КЛАСТЕРИЗАЦИЯ НА ПАЗАРА НА ТРУДА В БЪЛГАРИЯ

Десислава Стефанова
Университет за национално и световно стопанство, София

Резюме

Въз основа на седем показателя, характеризиращи състоянието на трудовия пазар, 28-те области са групирани в шест еднородни кластера и са определени показателите, които допринасят в най-голяма степен за възникването на регионални различия на трудовия пазар през 2010 г. Получените резултати дават основание да се смята, че приложението на кластерния анализ за период от няколко последователни години ще позволи да бъдат разкрити показателите с трайно изразени разграничителни способности. На тази основа е възможно да бъдат формулирани коректни препоръки за въздействие върху регионалните различия на трудовия пазар.

Ключови думи: регионални различия на трудовия пазар, кластер анализ на К-средните, коефициент на заетост, коефициент на безработица, коефициент на икономическа активност.

Key words: regional differences on labour market, K-Means Cluster, coefficient of employment, coefficient of unemployment, coefficient of economic activity.

JEL: J01, J08, J21.

Увод

Един от съществените проблеми на пазара на труда са регионалните различия в някои основни показатели, характеризиращи неговото състояние [8, 13]. Те се проявяват по-силно на по-ниски териториални нива и предизвикват значителни вътрешнорегионални различия в районите за планиране на ниво NUTS 2 [9]. Ефективното въздействие върху тях изисква разработването на системи от мерки, които да се прилагат съобразно със състоянието на трудовите пазари в съответните административно-териториални единици (области и общини). За дефинирането на определени мерки е целесъобразно областите със сходни показатели за пазара на труда да бъдат обединени в еднородни групи (кластери). Това ще позволи за всеки един кластер да се предложи обща рамка от мерки за въздействие върху различията на трудовия пазар. Впоследствие за всяка една област съответната система от мерки ще се конкретизира и допълни в областната ѝ стратегия за развитие.

Целта на настоящата статия е въз основа на показатели, характеризиращи състоянието на трудовия пазар, 28-те области да бъдат групирани в кластери.

1. Показатели, характеризиращи състоянието на трудовия пазар

В настоящата статия под понятието пазар на труда се разбира: система от законодателни актове, документи и институции, която регламентира отношенията между субектите, осъществяващи сделка по покупко-продажба на трудова услуга. Така възприетото работно понятие не отразява всички аспекти на пазара на труда. Въпреки това насочва вниманието към неговата сложност и двата основни процеса – безработица и заетост, които се формират на него и го характеризират. Тези процеси се оценяват чрез коефициента на безработица и коефициента на заетост. Първият се определя от отношението между безработните лица и икономически активното население, а вторият от отношението между заетите лица и населението на същата възраст [15]. Променливите използвани при пресмятането на двата показателя водят до съществени различия между тях. Това поражда спорове в научните среди относно това кой от двата коефициента характеризира по-добре състоянието на икономиката и в частност пазара на труда. Една част от икономистите смята, че коефициента на безработица преувеличава слабостите на икономиката, а другата смята, че коефициента на заетост преувеличава силните ѝ страни [3, 4].

В настоящата статия се възприема тезата на проф. Л. Дулевски, че при анализ на пазара на труда по-широко следва да се прилага интегрирания подход. Той намира израз в система от множество показатели за характеризиране състоянието на пазара на труда [3, 4]. В тази връзка в настоящото изследване освен коефициентите на заетост и безработица се използват и следните показатели: цена на труда; производителност на труда; относителен дял на населението на възраст между 25 и 64 години с висше образование; относителен дял на населението на възраст между 25 и 64 години с основно и по-ниско образование; и коефициент на икономическа активност. Между тези показатели съществува логическа връзка и взаимодействие, която дава представа за състоянието на пазара на труда. Така например по-високото образование на икономически активното население е пред-

поставка за по-висока производителност и цена на труда. Същевременно по-ниското образование на работната сила се свързва с по-високо равнище на безработица, а то от своя страна поставя изисквания за по-висока производителност на труда. В настоящото изследване за цена на труда е използвана средната годишна работна заплата на наетите лица по трудово и служебно правоотношение, а за производителност на труда – добавената стойност по факторни разходи на един нает. В таблица 1. са представени наименованите стойности на показателите, характеризиращи състоянието на трудовия пазар, чрез които ще се извърши кластеризацията на 28-те области.

Към техния брой и познавателни възможности е вероятно да бъдат отправени критични бележки и препоръки, които авторът ще приеме с благодарност. Въпреки това следва да се отбележи, че при квантифициране на многомерен признак, какъвто в случая е пазарът на труда изисквания за изчерпателност не могат да се предявяват, тъй като винаги остава нещо неуловимо от системата от едномерни признаци (в случая показателите, представени в табл. 1) [11].

Табл. 1. Показатели, характеризиращи състоянието на трудовия пазар по области през 2010 г.

Области	Коефициент на безработица, %	Коефициент на заетост 15-64 навършени години, %	Средна годишна работна заплата на наетите лица по трудово и служебно правоотношение, лв.	Добавена стойност по факторни разходи на един нает, лв./нает	Относителен дял на населението на възраст между 25-64 години с висше образование, %	Относителен дял на населението на възраст между 25-64 години с основно и по-ниско образование, %	Коефициент на икономическа активност - 15-64 навършени години, %
Благоевград	5,80	64,90	5 683,00	11,42	15,90	23,60	68,80
Бургас	9,60	57,20	7 092,00	8,07	15,60	32,20	63,30
Варна	8,20	60,80	7 533,00	8,07	25,60	20,20	66,30
Велико Търново	13,10	51,40	6 119,00	8,40	23,00	17,70	59,30
Видин	13,10	48,70	5 422,00	8,98	18,90	17,50	56,10
Враца	8,60	53,10	8 509,00	6,24	20,00	16,80	58,10
Габрово	7,50	61,60	6 591,00	9,35	24,20	14,20	66,60
Добрич	16,40	55,20	6 130,00	9,00	17,90	29,40	66,10
Кърджали	4,90	46,30	5 643,00	8,20	10,30	49,50	48,70
Кюстендил	9,00	60,80	5 891,00	10,32	17,20	16,60	66,80
Ловеч	8,00	59,80	5 936,00	10,07	21,60	17,40	65,10
Монтана	12,20	55,40	6 028,00	9,19	15,40	15,90	63,10
Пазарджик	15,50	57,80	6 087,00	9,50	15,20	25,60	68,50
Перник	6,90	61,90	5 890,00	10,51	16,30	14,00	66,60
Плевен	12,80	52,40	6 093,00	8,60	17,50	21,20	60,10
Пловдив	8,50	59,40	6 462,00	9,19	22,50	21,00	64,90
Разград	22,60	46,40	6 071,00	7,64	12,10	44,70	59,90
Русе	6,80	62,00	6 354,00	9,76	20,40	20,70	66,50
Силистра	11,40	50,00	5 771,00	8,66	13,10	39,30	56,60
Сливен	16,70	54,00	6 000,00	9,00	18,40	26,40	65,00
Смолян	19,20	53,30	6 134,00	8,69	17,20	28,50	66,00
София, столица	6,60	70,60	10 547,00	6,69	44,80	5,30	75,60
София, област	7,80	62,50	7 749,00	8,07	13,10	17,60	67,90
Стара Загора	6,70	63,50	8 080,00	7,86	19,30	16,10	68,10
Търговище	11,70	54,20	5 620,00	9,64	13,10	36,40	61,50
Хасково	13,20	60,50	5 638,00	10,73	19,40	24,30	69,80
Шумен	28,80	50,10	6 214,00	8,06	18,80	33,90	70,50
Ямбол	14,80	56,90	6 034,00	9,43	19,20	21,80	67,00

Източник: НСИ.

2. Кластерен анализ с К-средни величини

Методите за формиране на кластери се разделят на йерархични и нейерархични. Поради някои недостатъци на първите (вж. Манов, 2002) в настоящото изследване се предпочитат методите за нейерархична кластеризация. Най-широко използвания от тях е метода на К-средните величини [6].

Алгоритъмът на работа изисква седемте показателя, които са наименовани величини да се трансформират в ненаименовани. За целта се прилага формулата на класическата стандартизация [10, 11]:

$$z_{ij} = \frac{x_{ij} - \overline{x}_j}{S_j}, \qquad (1)$$

където

z_{ij} е стандартизираното значение на j-ия признак при i-та единица;

x_{ij} – значението на j-ия признак при i-ата единица;

$\overline{x}_j$ – средната за съответния j-и признак и се пресмята чрез формула (2):

$$\overline{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}, \qquad (2)$$

където

n е броят на единиците в съответната съвкупност;

S_j – стандартното отклонение за j-ия признак и се пресмята се чрез следната формула:

$$S_j = \left[\frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \overline{x}_j)^2\right]^{\frac{1}{2}} \qquad (3)$$

След превръщането на наименованите величини в ненаименовани се пресмята евклидовото разстояние между единиците в многомерното пространство. Това се извършва чрез формула (4) [6, 10]:

$$C_{rs} = \left[\sum_{j=1}^{m}(z_{qj} - z_{pj})^2\right]^{\frac{1}{2}}, \qquad (4)$$

където

z_{qj} е стандартизираното значение на j-ия признак при q-ата единица;

z_{pj} – стандартизираното значение на j-ия признак при p-ата единица;

m – броят на признаците.

Въз основа на така пресметнатите евклидови разстояния между отделните единици се съставя матрицата на разстоянията. От нея става ясно между кои единици разстоянието е най-голямо. Обикновено те се приемат и за начални кластерни центрове. Останалите единици се присъединяват към съответния кластер на принципа на най-близкото разстояние до неговия център. Включването на всяка нова единица към кластера изисква да се изчислят промените в неговия кластерен център. Когато те надхвърлят предварително зададения критерий за сходство трябва отново да се пресметнат разстоянията между единиците и новите центровете, и ако е необходимо да се извърши прегрупиране според условието за най-малко разстояние [6].

3. Кластеризация на трудовия пазар в България през 2010 г.

Чрез формула (1) наименованите величини от табл. 2 се трансформират в ненаименовани. Техните стандартизирани значения са представени в табл. 2.

Табл. 2. Стандартизирани значения на показателите, характеризиращи състоянието на трудовия пазар по области през 2010 г.

Области	Коефициент на безработица	Коефициент на заетост 15-64 навършени години	Средна годишна работна заплата на наетите лица по трудово и служебно правоотношение	Добавена стойност по факторни разходи на един нает	Относителен дял на населението на възраст между 25-64 години с висше образование	Относителен дял на населението на възраст между 25-64 години с основно и по-ниско образование	Коефициент на икономическа активност - 15-64 навършени години
Благоевград	-1,085	1,395	-0,727	2,208	-0,467	-0,026	0,835
Бургас	-0,381	0,067	0,565	-0,733	-0,516	0,853	-0,205
Варна	-0,641	0,688	0,970	-0,733	1,104	-0,373	0,362
Велико Търново	0,267	-0,933	-0,327	-0,443	0,683	-0,628	-0,962

(продължение)

Области	Коефициент на безработица	Коефициент на заетост 15-64 навършени години	Средна годишна работна заплата на наетите лица по трудово и служебно правоотношение	Добавена стойност по факторни разходи на един нает	Относителен дял на населението на възраст между 25-64 години с висше образование	Относителен дял на населението на възраст между 25-64 години с основно и по-ниско образование	Коефициент на икономическа активност - 15-64 навършени години
Видин	0,267	-1,399	-0,966	0,066	0,019	-0,649	-1,567
Враца	-0,566	-0,640	1,865	-2,339	0,197	-0,720	-1,189
Габрово	-0,770	0,826	0,106	0,391	0,877	-0,986	0,419
Добрич	0,879	-0,278	-0,317	0,083	-0,143	0,567	0,324
Кърджали	-1,252	-1,813	-0,764	-0,619	-1,375	2,620	-2,967
Кюстендил	-0,492	0,688	-0,536	1,242	-0,257	-0,741	0,457
Ловеч	-0,678	0,516	-0,495	1,023	0,456	-0,659	0,135
Монтана	0,101	-0,243	-0,411	0,250	-0,548	-0,812	-0,243
Пазарджик	0,712	0,171	-0,357	0,522	-0,581	0,179	0,778
Перник	-0,881	0,878	-0,537	1,409	-0,403	-1,006	0,419
Плевен	0,212	-0,761	-0,351	-0,268	-0,208	-0,271	-0,811
Пловдив	-0,585	0,447	-0,013	0,250	0,602	-0,291	0,097
Разград	2,027	-1,796	-0,371	-1,110	-1,083	2,130	-0,848
Русе	-0,900	0,895	-0,112	0,751	0,261	-0,322	0,400
Силистра	-0,048	-1,175	-0,646	-0,215	-0,921	1,578	-1,473
Сливен	0,934	-0,485	-0,436	0,083	-0,062	0,261	0,116
Смолян	1,397	-0,606	-0,313	-0,189	-0,257	0,475	0,305
София, столица	-0,937	2,379	3,734	-1,944	4,214	-1,895	2,121
София, област	-0,715	0,981	1,168	-0,733	-0,921	-0,638	0,665
Стара Загора	-0,918	1,154	1,471	-0,917	0,083	-0,792	0,703
Търговище	0,008	-0,450	-0,785	0,645	-0,921	1,282	-0,546
Хасково	0,286	0,636	-0,768	1,602	0,100	0,046	1,024
Шумен	3,176	-1,158	-0,240	-0,742	0,002	1,027	1,156
Ямбол	0,582	0,015	-0,405	0,461	0,067	-0,209	0,494

Стандартизираните ненаименовани величини се обработват с програмния продукт SPSS 16.0. Въз основа на получена дендрограма броя на кластерите се определя на шест. Използва се критерий за сходство 0,001 и опцията Use running means, т.е. центровете на кластерите се актуализират след присъединяването на всяка единица [2]. Първоначалните кластерни центрове са представени в табл. 3. От тях може да се направят някои заключения за разпределените 28 области. В първи кластер се отнасят области, характеризиращи се със сравнително висока добавена стойност по факторни разходи на един нает, а във втори области с нисък относителен дял на населението на възраст между 25-64 години с висше образование. В трети кластер се включват области с относително високи коефициенти на заетост, средна годишна работна заплата и коефициенти на икономическа активност. В четвърти и пети кластер попадат области със сравнително нисък коефициент на заетост, а в шести области с относително висока средна годишна работна заплата.

Табл. 3. Първоначални кластерни центрове

Показатели	Кластери					
	1	2	3	4	5	6
Коефициент на безработица	-1,08	-1,25	-0,94	3,18	0,27	-0,92
Коефициент на заетост 15-64 навършени години	1,40	-1,81	2,38	-1,16	-1,40	1,15
Средна годишна работна заплата на наетите лица по трудово и служебно правоотношение	-0,73	-0,76	3,73	-0,24	-0,97	1,47
Добавена стойност по факторни разходи на един нает	2,21	-0,62	-1,94	-0,74	0,07	-0,92
Относителен дял на населението на възраст между 25-64 години с висше образование	-0,47	-1,38	4,21	0,00	0,02	0,08

(продължение)

Показатели	Кластери					
	1	2	3	4	5	6
Относителен дял на населението на възраст между 25-64 години с основно и по-ниско образование	-0,03	2,62	-1,90	1,03	-0,65	-0,79
Коефициент на икономическа активност 15-64 навършени години	0,84	-2,97	2,12	1,16	-1,57	0,70

Крайните кластерни центрове са определени след шестата итерация. Те са представени в табл. 4. На принципа на най-близкото разстояние към тях се присъединяват и съответните области. Към първи кластер се включват осем области: Благоевград, Габрово, Кюстендил, Ловеч, Перник, Пловдив, Русе и Хасково. Във втори кластер попадат областите Кърджали и Силистра, а в т. нар. трети кластер е единствено област София-столица. Към четвърти кластер се присъединяват областите Разград и Шумен, а към пети Велико Търново, Видин, Добрич, Монтана, Пазарджик, Плевен, Сливен, Смолян, Търговище, Ямбол. В последния шести кластер попадат областите Бургас, Варна, Враца, София-област, Стара Загора (вж. табл. 5).

Табл. 4. Крайни кластерни центрове

Показатели	Кластери					
	1	2	3	4	5	6
Коефициент на безработица	-0,64	-0,65	-0,94	2,60	0,54	-0,64
Коефициент на заетост 15-64 навършени години	0,79	-1,49	2,38	-1,48	-0,50	0,45
Средна годишна работна заплата на наетите лица по трудово и служебно правоотношение	-0,39	-0,71	3,73	-0,31	-0,47	1,21
Добавена стойност по факторни разходи на един нает	1,11	-0,42	-1,94	-0,93	0,12	-1,09
Относителен дял на населението на възраст между 25-64 години с висше образование	0,15	-1,15	4,21	-0,54	-0,20	-0,01
Относителен дял на населението на възраст между 25-64 години с основно и по-ниско образование	-0,50	2,10	-1,90	1,58	0,02	-0,33
Коефициент на икономическа активност 15-64 навършени години	0,47	-2,22	2,12	0,15	-0,21	0,07

Табл. 5. Разпределение на областите в кластери

Кластери	Области
Първи кластер	Благоевград, Габрово, Кюстендил, Ловеч, Перник, Пловдив, Русе, Хасково
Втори кластер	Кърджали, Силистра
Трети кластер	София-столица
Четвърти кластер	Разград, Шумен
Пети кластер	Велико Търново, Видин, Добрич, Монтана, Пазарджик, Плевен, Сливен, Смолян, Търговище, Ямбол
Шести кластер	Бургас, Варна, Враца, София-област, Стара Загора

От началните и крайните кластерни центрове могат да се направят някои обобщения за областите, които са присъединени към тях:

- В първи кластер попадат 8 области, характеризиращи се с относително висока добавена стойност по факторни разходи на един нает и сравнително ниски спрямо останалите области коефициент на безработица и средна годишна работна заплата на наетите по трудово и служебно правоотношение;
- Във втори кластер се включват области с ниски коефициенти на икономическа активност и заетост, ниска средна годишна работна заплата, нисък относителен дял на населението на възраст между 25-64 години с висше образование и висок относителен дял на населението между 25-64 години с основно и по-ниско образование;
- В т. нар. трети кластер попада единствено област София, столица. Тя се разграничава от останалите области по ниското равнище на безработица и сравнително ниския относителен дял на населението на възраст между 25-64 години с основно и по-ниско образование, както и с високи стойности на коефициента на заетост, коефициента на икономическа активност и средната годишна работна заплата;
- Към четвърти кластер се присъединяват области с високо равнище на безработица, ниско равнище на заетост и висок относителен дял на населението на възраст между 25-64 години с основно и по-ниско образование;
- В пети кластер се включват области със сравнително високо равнище на безработи-

ца, ниска икономическа активност, относително висока добавена стойност по факторни разходи на един нает и ниска средна годишна работна заплата;

- Към шести кластер се включват области със сравнително висока средна годишна работна заплата и нисък относителен дял на населението на възраст между между 25-64 години с основно и по-ниско образование;
- Може да се смята, че състоянието на пазара на труда е най-благоприятно в т. нар. трети кластер, където попада единствено област София-столица. От матрицата на разстоянията (вж. табл. 6) е видно, че най-близо до него е разположен шести кластер, а най-отдалечени са втори и четвърти кластер;
- Относно показателите средна годишна работна заплата на наетите лица по трудово и служебно правоотношение, коефициент на безработица и коефициент на заетост се наблюдават най-големите разстояния между центровете на кластерите (вж. табл. 4). Това означава, че тези показатели допринасят в най-голяма степен за обособяването на шестте кластера и констатираните регионални различия на трудовия пазар. Този извод се потвърждава и от представените в табл. 7. F-отношения. Тъй като областите са обособени в кластери не чрез случаен подбор, а в съответствие с изискването за най-близко разстояние, резултатите за равнище на значимост (Sig.) не могат да се използват за проверка на хипотези за средни величини [7]. Въпреки това от размерите на F-отношенията могат да се направят изводи за разграничителните способности на седемте показателя, а именно: колкото по-голям е размера на F-отношението, толкова по-значим е приноса на съответния показател за извършената кластеризация и регионалните различия на трудовия пазар.

Табл. 6. Матрица на разстоянията между крайните кластерни центрове

Кластери	1	2	3	4	5	6
1		4,827	7,081	4,965	2,204	2,776
2	4,827		10,033	4,155	3,466	4,510
3	7,081	10,033		9,126	7,804	5,951
4	4,965	4,155	9,126		3,005	4,529
5	2,204	3,466	7,804	3,005		2,607
6	2,776	4,510	5,951	4,529	2,607	

Табл. 7. F-отношения

	Cluster		Error		F	Sig.
	Mean Square	Df	Mean Square	Df		
Коефициент на безработица	4,693	5	0,206	22	22,775	0,000
Коефициент на заетост 15-64 навършени години	4,580	5	0,232	22	19,767	0,000
Средна годишна работна заплата на наетите лица по трудово и служебно правоотношение	5,157	5	0,101	22	51,261	0,000
Добавена стойност по факторни разходи на един нает	4,358	5	0,282	22	15,442	0,000
Относителен дял на населението на възраст между 25-64 години с висше образование	4,306	5	0,294	22	14,633	0,000
Относителен дял на населението на възраст между 25-64 години с основно и по-ниско образование	3,987	5	0,367	22	10,874	0,000
Коефициент на икономическа активност - 15-64 навършени години	3,333	5	0,515	22	6,467	0,001

В заключение следва да се отбележи, че направения кластерен анализ се отнася за една година (2010 г.). Чрез него трудовите пазари по области се систематизират в еднородни групи и се разкриват показателите, които допринасят в най-голяма степен за появата на регионални различия на трудовия пазар. Тъй като една година е твърде кратък период от време, отразяващ моментно състояние на трудовите пазари по области, настоящото изследване не се ангажира с формулирането на препоръки за въздействие върху регионалните различия. Тяхното извеждане и прецизиране изисква да се разкрият показателите, характеризиращи състоянието на регионалните трудови пазари, с трайно изразени разграничителни способности. Изпълнението на тази задача е обект на бъдеща изследователска работа, свързана с приложението на кластерния анализ за период от няколко последователни години.

Литература

1. Владимирова, К. *Икономика на труда*. Сиела. С. 2009.

2. Гоев, В. *Статистическа обработка и анализ на информацията от социологически, маркетингови и политически изследвания със SPSS*. Унив. изд. „Стопанство". С. 1996.
3. Дулевски, Л. *Пазарът на труда. Теоретико-методологически подходи и предизвикателства пред Република България*. Тракия-М. С. 2009.
4. Дулевски, Л. *Пазар на труда*. Сиела С. 2010.
5. *Закон за регионалното развитие*. ДВ. бр. 50/2008.
6. Манов, А. *Многомерни статистически методи със SPSS*. Унив. изд. „Стопанство". С. 2002.
7. Манов, А. *Статистика със SPSS*. Тракия-М. С. 2001.
8. *Актуализирана стратегия по заетостта на Република България 2008-2015*. Министерство на труда и социалната политика. С. 2008.
9. *Оперативна програма „Регионално развитие" 2007-2013*. Министерство на регионалното развитие и благоустройството. Главна дирекция „Програмиране на регионалното развитие. С. 2007.
10. Мишев, Г. *Международни статистически сравнения*. ВИИ „Карл Маркс". С. 1988.
11. Мишев, Г. *Квантификация на многомерни признаци*. Статистика. 1992. кн. 3. С. стр. 3-29.
12. Павлов, Н. *Осигурителни схеми*. Фабер. С. 2011.
13. *Пазарът на труда в условия на финансова и икономическа криз*а – предизвикателства и възможни решения. Становище на икономическия и социален съвет. С. 2009.
14. Шопов, Д., Стефанов, Л., Паунов, М., Дулевски, Л. *Икономика на труда*. С. 1999.
15. Национален статистически институт. [www.nsi.bg].

K-MEANS CLUSTER ANALYSIS ON THE LABOUR MARKET IN BULGARIA

Desislava Stefanova
University of National and World Economy, Sofia, Bulgaria

Abstract

The 28 districts, which are included in the regions on level NUTS 2, are combined in 6 different clusters on the basis of 7 indicators. The last ones are used for labour market characterization. Through K-means cluster analysis are revealed the indicators with higher contribution to the regional differences on the labour market in 2010. The results from the analysis give grounds for the following statement: the application of K-means cluster analysis for a period of some consecutive years will reveal the indicators with long lasting dividing characteristics. On this basis is possible some recommendations for influence on the labour market differences to be formulated.