   | growth factor | Dim     | Mean   | SD    | Obs.   |
| 1       | 0.99                                | overall | 100.64 | 4.61  |        | 0.99          | overall | 100.53 | 5.62  |        |
|         |                                     | between |        | 2.55  | n = 21 |               | between |        | 2.60  | n = 26 |
|         |                                     | within  |        | 3.88  |        |               | within  |        | 5.00  |        |
| 2       | 1.01                                | overall | 100.07 | 8.76  |        | 0.88          | overall | 116.18 | 33.71 |        |
|         |                                     | between |        | 3.07  | n = 5  |               | between |        |       | n = 1  |
|         |                                     | within  |        | 8.30  |        |               | within  |        | 33.71 |        |
| 3       | 0.97                                | overall | 119.07 | 33.23 |        | 1.07          | overall | 121.95 | 35.69 |        |
|         |                                     | between |        | 4.08  | n = 2  |               | between |        |       | n = 1  |
|         |                                     | within  |        | 33.10 |        |               | within  |        | 35.69 |        |

As the second step, we analysed the clusters created by both algorithms for the disposition time of both criminal and non-criminal cases. The time trajectories of created clusters can be seen in Figure 2.

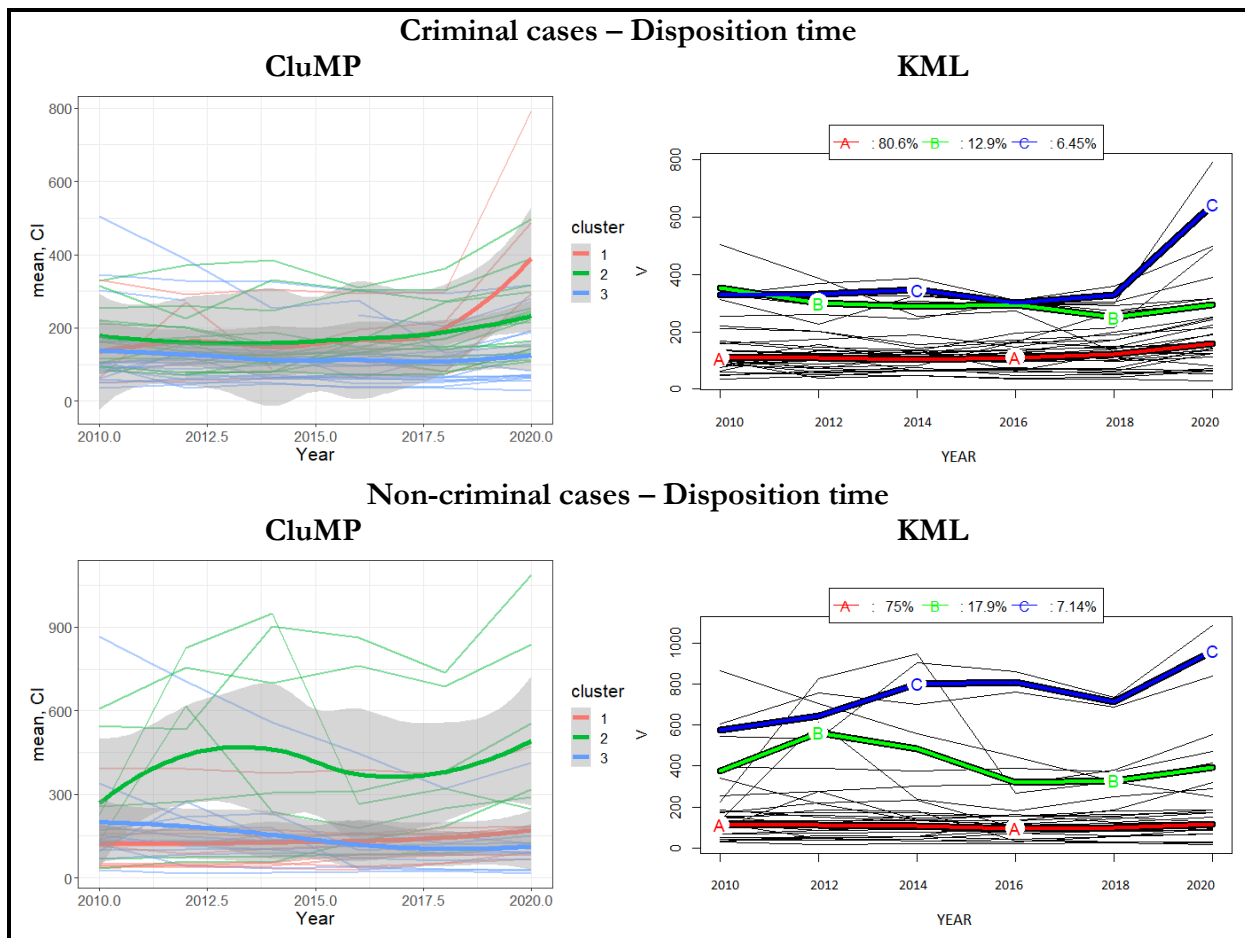


Figure 2 - Resulting clusters for duration time: criminal cases and non-criminal cases

When we examine the descriptive statistics presented in Table 4, we can see that the results show similar performance of the methods, with only minor differences, as observed for the clearance rate variable. The CluMP, again, tends to form clusters with more stable and consistent values of disposition times, capturing less variability within clusters but showing higher between-cluster variability in some cases. The KML often results in clusters with higher mean disposition times and greater within-cluster variability, suggesting it groups more extreme cases.

Table 4 - Descriptive statistics of clusters for disposition time of criminal and non-criminal cases

| Cluster | Criminal cases - Disposition Time     |         |        |        |        |               |         |        |        |        |
|---------|---------------------------------------|---------|--------|--------|--------|---------------|---------|--------|--------|--------|
|         | CluMP                                 |         |        |        |        | KML           |         |        |        |        |
|         | growth factor                         | Dim     | Mean   | SD     | Obs.   | growth factor | Dim     | Mean   | SD     | Obs.   |
| 1       | 1,41                                  | overall | 206.24 | 169.09 |        | 1.02          | overall | 84.18  | 32.19  |        |
|         |                                       | between |        | 114.11 | n = 5  |               | between |        | 25.62  | n = 14 |
|         |                                       | within  |        | 125.96 |        |               | within  |        | 20.48  |        |
| 2       | 1.08                                  | overall | 180.77 | 89.43  |        | 1.13          | overall | 162.27 | 62.45  |        |
|         |                                       | between |        | 82.73  | n = 15 |               | between |        | 26.06  | n = 11 |
|         |                                       | within  |        | 39.63  |        |               | within  |        | 57.21  |        |
| 3       | 1.00                                  | overall | 120.92 | 88.21  |        | 1.02          | overall | 325.11 | 107.41 |        |
|         |                                       | between |        | 81.51  | n = 18 |               | between |        | 45.86  | n = 6  |
|         |                                       | within  |        | 37.65  |        |               | within  |        | 98.66  |        |
| Cluster | Non-criminal cases – Disposition Time |         |        |        |        |               |         |        |        |        |
|         | CluMP                                 |         |        |        |        | KML           |         |        |        |        |
|         | growth factor                         | Dim     | Mean   | SD     | Obs.   | growth factor | Dim     | Mean   | SD     | Obs.   |
| 1       | 1.10                                  | overall | 136.80 | 106.51 |        | 1.01          | overall | 107.01 | 60.53  |        |
|         |                                       | between |        | 109.68 | n = 9  |               | between |        | 46.67  | n = 21 |
|         |                                       | within  |        | 21.23  |        |               | within  |        | 39.66  |        |
| 2       | 1.16                                  | overall | 403.12 | 301.02 |        | 1.04          | overall | 411.03 | 204.63 |        |
|         |                                       | between |        | 271.52 | n = 7  |               | between |        | 104.97 | n = 5  |
|         |                                       | within  |        | 160.88 |        |               | within  |        | 180.98 |        |
| 3       | 0.90                                  | overall | 146.03 | 147.56 |        | 1.1           | overall | 751.50 | 157.51 |        |
|         |                                       | between |        | 135.96 | n = 12 |               | between |        | 37.48  | n = 2  |
|         |                                       | within  |        | 67.75  |        |               | within  |        | 155.06 |        |

We compute the silhouette index to describe the performance of both algorithms for each clustering, as shown in Table 5.

Table 5 - Values of the Silhouette index for the CluMP and the KML clustering algorithms and each examined variable

|       | Silhouette index |                  |                           |                  |
|-------|------------------|------------------|---------------------------|------------------|
|       | Criminal cases   |                  | Other than criminal cases |                  |
|       | Clearance rate   | Disposition time | Clearance rate            | Disposition time |
| CluMP | <b>0.55</b>      | <b>0.6</b>       | <b>0.75</b>               | 0.53             |
| KML   | 0.36             | 0.39             | 0.74                      | <b>0.64</b>      |

The CluMP algorithm overperforms the KLM in three of four cases. This case can be attributed to the fact that the KML algorithm may be better at identifying clusters that are broadly defined but exhibit significant internal variability, potentially capturing outliers more effectively.

The results achieved by the CluMP algorithm are summarised in Tables 6 and 7. Due to the distinct nature of criminal and non-criminal cases, we categorise the results into two separate parts. Table 6 presents the clustering based on judicial efficiency related only to criminal cases, and Table 7 summarises the results for non-criminal cases. Our approach allows us to cluster countries not only based on the level of selected indicators but also according to their trajectories during the examined period, for example. Countries included in cluster 1.1 exhibit outstanding efficiency based on the clearance rate, but the trend during the examined period appears to be rather negative.

Table 6 - Criminal cases judiciary efficiency: panel clustering with CluMP algorithm

| Cluster no. | Clearance (the higher the better)  | Cluster characteristics  | Cluster no. | Duration (the lower the better)  | Cluster characteristics   |
|-------------|--|--|-------------|--|---|
| 1.1         | Armenia, Croatia, Estonia, Georgia, Malta, North Macedonia   | High clearance rate with declining trend (Top-performers with worsening results) | 2.1         | Albania, Andorra, Armenia, Azerbaijan, Malta   | High duration time with a growing trend (Under-performers with unfavorable trend) |
| 1.2         | Austria, Bulgaria, Denmark, Finland, Hungary, Lithuania, Netherlands, Norway, Poland, Republic of Moldova, Romania, Spain, Sweden, Switzerland, Slovakia | Average clearance rate with stable trend   | 2.2         | Austria, Croatia, Cyprus, Italy, Monaco, Montenegro, North Macedonia, Republic of Moldova, Romania, Slovenia, Spain, Sweden, Ukraine, Türkiye, UK  | Average duration time with rather stable trend                                    |
| 1.3         | Azerbaijan, Bosnia and Herzegovina, Cyprus, France, Italy, Latvia, Montenegro, Serbia, Slovenia, Ukraine, Türkiye  | Low clearance rate with stable trend   | 2.3         | Bosnia and Herzegovina, Bulgaria, Czechia, Denmark, Estonia, Finland, Georgia, Germany, Hungary, Latvia, Lithuania, Netherlands, Norway, Poland, Portugal, Serbia, Switzerland, Slovakia | Low duration time with stable trend (Top-performers)                              |

Overall best performing countries with respect to criminal cases are Estonia and Georgia. Good results in both set of indicators have been also achieved by Croatia, North Macedonia, Bulgaria, Croatia, Denmark, Hungary, Lithuania, Netherlands, Norway, Poland, Switzerland and Slovakia. The lack of panel data for some countries has led to these countries being assessed in only one of two areas (clearance or duration).

Table 7 – Non-criminal cases judiciary efficiency: panel clustering with CluMP algorithm

| Cluster no. | Clearance (the higher the better)   | Cluster characteristics   | Cluster no. | Duration (the lower the better)  | Cluster characteristics  |
|-------------|---|---|-------------|--|--|
| 3.1         | Azerbaijan, Bosnia and Herzegovina, Bulgaria, Croatia, Czechia, Estonia, Finland, France, Georgia, Hungary, Italy, Latvia, Lithuania, Malta, Netherlands, Norway, Rep. of Moldova, Slovenia, Switzerland, Ukraine | Average clearance rate with stable trend  | 4.1         | Azerbaijan, Czechia, Italy, Lithuania, Norway, Poland, Rep. of Moldova, Romania, Ukraine                                 | Low duration time with almost stable trend (Top-performers)                          |
| 3.2         | Cyprus, Montenegro, Poland, Sweden, Slovakia  | Low to average clearance rate with increasing trend (Improvers)                             | 4.2         | Bosnia and Herzegovina, Bulgaria, Cyprus, France, Georgia, Montenegro, Serbia  | Very long duration time with growing trend (Under-performers with unfavorable trend) |
| 3.3         | North Macedonia, Serbia   | High clearance rate with decreasing trend (Top-performers with gradually worsening results) | 4.3         | Croatia, Denmark, Estonia, Finland, Latvia, Malta, Netherlands, North Macedonia, Slovenia, Sweden, Switzerland, Slovakia | Average to low duration with decreasing trend (Fast-improvers)                       |

The results for non-criminal cases differ slightly. We can again distinguish between clusters where countries have good performance and trends, and those where efficiency is relatively weak. Azerbaijan, Czechia, Italy, North Macedonia, Norway, the Republic of Moldova and Ukraine are countries with the best overall results. Sweden and Slovakia are two countries improving significantly in both examined indicators. In the case of several countries, we had panel data available for only one type of indicator. Countries with unsatisfactory results and a negative trend in both measures are Cyprus and Montenegro.

Our findings contribute to the growing body of literature on judicial efficiency. While machine learning has been applied to various judicial problems, such as classifying civil branches (Farzammehr, 2021) or predicting verdicts (Rani et al., Alghazzawi et al., 2022; Sert et al., 2022; Mumcuoğlu et al., 2021), our work explicitly addresses the dynamic efficiency of judicial systems. Since no similar study based on panel clustering has been conducted to date, the comparability of our results is limited. However, some of the results can still be discussed in more detail. Our approach emphasises dynamic changes in selected efficient indicators over time, which is entirely consistent with Ahmed (2021), who argues that the judiciary's efficiency changes dynamically over time. The justification for this approach is also stated by Sarantis (2017), who highlights the importance of analysing the evolution of indicators such as disposition time and clearance rate to understand changes in court efficiency better. This finding is also consistent with the results of

other studies, which have shown that various factors dynamically influence judicial efficiency over time. The adoption of digital technologies and AI over time is seen as a crucial factor leading to changes in judicial process efficiency (Emelianova, 2021; Ahmed, 2021).

## 5. Conclusions

Our research provides new insights into the application of cluster analysis for assessing and classifying courts based on efficiency indicators. The efficacy issue is both economically and socially significant, yet it remains underexamined. However, recent research has produced some findings in this area, relying primarily on methods based on cross-sectional data and occasionally also on time series. In contrast, our research takes a further step ahead by employing panel data and introducing a novel two-step clustering algorithm (the CluMP). This approach enhances the accuracy of results and enables more straightforward interpretation. Additionally, CluMP can also be applied to cluster other entities based on the panel data indicators, presenting a rather extensive range of possibilities for its further application. We encourage other researchers to test it on different datasets, as we see ample opportunities for further research and improvements.

The results of CluMP were critically compared with the results obtained by the KML algorithm, which serves a similar purpose. Both methods achieve quite a high silhouette index. The CluMP was able to recognise patterns more precisely and divide our data into more evenly grouped groups. On the other hand, the KML algorithm may appear rougher, but it can bring focus to outliers in the objects. These differences may reflect the underlying methods of the algorithms, with CluMP possibly being more conservative in its clustering approach, while KML might capture more nuanced but variable patterns. Overall, the choice between the CluMP and the KML depends on the specific goals of the clustering task. If the objective is to form more distinct clusters with less internal variability, the CluMP may be a preferable option. If the goal is to capture broader trends, even if it leads to more internal cluster variability, the KML might be the better choice.

Our approach can be applied to a wide range of research problems in economic. In our case, we use it to cluster the countries based on their judicial efficiency. The results obtained by CluMP reveal similarities and differences in judicial efficiency and its trajectories among European countries. Based on the results, it is possible to identify examples of good practice as well as countries where the situation is critically deteriorating. These findings can be helpful in formulating public policies in the field of justice. However, our approach has several limitations. The availability of panel data for the large number of countries used in the sample is limited. Moreover, due to some significant differences in judicial systems and administration, the results must still be interpreted with caution. Thanks to the comparable and standardised data obtained using the international CEPEJ database, an inevitable comparison is possible. However, international comparison of court performance is a multidimensional and complex problem influenced by other factors. There is an expansive room for further research in this area. Since this is essentially a pioneering attempt, it did not aim to examine various aspects of judicial efficiency with its determinants. Hence, a more detailed examination based on the data used is desirable in the future.

## Acknowledgements

This work was supported by Slovak Scientific Grant Agency (VEGA) under the contract No. 1/0411/24. The authors also gratefully acknowledge financial support from the BRRIDGE

project, funded by the European Union's Horizon Europe research and innovation program under grant agreement No. 101079219.

## References

- Alghazzawi, D., Bamasag, O., Albeshri, A., Sana, I., Ullah, H., Asghar, M. Z. (2022). Efficient prediction of court judgments using an LSTM+CNN neural network model with an optimal feature set. *Mathematics*, 10(5), 683.
- Ahmed, R. K., Muhammed, K. H., Pappel, I., Draheim, D. (2021). Impact of e-court systems implementation: a case study. *Transforming Government: People, Process and Policy*, 15(1), 108-128.
- Beck, T., & Levine, R. (2005). Legal institutions and financial development. *Handbook of new institutional economics*, Boston, pp. 251-278, MA: Springer US.
- Boyd, C. L., Hoffman, D. A., Obradovic, Z., Ristovski, K. (2013). Building a Taxonomy of Litigation: Clusters of Causes of Action in Federal Complaints. *Journal of Empirical Legal Studies*, 10(2), 253-287.
- Cappellina, B. (2020). Legitimising EU Governance through Performance Assessment Instruments. European Indicators for a Judicial Administration Policy. *International Review of Public Policy*, 2(2), 141-158.
- Castelliano, C., Grajzl, P., Watanabe, E. (2023). Does electronic case-processing enhance court efficacy? New quantitative evidence. *Government Information Quarterly*, 40(4), 101861.
- CEPEJ. (2016). CEPEJ Guidelines. European Commission for the efficiency of justice (CEPEJ). <https://rm.coe.int/cepej-guidelines/16809f007a>. (accessed in July 9, 2025)
- Demirgüç-Kunt, A., Maksimovic, V. (2002). Funding growth in bank-based and market-based financial systems: evidence from firm-level data. *Journal of financial Economics*, 65(3), 337-363.
- Emelianova, Tatiana V. (2021). "Affect of Artificial Intelligence Artificial intelligence Technologies and Digitalisation on Jurisprudence and Education." *Artificial Intelligence in Industry 4.0: A Collection of Innovative Research Case-studies that are Reworking the Way We Look at Industry 4.0 Thanks to Artificial Intelligence*, pp. 165-179. Cham: Springer International Publishing.
- Falavigna, G., Ippoliti, R., Ramello, G. B. (2018). DEA-based Malmquist productivity indexes for understanding courts reform. *Socio-Economic Planning Sciences*, 62, 31-43.
- Farzammehr, M. A. (2021). Application of hierarchical clustering on principal components to evaluate the performance of justice system by judicial indicators. *Journal of Statistical Modelling: Theory and Applications*, 2(2), 143-158.
- Ferreira, L., Hitchcock, D.B. (2009). A comparison of hierarchical methods for clustering functional data. *Communications in Statistics - Simulation and Computation*, 38, 1925-1949.
- Genolini, Ch., Falissard, B. (2011). Kml: A package to cluster longitudinal data. *Computer methods and programs in biomedicine*, 104, 112-121.

Giacalone, M., Nissi, E., Cusatelli, C. (2020). Dynamic efficiency evaluation of Italian judicial system using DEA based Malmquist productivity indexes. *Socio-Economic Planning Sciences*, 72, 100952.

Ippoliti, R., Melcarne, A., Ramello, G. B. (2015). The impact of judicial efficiency on entrepreneurial action: A European perspective. *Economic Notes: Review of Banking, Finance and Monetary Economics*, 44(1), 57-74.

Kapopoulos, P., Rizos, A. (2024). Judicial efficiency and economic growth: Evidence based on European Union data. *Scottish Journal of Political Economy*, 71(1), 101-131.

Kastellec, J. P. (2010). The statistical analysis of judicial decisions and legal rules with classification trees. *Journal of Empirical Legal Studies*, 7(2), 202-230.

Lepore, L., Paolone, F., Cambrea, D. R. (2018). Ownership structure, investors' protection and corporate valuation: The effect of judicial system efficiency in family and non-family firms. *Journal of management and governance*, 22, 829-862.

Lepore, L., Paolone, F., Pisano, S., Alvino, F. (2017). A cross-country comparison of the relationship between ownership concentration and firm performance: does judicial system efficiency matter?. *Corporate Governance: The International Journal of Business in Society*, 17(2), 321-340.

Lombardo, D., Pagano, M. (1999). Law and equity markets: A simple model. *Available at SSRN 209312*.

MacQueen, J., B. (1967). Some Methods for classification and Analysis of Multivariate Observations. *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 281-297.

Mumcuoğlu, E., Öztürk, C. E., Ozaktas, H. M., Koç, A. (2021). Natural language processing in law: Prediction of outcomes in the higher courts of Turkey. *Information Processing & Management*, 58(5), 102684.

Magalhães, P. C., Garoupa, N. (2020). Judicial Performance and Trust in Legal Systems: Findings from a Decade of Surveys in over 20 European Countries. *Social Science Quarterly (Wiley-Blackwell)*, 101(5).

Medina, D., dalla Pellegrina, L., Garoupa, N. (2022). Unfolding judicial ideology: a data-generating priors approach with an application to the Brazilian Supreme court. *Review of Law & Economics*, 18(1), 1-54.

Rani, A. J. M., Bharathwaj, K. S., Swaroopan, N. J., Kumar, K. H., Geetha, R. (2023). A Legal Prediction Model Using Support Vector Machine and K-Means Clustering Algorithm for Predicting Judgements and Making Decisions. *International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)*, 1-6.

Sarantis, D. (2017). The challenge of accelerating Greek judicial procedure. *Electronic Government and the Information Systems Perspective: 6th International Conference, EGOVIS 2017, Lyon, France, August 28-31, 2017, Proceedings 6*, 251-260. Springer International Publishing.

Sert, M. F., Yıldırım, E., Haşlak, İ. (2022). Using artificial intelligence to predict decisions of the Turkish constitutional court. *Social Science Computer Review*, 40(6), 1416-1435.

Sil, R. (2022). Random Forest Based Legal Prediction System. *Proceedings of International Joint Conference on Advances in Computational Intelligence: IJCACI 2021*, 623-633. Singapore: Springer Nature Singapore.

Sobíšek, L. (2017). Clustering and regression analysis of micro panel data, Dissertation thesis, 2017. <https://vskp.vse.cz/eid/52823>. (accessed in July 9, 2025)

Sobisek, L., Stachova, M., Fojtik, J. (2018). Novel Feature-Based Clustering of Micro-Panel Data (CluMP). Working paper version online: [www.arxiv.org](http://www.arxiv.org). <https://doi.org/10.48550/arXiv.1807.05926>.

Tsintzos, P., Plakandaras, V. (2020). The judiciary system as a productivity factor; the European experience. *Economics Letters*, 192, 109257.

Vasconcelos, F.F., Sátiro, R.M., Fávero, L.P.L., Bortoloto, G.T. Corrêa, H.L. (2023) Analysis of Judiciary Expenditure and Productivity Using Machine Learning Techniques. *Mathematics*, 11, 3195.

Voigt, S. (2016). Determinants of judicial efficiency: a survey. *European Journal of Law and Economics*, 42, 183-208.

Voigt, S., El-Bialy, N. (2016). Identifying the determinants of aggregate judicial performance: taxpayers' money well spent?. *European Journal of Law and Economics*, 41, 283-319.

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function, *Journal of the American Statistical Association*, 58, 236-244.

## Appendix

Table - 1 Results of cluster analysis based on panel data capturing judiciary clearance rate carried out by CluMP and KML algorithms.

|           |  | Criminal cases |   | Other than criminal cases   |  |
|-----------|--|----------------|---|---|--|
|           |  | CluMP          | KML   | CluMP   | KML  |
| Cluster 1 | Armenia, Croatia, Estonia, Georgia, Malta, North Macedonia   |                | Armenia, Austria, Azerbaijan, Bosnia and Herzegovina, Bulgaria, Cyprus, Denmark, Finland, France, Hungary, Italy, Latvia, Lithuania, Malta, Netherlands, Norway, Poland, Republic of Moldova, Romania, Serbia, Slovenia, Spain, Sweden, Switzerland, Ukraine, Slovakia, Türkiye | Azerbaijan, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Hungary, Italy, Latvia, Lithuania, Malta, Netherlands, Norway, Republic of Moldova, Slovenia, Switzerland, Ukraine | Azerbaijan, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Hungary, Italy, Latvia, Lithuania, Malta, Montenegro, Netherlands, Norway, Poland, Republic of Moldova, Slovenia, Sweden, Switzerland, Ukraine, Slovak Republic |
| Cluster 2 | Austria, Bulgaria, Denmark, Finland, Hungary, Lithuania, Netherlands, Norway, Poland, Republic of Moldova, Romania, Spain, Sweden, Switzerland, Slovakia Azerbaijan, Bosnia and Herzegovina, Cyprus, |                | Estonia, Georgia, North Macedonia   | Cyprus, Montenegro, Poland, Sweden, Slovak Republic   | North Macedonia  |
| Cluster 3 | France, Italy, Latvia, Montenegro, Serbia, Slovenia, Ukraine, Türkiye  |                | Croatia, Montenegro   | North Macedonia, Serbia   | Serbia   |

Table - 2 Results of cluster analysis based on panel data capturing judiciary disposition time carried out by CluMP and KML algorithms.

|           |  | Criminal cases  |  | Other than criminal cases   |   |
|-----------|--|---|--|---|---|
|           |  | CluMP   | KML  | CluMP   | KML   |
| Cluster 1 | Albania, Andorra, Armenia, Azerbaijan, Malta   |   | Austria, Azerbaijan, Bulgaria, Denmark, Estonia, Georgia, Hungary, Lithuania, Netherlands, Norway, Poland, Romania, Switzerland, Slovak Republic | Azerbaijan, Czech Republic, Italy, Lithuania, Norway, Poland, Republic of Moldova, Romania, Ukraine | Azerbaijan, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, Georgia, Latvia, Lithuania, Netherlands, North Macedonia, Norway, Poland, Republic of Moldova, Romania, Slovenia, Sweden, Switzerland, Ukraine, Slovak Republic |
|           | Austria, Croatia, Cyprus, Italy, Monaco, Montenegro, North Macedonia, Republic of Moldova, Romania, Slovenia, Spain, Sweden, Ukraine, Türkiye, UK - England and Wales                                  | Armenia, Croatia, Finland, Latvia, Montenegro, North Macedonia, Republic of Moldova, Slovenia, Spain, Sweden, Ukraine | Bosnia and Herzegovina, Bulgaria, Cyprus, France, Georgia, Montenegro, Serbia  |   | France, Italy, Malta, Montenegro, Serbia  |
| Cluster 2 | Bosnia and Herzegovina, Bulgaria, Czech Republic, Denmark, Estonia, Finland, Georgia, Germany, Hungary, Latvia, Lithuania, Netherlands, Norway, Poland, Portugal, Serbia, Switzerland, Slovak Republic | Bosnia and Herzegovina, Cyprus, Italy, Malta, Serbia, Turkey  | Croatia, Denmark, Estonia, Finland, Latvia, Malta, Netherlands, North Macedonia, Slovenia, Sweden, Switzerland, Slovak Republic                  |   | Bosnia and Herzegovina, Cyprus  |
| Cluster 3 |  |   |  |   |   |