Summary: The main objective of this article is proving how the regional diversification in the size of general government sector influences the economies of EU countries. To achieve this, presenting both the size of the general government sector and of the economy, using variables which enable comparison in time and space, is essential. Bearing this in mind, the general government sector has been depicted by nine variables and the economy has been described by thirteen explanatory variables. The complexity of the phenomenon imposes the implementation of an unconventional approach in this field of exploration. Our approach is based on Intelligent Data Analysis (IDA) - a methodology that includes a set of techniques that can be applied for extracting useful knowledge from large amounts of data. In order to indicate the impact of regional diversification in the size of the general government sector on the EU countries’ economies, probabilistic techniques were applied – Bayesian Networks. Analysis made in the study showed that the largest impact of the GGS size on the economy was identified in Portugal and Slovakia. The results of the studies show that the most "responsive" to the size of the GGS variable describing the economy was gross domestic product per inhabitant. The research proved that the economies of some countries
showed similarities in the effect of the size of the general government sector on the parameters of the economy. We have identified five groups of such countries.

**Keywords:** general government sector size, economy, public finance, impact of size of the general government sector on the economy, Bayesian networks, probabilistic inference.

**Streszczenie:** Głównym celem artykułu jest ukazanie wpływu regionalnego zróżnicowania rozmiaru sektora finansów publicznych na gospodarki państw Unii Europejskiej. Do osiągnięcia tego celu konieczne jest zobrazowanie rozmiaru zarówno sektora finansów publicznych, jak i gospodarki – za pośrednictwem zmiennych umożliwiających ich porównanie w czasie i w przestrzeni. Mając to na uwadze, sektor finansów publicznych zobrazowano za pośrednictwem dziesięciu zmiennych, a gospodarka została opisana za pomocą trzynastu zmiennych wyjaśniających. Złożoność tego zjawiska wymaga realizacji niekonwencjonalnego podejścia w tej dziedzinie badań. Podejście autorów opiera się na inteligentnej analizie danych (IDA) – metodologii obejmującej zestaw technik, które można zastosować do wydobywania użytecznej wiedzy z dużej ilości danych. W celu wskazania opisanych związków w artykule zastosowało techniki probabilistyczne – sieci Bayesa.

**Słowa kluczowe:** wielkość sektora instytucji rządowych i samorządowych, gospodarka, finanse publiczne, wpływ sektora instytucji rządowych i samorządowych na gospodarkę, sieci Bayesa, wnioskowanie probabilistyczne.

**1. Introduction**

The main objective of this article is to show the impact regional diversification in the size of the general government sector has on the economies of EU countries. In order to do so, it is necessary to depict both the general government sector size and the economy itself. The presented research problem is particularly important because these issues have not been entirely settled and exhaustively discussed. However, the number of trials of defining the size of the general government sector (more on this topic: Skica et. al. [2015a], Skica et al. [2015b]) has draws attention, as well as the kind of monogamous method of determining its size.

Approaches to expressing the size of the general government sector present in the research focus mainly (with few exceptions) on measures combining public spending with GDP. This position is present in works published by Peden [1991], Vedder and Gallaway [1998], Chao and Gruber [1998], and Buti et al. [2003], as well as by Pevcin [2004], Chobanov and Mladenova [2009], and Davies [2009]. More recent publications also do not avoid referring to the expenditure measure as a quantifier of the size of the general government sector, among them we should mention Forte and Magazzino [2011], Magazzino [2012], and Di Liddo, Magazzino and Porcelli [2015].

The method of expressing the size of the general government sector (in a universal way) mentioned above translates into how the research on the impact of the size of the general government sector on the economy is oriented. The dominant approach in research is that the explained variable describing the state of the economy is usually the economic growth. This is visible in some earlier works [Ram
1986; Bairam 1988; Barth, Bradley 1988; Conte, Darrat 1988; Grier, Tullock 1989; Engen, Skinner 1992]. This is also mentioned in more recent publications [Fölster, Henrekson 2001; Heitger 2001; Bose et al. 2003; Kustepeli 2005; Jiranyakul 2007; Maku 2009; Forte, Magazzino 2011]. This is also the case with the most up-to-date works by Di Liddo, Magazzino and Porcelli [2015], Gisore, Kiprop, Kalio, Ochieng and Kibet [2014], Martins and Veiga [2014], Odhiambo [2015], Aleksandrovich and Upadhyaya [2015], Sriyana [2016], Lingxiao, Peculea, Xu [2015], Al-Fawwaz [2015], Acha, Michael and Essien [2017], Dudzevičiūtė, Šimelytė and Liučvaitienė [2017], and Abdieva, Baigonushova and Ganiev [2017].

The methodology used in measuring the size of the general government sector and its optimization from the point of view of the condition of those economies in question is also significant. Where the dimensions of the general government sector expressed through the share of public expenditure in GDP and the condition of the economies described by the economic growth connect, several research methods and approaches are used. The first one is the inverted U-shaped curve, called the Armey curve [Armey 1995], the Rahn curve [Rahn, Fox 1996] or the BARS curve. The last name refers to Barro [1989], Armey [1995], Rahn and Fox [1996], and Scully [1994; 1995]. These are the economists who developed this method. Empirical verification of the BARS curve was performed by Karras [1997], Chen and Lee [2005], Chiuou-Wei, Zhu and Kuo [2010], Gunalp and Dincer [2010], and Altunc and Aydin [2013]. Another approach is based on the use of a wide range of econometric techniques, including the use of cross-section estimates or panel data. Earlier works were based on cross-section studies and conducted by Cameron [1982], Landau [1983], Summer and Heston [1984], Kormendi, Meguire [1985], Marlow [1986], Ram [1986], Saunders [1988], and Agell, Lindh and Ohlsson [1997]. Panel studies were carried by Fölster and Henrekson [2001], Dar and AmirKhalkhali [2002], Pevcin [2004], Agell, Ohlsson and Thoursie [2006], Romero-Avila and Strauch [2008], Colombier [2009], Afonso and Fuceri [2010], Bergh and Karlsson [2010], Afonso, Jelles [2011], as well as by Di Liddo, Magazzino, Porcelli [2015], Gisore, Kiprop, Kalio, Ochieng, Kibet [2014], Lahirushan and Gunasekara [2015], Mose, Kalio, Kiprop, Babu [2015], Asimakopoulos and Karavias [2016], Mehrara and Keshtgar [2016]. The most recent articles with the use of panel studies were written by Sáez, Álvarez-Garcia, Rodriguez [2017] and Leshoro [2017].

As proven, the literature includes search of the dependencies connecting the size of the general government sector and the public sector with the economy using econometric modelling. There are usually two results of such research. The first is identifying measurements of the size of the general government sector (or public sector); the second is estimating the dependencies connecting the size of the sector with economic growth. Those approaches leave at least two issues unresolved – one being the impact of diversification of the general government sector size on the economies of various countries (tested using the same research methodology), and second the optimization of the size of the sector due to the criterion of the economic
results obtained by the examined country. Considering the indicated cognitive gap, this article undertakes the first of these issues, and adopts the general government sector and the economies of the EU countries as the object of analysis.

Considering the universality of the above presented attempts to explain the interrelations between the variable explain and explanatory variable, the authors of this article decided to use the Bayesian networks approach (BNs), a method that has not been used in this kind of research before.

The authors also adopted a different approach to defining variables. In place of the commonly used explanatory variable (public expenditure to GDP), the authors introduced nine variables describing the size of the general government sector. By contrast, the explained variable (economic growth) was described using thirteen variables. As a result, the size of the general government sector (the explanatory variable), was described by the following variables: total general government expenditure (euros per inhabitant), the ratio of total taxes to GDP (% GDP), public sector employment (number of people), total general government expenditure (euros per inhabitant), government consolidated gross debt (% GDP), general government sector output (% GDP), gross value added (general government total value-added) (basic (current) prices), total general government expenditure (% GDP), general government gross fixed capital formation (% GDP). The economy was described using the following set of explained variables: activity rate (in %), balance of the current account (million euros), external balance of goods and services (million euros), FDI – foreign direct investment (million USD), gross domestic product in current prices per inhabitant (GDP per inhabitant), harmonized indices of consumer prices (HICPs) (annual average rate of change), inward FDI flows (million USD), outward FDI flows (million USD), real effective exchange rate (index 1999 = 100), unemployment rate (in %), potential output of total economy (million euros), gross capital formation (% GDP), gross domestic product in current prices per inhabitant – dynamic (percentage change).

2. Methods

In the research, Bayesian networks were applied to identify the regional diversification of the impact of variables describing the public finance sector on the economies of the EU countries were applied. BN – being a complex representation of knowledge about causal relationships between properties of objects – is pair (DAG, CPD) where the nodes and arcs form a directed, acyclic graph (DAG) and the conditional probability distribution (CPD) of a node (random variable) is defined for every possible outcome of the preceding causal node(s) [Pearl 1988]. If there exists a causal probabilistic dependence between two random variables in the DAG, the corresponding two nodes are connected by a directed edge [Murphy 1998] while the directed edge from a node A to a node B indicates that the random variable A causes
the random variable B. Each node is annotated with a conditional probability distribution (CPD), that represents $p(B|A)$, where A denotes the preceding of B in DAG. A graphical representation of the casual probabilistic relationships among a set of random variables and their conditional dependences provides a compact representation of a joint probability distribution.

Figure 1 presents a simple BN with just two variables and one dependency. In this depiction we assume only two possible causes of inflation’s fall: supply factors and demand factors. These two causes are independent and are ancestors and parents of inflation, analogically inflation is a descendent and child of both supply factors and demand factors.

$$P(SFs) = \begin{array}{cc} Yes & No \\ 0.1 & 0.9 \end{array}$$

$$P(DFs) = \begin{array}{cc} Yes & No \\ 0.2 & 0.8 \end{array}$$

$$P(I|SFs, DFs) = \begin{array}{ccc} SFs & DFs & Yes \\ Yes & Yes & 1 \\ No & No & 0 \\ Yes & No & 0.5 \\ No & Yes & 0.5 \end{array}$$

**Fig. 1.** DAG representing two independent possible causes of inflation. Supply factors are denoted by SFs, demand factors are denoted by DFs and inflation fall is denoted by I

Source: own work.

The whole concept of Bayesian networks is built on the Bayes theorem, which helps us to express the conditional probability distribution of cause given the observed evidence using the converse conditional probability of observing evidence given the cause: $P\ [Cause \mid Effect] = P\ [Effect \mid Cause] * (P\ [Cause]/P\ [Effect])$. Additionally, BNs nodes are conditionally independent of its all non-descendants given that node’s parents. Hence, the joint probability distribution of all random variables in the graph factorizes into a series of conditional probability distributions of random variables given their parents. Therefore, we can build a full probability model by only specifying the conditional probability distribution in every node [Spiegelhalter 1998]. Due to this fact the BN can be used for inference. The most common inference methods are predictive inference (from causes to effects) - from new information about causes to new beliefs about the effects, according to the directions of the network arcs.

To demonstrate the BN inference we refer to our example. Let us assume now that supply factors SFs have changed. We observe SFs = yes with probability 1 and
we wonder how the probability distribution of inflation fall changes, given the observed cause:

\[
P( I = yes | SFs = yes) = \frac{P(I, SFs, DFs)}{P(SFs, DFs)} = \alpha \sum_{DFs} P(I, SFs, DFs) = \alpha \left[ P(SFs = yes)P(DFs = yes)P(I = yes | SFs = yes, DFs = yes) + P(SFs = yes)P(DFs = no)P(I = yes | SFs = yes, DFs = no) \right] = \alpha 0.06
\]

\[
P( I = no | SFs = yes) = 0.04, \text{ where } \alpha = 10.
\]

The graphical interpretation after observing the change of supply factors is presented in Figure 2.

**Fig. 2.** DAG representing two independent potential causes of inflation fall with posterior probability distribution, i.e. after observing the evidence

Source: own work.

BNs can be constructed manually or learned from data. With the increasing availability of data, learning is evidently a more feasible alternative for developing a Bayesian network. The Bayesian network learning problem can be categorized as: 1) a parameter learning problem when the structure is known, and 2) a structure learning problem when the structure is unknown. Our research focused on the latter issue. Among various methods of structure learning [Koski, Noble 2012], the greedy search provides a way to obtain a good model in a reasonable time frame as compared to other methods, learning BN structure is NP–hard. For a fixed amount of computational time, a greedy search with random restarts produces better models than either simulated annealing or best-first search does [Chickering 2002]. Taking this into consideration, in our research the greedy search were used. BNs were developed with the help of a heuristic algorithm using the Bayesian function of network structure to distribution matching as a scoring function, called K2 [Jensen, Nielsen 2007].
BNs were applied in a wide range of areas in health services research: health economic evaluation, health quality measurement, health outcomes monitoring, cost-effectiveness analysis; but also in epidemiology, clinical research, medical decision-making, public health and the economy (see: [Quang 2010; Gadewadikar at al. 2010; Harding 2011; Cobb 2011; Sesen at al. 2013; Skica et al. 2015a])

3. Methodology and results

In our research, quantitative data for the EU Member States that were collected for the period 2000-2013 (inclusive) from the Eurostat, OECD and World Bank databases were used. The collected data were prepared for analysis.

1. Cleaning up the missing attributes/cases. It was assumed that for further research the attributes/cases in which the number of missing values is less than a third of the total number, were selected. Then, single missing values were completed using advanced methods based on generalized additive models and the method of \( k \)-nearest neighbours.

2. Integrating data in the form of decision tables. Data from the general government sector size and the economy domain were placed in separate files split by years. For each year the data were combined in tables consisting of all variables from the economy domain and only one decision attribute (called decision) from the general government sector size, located in the rightmost column.

3. Selecting data for the countries which are members of the European Union. Identification of the relationships between the economy and the general government sector size was carried out from 2001 to 2013 for 27 member states. Croatia was omitted due to the large number of missing data.

4. Transforming data describing continuous attributes. Two methods of discretization process were used: equal-width, where the width of the intervals is constant and equal frequency where the frequency of instances in the interval is constant. Analysis of the our previous research (Skica et al. (2016)) indicated that learning models (Bayesian Networks) for four discretization ranges provided the highest classification efficiency.

Next the BNs were generated considering each year separately (i.e. the first network was generated on the basis of data from the year 2000, the second from 2001, etc.) and based on data for 2000, 2007, 2009 and 2013 the predictive inference were conducted. We chose the data for 2000, 2007, 2009 and 2013 for the following reasons. The Year 2000 was the first one for which all of the data were available for all of the analyzed countries; 2007 was a year of prosperity not only in Poland but also in other EU countries. Conversely, 2009 was a year of recession in the EU economies. Finally, 2013 was the last year with the complete data for all of the EU economies at the time when the calculations was made.
Table 1. Results of predictive inference. Detailed description provided in the text

<table>
<thead>
<tr>
<th>Country</th>
<th>Activity rate (in %)</th>
<th>Balance of the current account (million euro)</th>
<th>External balance of goods and services (million euro)</th>
<th>FDI - foreign direct investment (million USD)</th>
<th>Gross domestic product in current prices (GDP per inhabitant)</th>
<th>Harmonized indices of consumer prices (annual average rate of change)</th>
<th>Inward FDI flows (million USD)</th>
<th>Outward FDI flows (million USD)</th>
<th>Real effective exchange rate (index 1999 = 100)</th>
<th>Unemployment rate (in %)</th>
<th>Potential output of total economy (million euro)</th>
<th>Gross capital formation (% GDP)</th>
<th>Gross domestic product in current prices per inhabitant (% change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td></td>
<td>0.79</td>
<td></td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>0.97</td>
<td>0.75</td>
<td>0.99</td>
<td>0.96</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulgaria</td>
<td>0.99</td>
<td>0.97</td>
<td>0.99</td>
<td>0.75</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyprus</td>
<td>0.87</td>
<td>0.75</td>
<td>0.99</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cz. Republic</td>
<td>0.94</td>
<td>0.99</td>
<td>0.95</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>0.94</td>
<td>0.71</td>
<td>0.71</td>
<td>0.99</td>
<td>0.78</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estonia</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>0.73</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>0.85</td>
<td>0.85</td>
<td></td>
<td>0.74</td>
<td>0.71</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>0.85</td>
<td>0.73</td>
<td></td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>0.89</td>
<td>0.82</td>
<td></td>
<td>0.73</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td>0.92</td>
<td>0.76</td>
<td>0.91</td>
<td>0.73</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>0.95</td>
<td>0.91</td>
<td>0.98</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>0.97</td>
<td>0.81</td>
<td></td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>0.85</td>
<td>0.85</td>
<td>0.79</td>
<td>0.98</td>
<td>0.98</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td>0.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latvia</td>
<td>0.71</td>
<td>0.97</td>
<td>0.99</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lithuania</td>
<td>0.95</td>
<td>0.99</td>
<td>0.97</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luxembourg</td>
<td>0.97</td>
<td>0.98</td>
<td>0.75</td>
<td>0.99</td>
<td>0.75</td>
<td>0.99</td>
<td></td>
<td></td>
<td></td>
<td>0.99</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malta</td>
<td>0.90</td>
<td>0.99</td>
<td>0.98</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.74</td>
<td>0.90</td>
<td></td>
<td>0.98</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>0.99</td>
<td>0.99</td>
<td>0.96</td>
<td>0.99</td>
<td>0.74</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portugal</td>
<td>0.89</td>
<td>0.99</td>
<td>1.00</td>
<td>0.96</td>
<td>0.98</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td>0.75</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romania</td>
<td>0.97</td>
<td>0.99</td>
<td>0.96</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slovakia</td>
<td>0.90</td>
<td>0.99</td>
<td>0.99</td>
<td>0.94</td>
<td>0.93</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td>0.75</td>
<td>0.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slovenia</td>
<td>0.90</td>
<td>0.99</td>
<td>0.97</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>0.99</td>
<td>0.99</td>
<td></td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>0.74</td>
<td>0.99</td>
<td>0.99</td>
<td>0.73</td>
<td>0.73</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.K.</td>
<td>0.94</td>
<td>0.71</td>
<td>0.99</td>
<td>0.73</td>
<td>0.87</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: own work.
Specifying the values of the variables (causes) describing the general government sector for each EU country, the effects were observed separately, i.e. changing the probability of distributions of attributes from the economy. It has been assumed that the result of inference is the values of the observed attributes (effects) that have reached a probability level greater than or equal to 0.7. In this way, on the basis of the developed model, for each EU country it was determined which variables in the area of the PFS economy were most strongly affected (see Table 1).

The largest impact of PFS size on the economy was observed in Portugal and Slovakia. The size of the PFS in these countries significantly affected as many as eight of the thirteen variables characterizing the economy. Interestingly, in both cases seven variables on the side of the economy occurred in both countries as listed below:

1) activity rate (in %),
2) balance of the current account (million euro),
3) external balance of goods and services (million euro),
4) FDI – foreign direct investment (million USD),
5) gross domestic product in current prices per inhabitant (GDP per inhabitant),
6) inward FDI flows (million USD),
7) outward FDI flows (million USD).

Additionally, in the case of Slovakia, the size of PFS significantly affected the potential output of the whole economy (million euros), while in Portugal it influenced the unemployment rate (in %). In the case of Italy, the size of PFS affected seven, and in the case of Bulgaria, Luxembourg and Poland, six variables from the economy. Only three variables were repeated in those three countries:

1) activity rate (in %),
2) external balance of goods and services (million euro),
3) gross domestic product in current prices per inhabitant (GDP per inhabitant).

In as many as nine countries, the influence of the PFS size on the five variables describing the economy was noted, and for the next nine countries, the size of PFS influenced four economic variables.

The size of the public finance sector in Austria and Spain affected two completely different economic variables, while PFS in Belgium affected only one variable – gross domestic product in current prices per inhabitant (GDP per inhabitant).

In order to visualize the above-mentioned relationships, it is necessary to divide all the analysed countries into four groups depending on the strength of the impact of PFS variables on the economy. Each of the groups of countries was assigned a number from 1 to 4. The result is shown on the Figure 3.

After applying the undertaken assumptions on the map, interesting relationships emerge. Countries in which the variables are influenced by the public sector were
Fig. 3. Groups of EU countries characterized by a similar level of influence of PFS size on the economy

Source: own work.

marked with shades of grey. It is clear that in the European Union it is possible to distinguish groups of countries, clearly grouped in that space. Given the geographical location, it can be clearly seen that countries with a medium PFS impact on the economy are arranged in clusters. Group 3, where PFS affects five variables on the part of the economy, are mostly countries located on the outskirts of Europe, namely the Scandinavian and island countries, with the only exception being Hungary. In turn, the countries assigned to group 2 (PFS affects 4 variables on the side of the economy) are mainly the countries of Central Europe.

Unfortunately, apart from the location criterion, it is difficult to find other ones that may affect the grouping of individual countries. We could, for example, use the following criteria:

1) entry date to the EU;
2) having/not having the euro as a currency;
3) the frequency of the country's excessive deficit procedure;
4) status of the country: social/non-social (i.e. the scale of expenditure allocated for social purposes in the total expenditure or in relation to GDP);
5) distribution of public finances between the central and the local governments (which part of public funds is assigned centrally and which is locally).
However, the authors of the study took into account the data on PFS and their economies for a very long period of time (2000-2013), which meant that other similarity criteria could not be taken into account due to their fluctuations in the analysed period of analysis. Other criteria, such as size in km$^2$ or population did not prove useful.

4. Summary and conclusions

The discussion about the role and size of the general government sector has been ongoing for a very long time. In addition to proving that the size of the general government sector affects the economy, it is difficult to find guidance on its optimal size. For the purposes of this study, the authors assumed that through the optimal size, the general government sector would understand the size of individual values describing the sector that exerts the strongest positive impact on economic variables. Analysis made in the study allowed firstly to indicate in which countries the size of the general government sector influences the variables on the side of the economy. It turned out that the largest impact on the economy of the size of the general government sector is exerted in Portugal and Slovakia, as many as eight out of thirteen variables describing the economy in these countries showed a dependence on variables expressing the size of the general government sector. The next stage of the research was the identification of countries that show similarity in terms of the impact of the general government sector on economic variables. Map 1 shows these countries marked with similar colors. The results of the studies show that the most "responsive" to the size of the GGS variable describing the economy was the variable of gross domestic product in current prices per inhabitant (GDP per inhabitant), which had all the signs of "responsive" in 22 countries. Variable activity rate (in %) were determined in 20 cases, while variables: external balance of goods and services (million euro) or FDI – foreign direct investment (million USD) showed variability in the case of eighteen countries. The above variables describing the economy are the most mildly affected by the sizes of GGS. The rest of the variables exceeded the value of the probability equal to 0.7 in a small number of cases, which suggests that the size of the GGS has no significant effect on the parameters describing the economy.

The economies of some countries showed similarities in the effect of the sizes of the general government sector on the parameters of the economy. In the case of countries such as Portugal, Slovakia, Slovenia, Bulgaria and Estonia, the sizes of the GGS affected the following economic variables:

1) activity rate (in %),
2) balance of the current account (million euros),
3) external balance of goods and services (million euros),
4) FDI – foreign direct investment (million USD),
5) gross domestic product in current prices per inhabitant (GDP per inhabitant).
The size of the general government sector in Italy, Czech Republic, Poland, Hungary, Romania, Lithuania and Latvia influenced the following economic variables:

1) activity rate (in %),
2) external balance of goods and services (million euros),
3) FDI – foreign direct investment (million USD),
4) gross domestic product in current prices per inhabitant (GDP per inhabitant).

However, the economic variables:

1) external balance of goods and services (million euros),
2) FDI – foreign direct investment (million USD),
3) gross domestic product in current prices per inhabitant (GDP per inhabitant),

were affected by the size of the general government sector in Denmark and Sweden.

The next three variables:

1) activity rate (in %),
2) external balance of goods and services (million euros),
3) FDI – foreign direct investment (million USD),

were influential in Greece and Malta.

Finally, in Germany and the UK the size of the general government sector affected the following economic variables:

1) activity rate (in %),
2) external balance of goods and services (million euros).

Bibliography


The impact of regional diversification in the size…


