

Използване на методите за оценка на характеристики за малки подсъвкупности при идентифициране на равнището на локалната безработица

ас. Александър Найденов

УНСС, катедра „Статистика и иконометрия“

e-mail: anaidenov@gmail.com

Резюме: Статията представя резултатите от експерименталното оценяване на равнището на безработица в общините на Република България, използвайки методите за оценка на характеристики на малки подсъвкупности. Необходимостта от подобен род оценки е продиктувана от обективно съществуващата нужда от данни с висока степен на териториална дезагрегация за целите на провеждането на коректно локализиране политики на въздействие. В изложението са изяснени основните понятия, свързани с използваната специфична методология и е даден един обобщен модел на подхода за анализ на изследваното явление. Направен е кратък обзор на източниците на информация, а резултатите от проведеното изследване са представени графично и таблично за всеки един използван метод, като в допълнение е направен и кратък критичен анализ. Във връзка с практическото приложение на специфичната методология са представени и част от формалния апарат и софтуерните

продукти, с чиято помощ е извършена оценката на равнището на безработица.

Ключови думи: малка подсъвкупност, община, равнище на безработица, експериментални резултати, оценяване, проблеми, софтуер.

JEL: C13.

Въведение

Необходимостта от коректно локализиране на политики на въздействие от различно естество изисква информационен ресурс с висока степен на териториална дезагрегация. Разглеждайки проблемите, касаещи пазара на труда и тяхното конкретно териториално изражение, се забелязва нуждата от детайлна статистическа информация относно равнището на безработица на общинско равнище. Това от своя страна е продиктувано най-вече от търсенето на подобен род информация във връзка с провеждането на политиката на Министерство на труда и социалната политика, както и реализирането на инвестиционни проекти от частния сектор (например

разширяване на банковата мрежа, откриване на нови магазини и други). Подробният анализ на съществуващата информационна осигуреност в Република България относно равнището на безработица за малки териториални подсъвкупности, каквито са общините, показва, че няма нито един надежден източник, който едновременно да отговаря на двете най-важни изисквания: добра методологична обоснованост и висока детайлност на данните [1]. Във връзка с това трябва да се отбележи, че последното преброяване на населението през 2001 година и Регистърът на безработните в бюрата по труда предоставят информация с висока степен на териториална дезагрегация, но пък притежават недобра методологична основа, тъй като и при двата източника понятието „безработен“ не е дефинирано коректно от гледна точка на изискванията на Международната организация на труда. Други недостатъци на регистъра са следните: той е създаден да обслужва единствено целите на самата Агенция по заетостта; разпоредбите, въз основа на които е организиран регистърът, са в пряка зависимост от промените в националното законодателство, засягащо заетостта; невъзможно е да бъдат обхванати лицата, които са безработни, но не са регистрирани в Бюрата по труда; налице е невъзможност за изчисляване на коефициента на безработица само с данни от Регистъра на безработните лица поради липса на информация за броя на заетите лица в него. За разлика от тези два източника, Наблюдението на работната (НРС) сила дефинира и изследва коректно явлението „безработица“, но поради ресурсната си ограниченост и специфичния модел на представителна извадка, то може да пре-

доставя информация на равнище област, но не и на по-ниско дезагрегационно ниво.

Единствена възможност към настоящия момент за преодоляване на тази информационна недостатъчност е използването на специфичната методология за оценка на характеристики за малки подсъвкупности¹. Тя представлява съвкупност от статистически методи за определяне на конкретната числова стойност на дадена характеристика за определена малка подсъвкупност², получена въз основа на данни от конкретно статистическо изследване. По своята същност те са модификация на пригубили популярност в статистическата литература, но адаптирани към специфичните цели на анализа методи. Повече информация за теоретичната обосновка на тези методи може да бъде намерена в [2].

Въпреки наличието в литературата на голям набор от методи за оценка на характеристики за малки подсъвкупности, на практика в българските условия за процеса на оценяване може да се използва малка част от тях. Това се дължи на недостатъчната информационна осигуреност, особено на ниво отделен индивид. Във връзка с това възниква необходимостта от адаптиране на част от съществуващите методи към конкретните информационни условия в Република България и съчетаване на разполагаемата информация от различните налични източници. Самият оценителен процес най-ясно може да се представи с помощта на схемата, показана на фигура 1.

От фигурата съвсем ясно се вижда, че единственият възможен вариант за осъществя-

¹ Small area estimation.

² Под малка подсъвкупност в теорията се разбира такава подсъвкупност, за която броят на единиците, попаднали в извадката, въз основа на които се извършва даден статистически анализ, е малък в сравнение с броя на единиците от същата подсъвкупност в генералната съвкупност [4]. Обикновено за малки подсъвкупности се считат тези, при които относителният дял на единиците, попаднали в извадката, за тази подсъвкупност е под 5 % от същата подсъвкупност на генералната съвкупност.

ване на оценителния процес е съчетаването на данни от „Наблюдението на работната сила“ и „Текущата демографска статистика“ на ниво община, използвайки директни оценки, оценки във вид на отношение, регресионни оценки и производните на тях комбинирани оценки.

На практика оценителният процес преминава през два основни етапа:

Първи етап – оценка на броя на заетите и броя на безработните по общини въз основа на някой от методите.

Втори етап – изчисляване на коефициента на безработица по общини въз основа на оцененния брой, използвайки на следната формула:

$$UR_d \% = \frac{UN_d}{EM_d + UN_d} \cdot 100, \quad (1)$$

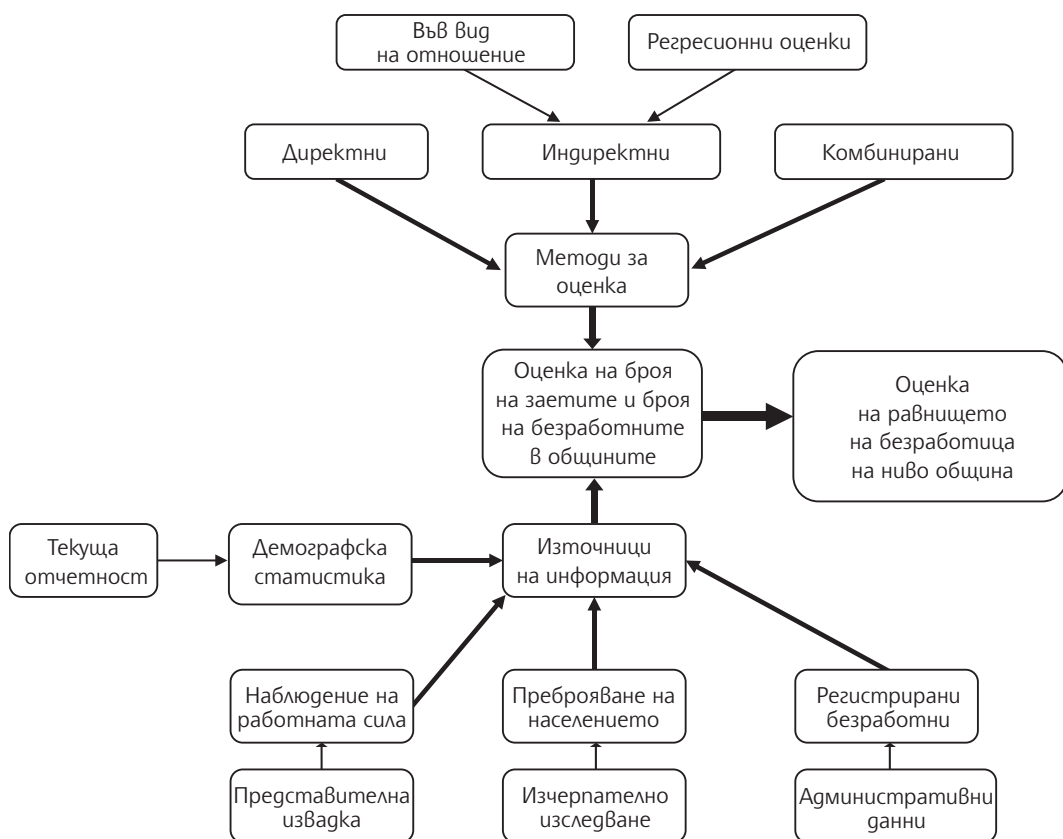
където:

UR_d е коефициентът на безработица за d -тата община, $d = 1, 2, \dots, k$;

UN_d – броят на безработните в d -тата община;

EM_d – броят на заетите в d -тата община;

k – броят общини на територията на Република България ($k = 264$).



Фигура 1. Методи за оценка и източници на информация при оценката на равнището на безработица по общини

В конкретния случай – при експерименталната оценка на равнището на безработица в общините на Република България, са използвани данни от „Наблюдение на работната сила“ през 2007 година и от „Демографска статистика“ – към 31.12.2006 г. Използването на информацията в такъв вид дава възможност да се избегне сезонността в явлението безработица, както и да се осигури относително голям брой случаи, въз основа на които да се осъществят оценителните процедури.

Общият брой на единиците (лица на 15 и повече навършени години), попаднали в наблюдението за 2007 година, е 117 458, а населението в този възрастов интервал възлиза на 6 647 375.

1. Резултати от експерименталното оценяване на равнището на безработица по общини

Преди разглеждането на конкретните резултати е необходимо да се обърне внимание на факта, че за общините Малко Търново, Маджарово, Трекляно, Априлци, Георги Дамяново, Копревщица, Антон, Чавдар, Алфатар, Гурково, Бяла, Макреш и Златарица в извадката не попада нито едно лице. Трябва да се отбележи също, че в 107 от общините броят на анкетираните лица е под 200 единици. Според последните две уточнения в тези общини ще възникнат затруднения при оценяването на броя на заетите и броя на безработните, тъй като оценяването на характеристики въз основа на по-малко от 200 единици в извадката често пъти води до подвеждащи изводи [5], а липсата на каквато и да било информация за част от общините ще доведе до невъзможност за получаване на оценки за всички общини при директния метод.

Експерименталните резултати, получени въз основа на горепосочените методи, съчетавайки данни от двата информационни източника, най-общо могат да бъдат представени в таблица 1.

Директната оценка за броя на заетите и броя на безработните по общини се получава след претегляне на единичните данни за лицата (от файла с данни на „Наблюдението на работната сила“) с индивидуалните постстратификационни тегла в зависимост от това, съответното лице в коя постстрата попада, използвайки следните формули:

- За броя на заетите:

$$EM_d^{dir} = \sum_{i=1}^p y_{id} w_{id} \quad , \quad (2)$$

където:

EM_d^{dir} е броят на заетите в d-тата община, $d = 1, 2, \dots, k$, получени в резултат на директно оценяване;

p – броят лица, попаднали в извадката от d-тата община;

y_{id} – дихотомната индикаторна величина, приемаща стойности 1 – ако лицето е заето, и 0 – ако то не е заето, за d-тата община;

w_{id} – теглото на i-тото лице в d-тата община, представляващо реципрочна стойност на вероятността за включване на това лице в извадката при излъчването на последната;

k – броят общини на територията на Република България ($k = 264$).

- За броя на безработните:

Начинът на изчисляване е аналогичен на този при броя на заетите и има следният вид:

$$UN_d^{dir} = \sum_{i=1}^p x_i w_i \quad , \quad (3)$$

Таблица 1. Резултати от приложените методи за оценка на равнището на безработица по общини

	Директна оценка		Регресионна оценка – обикновена		Комбинирана оценка – тип 1		Регресионна оценка – логнормална		Комбинирана оценка – тип 2		Регресионна с регионални елементи		Комбинирана оценка – тип 3		Бейсовска оценка		Комбинирана оценка – тип 4	
	Община	КБ (%)	Община	КБ (%)	Община	КБ (%)	Община	КБ (%)	Община	КБ (%)	Община	КБ (%)	Община	КБ (%)	Община	КБ (%)	Община	КБ (%)
Най-висока безработица	Община	52,0	Трекляно	66,1	Трекляно	66,1	Минерални бани	25,7	Никола Козлево	26,1	Ружинци	34,4	Каинарджа	34,1	Ружинци	44,5	Ново село	43,9
	Сливо поле	45,8	Бойница	57,9	Бойница	57,3	Гърмен	24,9	Минерални бани	24,6	Каинарджа	33,6	Ружинци	34,0	Ново село	43,6	Ново село	43,3
	Каолиново	34,5	Чавдар	55,9	Чавдар	55,9	Сатовча	24,2	Каинарджа	23,8	Ново село	30,7	Никола Козлево	31,3	Чупрене	43,3	Макреш	42,6
	Главиница	29,5	Макреш	52,4	Макреш	52,4	Руен	23,4	Гърмен	22,9	Борино	30,1	Ново село	30,5	Макреш	42,6	Каинарджа	41,7
	Копел	25,8	Ковачевци	51,8	Ковачевци	51,6	Каинарджа	22,4	Сатовча	22,3	Тунджа	29,3	Самуил	29,9	Каинарджа	41,8	Чупрене	41,3
	Исперих	25,4	Антон	48,4	Антон	48,4	Марица	22,2	Венец	21,8	Минерални бани	29,1	Борино	28,2	Бойница	41,7	Бойница	40,8
	Дебин	25,2	Маждарово	46,9	Маждарово	46,9	Никола Козлево	22,2	Самуил	21,7	Самуил	28,9	Черноочене	28,0	Георги Дамяново	39,7	Георги Дамяново	39,7
	Върбница	23,2	Челопеч	46,3	Челопеч	46,3	Черноочене	21,9	Добрич-сеска	21,7	Черноочене	28,9	Ситово	28,0	Ковачевци	39,6	Ковачевци	39,3
	Добрич-сеска	22,8	Чупрене	46,6	Чупрене	45,7	Аврен	21,7	Ружинци	21,4	Чупрене	28,6	Минерални бани	27,7	Борино	39,2	Невестино	37,4
	Разраг	22,5	Георги Дамяново	43,2	Георги Дамяново	43,2	Добрич-сеска	21,6	Черноочене	21,2	Якимово	28,3	Венец	27,3	Якимово	38,1	Борино	37,1
	Дончи-цирик	1,4	Костендрил	5,8	Костендрил	5,8	Плевен	5,3	Кюстендил	5,4	Първомай	4,1	Първомай	4,1	Плевен	3,3	Плевен	3,3
	Гоце Делчев	1,2	Русе	5,4	Русе	5,4	Стара Загора	5,3	Плевен	5,3	Сопот	4,1	Сопот	4,0	Казанлык	3,1	Русе	3,2
Най-ниска безработица	Дона Митрополит	1,1	Плевен	5,3	Плевен	5,3	Чипровци	5,2	Стара Загора	5,3	Казанлык	3,9	Казанлык	3,9	Русе	3,1	Казанлык	3,1
	Белово	1,1	Плевен	5,3	Плевен	5,3	Плевен	5,2	Перник	5,2	Карлово	3,4	Карлово	3,4	Карлово	2,8	Карлово	2,8
	Пещера	1,0	Габрово	5,2	Габрово	5,2	Бургас	5,1	Перник	5,1	Бургас	3,2	Стара Загора	3,1	Асеновград	2,5	Асеновград	2,5
	Карлово	1,0	Перник	5,2	Перник	5,2	Перник	5,1	Бургас	5,1	Бургас	3,1	Асеновград	3,1	Стара Загора	2,2	Стара Загора	2,3
	Ехово	0,9	Перник	5,1	Габрово	5,2	Габрово	5,0	Троян	5,0	Пловдив	2,9	Бургас	2,9	Бургас	1,9	Бургас	1,9
	Първомай	0,9	Бургас	5,1	Бургас	5,1	Троян	5,0	Пловдив	5,0	Габрово	2,3	Варна	2,4	Варна	1,4	Варна	1,5
	Чирпан	0,9	Пловдив	5,0	Пловдив	5,0	Пловдив	5,0	Габрово	5,0	Варна	2,3	Пловдив	2,3	Пловдив	1,3	Пловдив	1,3
	Севилиево	0,8	Столична	4,6	Варна	4,6	Варна	4,5	Варна	4,4	Пловдив	2,3	Пловдив	2,3	Столична	0,4	Столична	0,4
	Раковски	0,4	Варна	4,4	Варна	4,4	Столична	4,2	Столична	4,2	Столична	0,9	Столична	0,9	Столична	0,4	Столична	0,4

Източник: Изчисления на автора на база данни на НСИ

където:

повечето символи са известни с изключение на: UN_d^{dir} е броят на безработните в d -тата община, $d = 1, 2, \dots, k$, получени в резултат на директно оценяване;

x_{id} – дихотомната индикаторна величина, приемаща стойности 1 – ако лицето е безработно, и 0 – ако то не е безработно, за d -тата община.

При разглеждане на резултатите впечатление прави липсата на данни за равнището на безработица в 129 от 264 или около 50 % от общините. В 107 общини има по-малко от 200 единици, а в останалите случаи това се дължи на липса на данни за безработни лица в изследваните общини (9 на брой) и отсъствие на каквато и да било информация за работната сила поради невключване на тези общини в извадката на Наблюдението (13 на брой). От таблица 1 се вижда, че с най-висока безработица са Сливо поле (52,0 %), Каолиново (45,8 %), Главиница (34,5 %), Котел (29,5 %), Исперих (25,8 %), Девин (25,4 %), Върбица (25,2 %), Добрич-селска (23,2 %), Разград (22,8 %) и Долни чифлик (22,5 %). С най-ниско ниво на безработица са Гоце Делчев (1,4 %), Долна Митрополия (1,2 %), Белово (1,1 %), Пещера (1,1 %), Карлово (1 %), Елхово (1 %), Първомай (0,9 %), Чирпан (0,9 %), Севлиево (0,8 %) и Раковски (0,4 %).

Докато резултатите за най-високо равнище на безработица изглеждат реалистични, тези за най-ниско сочат наличието на някаква неточност, тъй като обикновено в практиката се оказва, че с най-ниско равнище на безработица са общините на големите градове. Това разминаване най-вероятно се дължи на сравнително малкия обем на извадката в тези общини и относително малкия брой лица, идентифицирани като безработни в момента на провеждане на изследването. Също така една възможна причина за прекалено ниската безработица може да се окаже и генерирането

на нереалистични, от гледна точка на оценката на ниво община, постстратификационни тегла. Те от своя страна изкривяват резултатите поради факта, че при наличие на малко на брой единици в дадена постстрата, всяка една от тях получава относително голямо тегло, което впоследствие повлиява на цялостната оценка на безработицата в общината, в която са попаднали тези единици. Често пъти директните оценки имат само ориентировъчен характер, що се отнася до оценки за малки подсъвкупности, затова в практика се използват индиректни по своята същност методи.

Комбинирайки данните от „Демографската статистика“ относно броя на лицата на 15 и повече навършени години по общини и общо за страната и броя на заетите общо за страната от НРС, получаваме броя на заетите по общини. По същия начин комбинираме данните с тези, отнасящи се за броя на безработните лица и получаваме броя на безработните по общини. Формално **оценката във формата на отношение – общо за страната**, става по следния начин:

- За броя на заетите:

$$EM_d^{ratio1} = \frac{N_d}{N} \cdot EM, \quad (4)$$

където:

EM_d^{ratio1} е броят на заетите в d -тата община, $d = 1, 2, \dots, k$, получени в резултат на оценяване във формата на отношение, общо за страната;

N_d – броят лица на 15 и повече навършени години в община d , определен въз основа на данни от „Демографската статистика“;

N – общият брой лица на 15 и повече навършени години за страната като цяло, определен въз основа на данни от „Демографската статистика“;

EM – броят на заетите лица за страната като цяло, определен въз основа на данни от НРС.

- За броя на безработните:

$$UN_d^{ratio} = \frac{N_d}{N} \cdot UN, \quad (5)$$

където:

UN_d^{ratio} е броят на безработните в d-тата община, $d = 1, 2, \dots, k$, получени в резултат на оценяване във формата на отношение, общо за страната;

UN – броят на безработните лица за страната като цяло, определен въз основа на данни от НРС.

Най-общо резултатите от направените изчисления са представени в таблица 2.

Прави впечатление, че въпреки реалистичните резултати относно абсолютния брой на заетите и безработните по общини, коефициентът на безработица е еднакъв за всички общини, независимо от размера им, и възлиза на 6,9 %. Това се дължи на следното:

$$\begin{aligned} UR_d^{ratio} &= \frac{UN_d^{ratio}}{EM_d^{ratio} + UN_d^{ratio}} \cdot 100 = \\ &= \frac{\frac{N_d}{N} \cdot UN}{\frac{N_d}{N} \cdot EM + \frac{N_d}{N} \cdot UN} \cdot 100 = \\ &= \frac{\frac{N_d}{N} \cdot UN}{\frac{N_d}{N} \cdot (EM + UN)} \cdot 100 = \frac{UN}{(EM + UN)} \cdot 100 = UR \end{aligned} \quad (6)$$

Таблица 2. Брой заети, брой безработни и коефициент на безработица, изчислени въз основа на оценка във вид на отношение – общо за страната

Община	Брой заети	Брой безработни	Коефициент на безработица (%)
Банско	5557,1	410,4	6,9
Белица	3809,3	281,3	6,9
Благоевград	32565,3	2405,2	6,9
Гоце Делчев	13043,5	963,3	6,9
Гърмен	5733,2	423,4	6,9
Кресна	2482,7	183,4	6,9
Петрич	23487,2	1734,7	6,9
Разлог	9029,7	666,9	6,9
Сандански	17764,8	1312,0	6,9
Сатовча	7014,2	518,0	6,9
...	...	...	...
Никола Козлево	2637,4	194,8	6,9
Нови пазар	7846,0	579,5	6,9
Смядово	3302,3	243,9	6,9
Хитрино	2860,0	211,2	6,9
Шумен	43217,1	3191,9	6,9
Болярово	2067,8	152,7	6,9
Елхово	7652,3	565,2	6,9
Стралджа	5910,8	436,6	6,9
Тунджа	12414,2	916,9	6,9
Ямбол	33024,8	2439,1	6,9

Източник: Изчисления на автора на база данни на НСИ.

Тоест коефициентите на безработица по общини са равни на коефициента на безработица общо за страната за 2007 година. Това от своя страна е основен недостатък на този метод, тъй като той дава добри резултати за броя на заетите и безработните по общини, но коефициентът на безработица остава един и същ т.е. не е възможно да се извърши сравнение на общините от гледна точка на равнището на безработица в тях, следователно не може да се вземе и адекватно управленско решение въз основа на тези резултати.

По аналогичен начин се получават и оценки, когато се използват данните не на ниво общо за страната, **а на ниво област**. Тогава оценките придобиват вида:

- За броя на заетите :

$$EM_d^{ratio2} = \frac{N_{ds}}{N_s} \cdot EM_s, \quad (7)$$

където:

EM_d^{ratio2} е броят на заетите в d-тата община, $d = 1, 2, \dots, k$, получени в резултат на оценяване във формата на отношение, по области;

Таблица 3. Брой заети, брой безработни и коефициент на безработица, изчислени въз основа на оценка във вид на отношение – по области

Община	Брой заети	Брой безработни	Коефициент на безработица (%)
Банско	6403,2	152,7	2,3
Белица	4389,3	104,7	2,3
Благоевград	37523,9	894,8	2,3
Гоце Делчев	15029,5	358,4	2,3
Гърмен	6606,2	157,5	2,3
Кресна	2860,8	68,2	2,3
Петрич	27063,5	645,4	2,3
Разлог	10404,6	248,1	2,3
Сандански	20469,7	488,1	2,3
Сатовча	8082,2	192,7	2,3
...	...	...	...
Никола Козлево	2424,0	512,2	17,4
Нови пазар	7211,3	1523,9	17,4
Смядово	3035,2	641,4	17,4
Хитрино	2628,6	555,5	17,4
Шумен	39721,0	8393,8	17,4
Болярово	2142,6	165,2	7,2
Елхово	7928,9	611,3	7,2
Стралджа	6124,5	472,2	7,2
Тунджа	12863,0	991,8	7,2
Ямбол	34218,8	2638,3	7,2

Източник: Изчисления на автора на база данни на НСИ.

N_{ds} – броят лица на 15 и повече навършени години в община d , принадлежаща на s -тата област, определен въз основа на данни от „Демографската статистика“;

N_s – броят лица на 15 и повече навършени години в s -тата област, определен въз основа на данни от „Демографската статистика“;

EM_s – броят заети лица в s -тата област, определен въз основа на данни от „Наблюдението на работната сила“.

- За броя на безработните:

$$UN_d^{ratio2} = \frac{N_d}{N_s} \cdot UN_s, \quad (8)$$

където:

UN_d^{ratio2} е броят на безработните в d -тата община, $d = 1, 2, \dots, k$, получени в резултат на оценяване във формата на отношение, по области;

UN_s – броят безработни лица в s -тата област, определен въз основа на данни от „Наблюдението на работната сила“.

След изчисляване на броя на заетите, броя на безработните и равнището на безработица по общини се получават следните резултати (таблица 3).

Графичният облик на резултатите е показан на фигура 2.

Това, което и тук прави впечатление, е, че така оценените коефициенти на безработица възпроизвеждат същите от по-високо териториално ниво – област. Това се дължи на следното:

$$\begin{aligned} UR_d^{ratio2} &= \frac{UN_d^{ratio2}}{EM_d^{ratio2} + UN_d^{ratio2}} \cdot 100 = \\ &= \frac{\frac{N_{ds}}{N_s} \cdot UN_s}{\frac{N_{ds}}{N_s} \cdot EM_s + \frac{N_{ds}}{N_s} \cdot UN_s} \cdot 100 = \\ &= \frac{\frac{N_{ds}}{N_s} \cdot UN_s}{\frac{N_{ds}}{N_s} \cdot (EM_s + UN_s)} \cdot 100 = \frac{UN_s}{(EM_s + UN_s)} \cdot 100 = UR_s \end{aligned} \quad (9)$$

За разлика от предходните методи, както се вижда от таблица 1, **регресионните оценки** дават съвсем реалистични резултати. Тук обаче трябва да се отбележи, че при този тип оценки, с цел получаване на достоверни и стабилни стойности на параметрите на регресионните модели, а оттам и правдоподобни оценки за броя на заетите



Фигура 2. Равнище на безработица по общини въз основа на оценка във вид на отношение тип 2

и безработните, относно изходните данни (данните от „Наблюдение на работната сила“), трябва да бъдат взети под внимание следните *особености*:

- За оценка на регресионните коефициенти са използвани данни за общините, които отговарят от своя страна на следните критерии:
 - броят на единиците от съответната община, попаднали в извадката, да не е по-малък от 200 единици;
 - да няма липсващи стойности за която и да е от използваните променливи;
 - да няма стойности на променливите, които екстремално да се различават от стойностите на останалите общини (т.нар. outliers).
- За реалната оценка на броя на заетите и броя на безработните са използвани само онези модели, които отговарят на изискванията на МНМК, т.е. моделите, които са адекватни, имат статистически значими регресионни коефициенти, липсва мултиколинеарност при факторните променливи, остатъците са нормално разпределени и R^2 е достатъчно висок³.

На първото условие се оказва, че отговарят 114 от общо 264 общини, въз основа на данните, от които са направени всички регресионни оценки. Относно второто условие, по-голу са представени само тези модели, които отговарят на посочените изисквания, а всички останали тествани модели излизат извън обхвата на настоящото изложение.

При изчисляването на оценките относно броя на заетите и броя на безработните по общини, въз основа на регресионните методи, са предприети следните стъпки:

- Съставяне на файл с данни от „Наблюдение на работната сила“, който включва всички променливи, необходими за провеждане на регресионния анализ за избраните 114 общини.
- Провеждане на стъпков регресионен анализ [3], при който се „отсяват“ незначимите факторни променливи от тези, които са се оказали значими, въз основа на т.нар. частен F-критерий, поотделно за броя на заетите и броя на безработните.
- Проверка за изпълнението на изискванията на регресионния анализ.
- Съставяне на регресионните уравнения за заетите и безработните с вече оценените параметри.
- Заместване на стойностите на факторните променливи с данните от текущата демографска статистика и получаване на оценки за броя на заетите и броя на безработните за всичките 264 общини.
- Изчисляване на равнището на безработица по общини въз основа на оценен брой заети и безработни.

Въз основа на разполагаемата информация, използвайки формалния апарат, отнасящ се до **обикновената линейна регресионна оценка**, и следвайки гореизложените стъпки, получаваме следния регресионен модел за броя на заетите и определящите го фактори:

$$\hat{E}M_d^{regr} = 0,826X_{1d} + 0,153X_{2d} + 0,193X_{3d}, \quad (10)$$

(0,065) (0,024) (0,085)

където:

- X_1 – броят на населението на възраст 25-49 г.;
- X_2 – общият брой на градското население;
- X_3 – броят мъже на възраст 50 и повече години⁴.

Резултатите относно регресията, засягаща броя на безработните, имат следния вид:

³ За оценителните процедури е използван софтуерният продукт SPSS Версия 13.

⁴ Числата, посочени в скоби, са стандартните грешки на съответните коефициенти.

$$\hat{Y}_d^{\text{regr}} = 480,584 + 0,330X_{1d}, \quad (11)$$

(121,589) (0,030)

където:

X_1 е броят на жените на възраст 15-24 г.

При съчетаването на данните от двата източника на информация (регресионен модел от „Наблюдение на работната сила“ и стойност на факторите от „Демографска статистика“) получаваме стойностите относно броя на заетите и броя на безработните по общини. Въз основа на този брой изчисляваме и коефициента на безработица по общини, визуално представен на следната картограма (Виж фигура 3).

От таблица 1 се вижда, че общините с най-високо равнище на безработица са: Трекляно (66,1 %), Бойница (57,9 %), Чавдар (55,9 %), Макреш (52,4 %), Ковачевци (51,8 %), Антон (48,4 %), Маджарово (46,9 %), Челопеч (46,8 %), Чупрене (46,6 %) и Георги Дамяново (43,2 %). Общините Варна (4,4 %), Столична (4,6 %), Пловдив (5 %), Бургас (5,1 %), Перник (5,1 %), Стара Загора

(5,2 %), Габрово (5,2 %), Плевен (5,3 %), Русе (5,4 %) и Кюстендил (5,8 %) са с най-ниска безработица. Така разгледаните резултати изглеждат достоверни, тъй като е логично в малките селски общини равнището на безработица да бъде високо, а в големите градски общини – ниско.

От таблица 1 също така се вижда, че е възможно изчисляването на т.нар. комбинирана оценка. Това най-общо става във вида:

$$Y_d^{\text{comp}} = \gamma \cdot Y_d^{\text{dir}} + (1 - \gamma) \cdot Y_d^{\text{regr}} \quad (12)$$

където:

Y_d^{comp} е самата комбинирана оценка (например брой заети) за d-тата малка подсвкупност (например община), изчислена като средна претеглена величина от две други оценки;

Y_d^{dir} – директната оценка за d-тата малка подсвкупност;

Y_d^{regr} – регресионната оценка за d-тата малка подсвкупност;

γ – подходящото избрано тегло, най-често в практиката изчислявано чрез формулата:



Фигура 3. Равнище на безработица по общини въз основа на обикновена линейна регресионна оценка⁵

⁵ За картографирането на резултатите е използван продуктът MapInfo, Версия 9.5.

$$\gamma_d = \frac{n_d}{N_d}, \quad (13)$$

където:

n_d е броят единици от съответната малка подсвкупност, попаднали в извадката;

N_d – общият брой единици в същата малка подсвкупност на генералната съвкупност.

Разглеждайки резултатите относно **комбинираните оценки**, това, което прави силно впечатление, е, че те се различават съвсем незначително от оценките, към които се отнасят. Така например при регресионните оценки за община Бойница нивото спада от 57,9 % на 57,3 %, а за община Перник от 5,1 % на 5,2 %. Това се дължи на факта, че броят на единиците, попаднали в извадката от дадена община, е прекалено малък, като в повечето случаи делът на тези единици от общия им брой в генералната съвкупност е под 3 %.

Известно подобрене в оценителния процес внасят **оценките от логлинеен тип**, като при тях чрез логаритмуване се преодолява голямото вариране в изходните данни. Конкретните логлинейни регресионни уравнения имат следния вид:

- за заетите:

$$\ln(\hat{E}M_d^{\ln, \text{reg}}) = \underset{(0,208)}{0,425} + \underset{(0,068)}{0,735 \ln(X_{1d})} + \underset{(0,042)}{+0,156 \ln(X_{2d})} + \underset{(0,039)}{+0,101 \ln(X_{3d})}, \quad (14)$$

където:

X_1 е броят на населението на Възраст 25-49 г.;

X_2 – общият брой на жени в градовете;

X_3 – броят мъже на Възраст 15-24 години.

- за безработните:

$$\ln(\hat{U}N_d^{\ln, \text{reg}}) = \underset{(0,011)}{0,893 \ln(X_{1d})} \quad (15)$$

където:



Фигура 4. Равнище на безработица по общини въз основа на логлинейна регресионна оценка

X_1 е броят на жените на възраст 15-24 г.

Следвайки вече познатите стъпки, получаваме следните резултати (Виж фигура 4).

От фигура 4 и таблица 1 ясно се вижда, че с най-високи стойности на коефициента на безработица са: Минерални бани (25,7 %), Гърмен (24,9 %), Сатовча (24,2 %), Руен (23,4 %), Кайнарджа (22,4 %), Марица (22,2 %), Никола Козлево (22,2 %), Черноочене (21,9 %), Аврен (21,7 %) и Добрич-селска (21,6 %). Най-малко засегнати от безработицата са общините Столична (4,2 %), Варна (4,5 %), Пловдив (5 %), Троян (5 %), Габрово (5 %), Перник (5,1 %), Бургас (5,1 %), Чипровци (5,2 %), Стара Загора (5,3 %) и Плевен (5,3 %).

Допълнително подобрение в оценките на равнището на безработица в общините внася добавянето на регионални променливи в регресионните уравнения, като по този начин се отчита влиянието на характерните особености на различните области, в които се намират съответните общини.

В този случай оценката на параметрите на регресионните модели се извършва по два начина: чрез обикновения метод на най-малките квадрати и байесовски йерархичен подход. Характерното при оценките от байесовски тип е възможността за използване на допълнителна априорна информация в процеса на оценяване⁶. Това от своя страна обаче изисква и прилагането на сложни итеративни процедури за оценка на търсените параметри на регресионните уравнения⁷.

Въз основа на разполагаемата информация и подходящия за оценките на параметрите софтуер се получават следните регресионни уравнения:

- за броя на заетите:

$$\ln(\hat{E}M_d^{\text{bayes}}) = \underset{(0,284)}{0,330} + \underset{(0,068)}{0,738 \ln(X_{1d})} + \underset{(0,042)}{0,155 \ln(X_{2d})} + \underset{(0,040)}{0,100 \ln(X_{3d})} + \underset{(0,028)}{0,131 \ln(U_d)} \quad (16)$$

където:

X_1 е броят на населението на възраст 25-49 г.;



Фигура 5. Равнище на безработица по общини въз основа на логлинейна регресия с регионални променливи

⁶ В случая са използвани предварителни данни за разпределението на параметрите на регресионното уравнение.

⁷ Известно облекчение в оценителните процедури внася използването на софтуерния продукт WinBUGS.

X_2 – общият брой на жени в градовете;
 X_3 – броят мъже на възраст 15-24 г.;
 U_i – размерът на БВП в млн. лв. в областта, към която принадлежи съответната община.

- за броя на безработните:

$$\ln(\hat{U}_d^{\text{bayes}}) = 5,583 + 0,614\ln(X_{1d}) - 0,550\ln(U_d), \quad (17)$$

(1,293) (0,094) (0,161)

където:

X_1 – броят жени на възраст 15-25 г.;
 U_i – размерът на БВП в млн. лв. в областта, към която принадлежи съответната община.

От гледна точка на получените резултати относно равнището на безработица от двата подхода не се забелязва съществена разлика (виж таблица 1), а графично данните могат да бъдат представени по следния начин, показан на фигура 5.

Съдейки по резултатите, поместени в таблица 1 и фигура 5, общините с най-висока безработица са Ружинци (44,5 %), Ново село (43,6 %), Чупрене (43,3 %), Макреш (42,6 %), Кайнарджа (41,8 %), Бойница (41,7 %), Георги Дамяново (39,7 %), Ковачевци (39,6 %), Борино (39,2 %) и Якимово (38,1 %), а тези с най-ниска – Столична (0,4 %), Пловдив (1,3 %), Варна (1,4 %), Бургас (1,9 %), Стара Загора (2,2 %), Асеновград (2,5 %), Карлово (2,8 %), Русе (3,1 %), Казанлък (3,1 %) и Плевен (3,3 %). Това, което прави впечатление тук, е сравнително големият размах между най-ниската (0,4 %) и най-високата оценка (44,5 %), възлизащ на 44,1 %. Въпреки това резултатите имат логичен облик и потвърждават направените до момента съждения за равнището на безработица по общини.

Заклучение

Приложението на методите, засягащи оценката на равнището на безработица

по общини, на практика е извършено, съчетавайки, по уникален за българските условия начин, данни от „Наблюдение на работна сила“ и „Демографска статистика“ на ниво община. Специалното комбиниране на три софтуерни продукта в процеса на работа дава практическа възможност за реализиране на оценяването, преодолявайки проблема с отсъствието на цялостно изграден програмен продукт за приложение на специфичната методология на теорията за оценка на характеристики за малки подсъвкупности.

Обръщайки внимание на конкретните резултати от процеса на оценяване, трябва да споменем, че като цяло има известно различие между тях, вероятно дължащо се на използването на различни по своята същност методи. Въпреки това оценките за броя на заетите, броя на безработните и произтичащия от тях коефициент на безработица са коректни и могат да бъдат използвани в процеса на вземане на управленски решения. Разглеждайки резултатите в обобщен вид, можем да заключим, че с най-високо равнище на безработица се откроя общините Никола Козлево, Кайнарджа, Минерални бани, Ново село, Якимово, Борино и Черноочене. Най-малко „засегнати“ от безработицата се оказаха столицата и големите градове – Пловдив, Бургас, Варна, Стара Загора и Плевен.

Една част от методите обаче дадоха някои нелогични резултати, като например директната оценка и оценките във вид на отношение. Причините се коренят във факта, че при директната оценка голяма част от оценките за общините (над 50 %) се базират на ограничен брой случаи, попаднали в извадката и същевременно на неадекватно генерирани постстратификационни тегла.

При оценките във вид на отношение се оказва, че нереалистичността се дължи на

факта, че по своята същност тези оценки възпроизвеждат стойностите на равнището на безработица на ниво области и общо за страната.

От гледна точка на регресионното оценяване като основни фактори, оказващи влияние върху оценявания брой заети, могат да се посочат броят на населението на възраст 25-49 г., общият брой на градското население и броят мъже на възраст 50 и повече години. Относно броя на безработните такива се оказват броят на жените на 15-24 години и брутният вътрешен продукт на равнище област.

В заключение може да се обобщи, че използвайки методите за оценка на малки подсъвкупности, се решава един много важен проблем, а именно липсата на надеждна информация за равнището на безработица на ниво община. Адаптирайки тази специфична методология съобразно конкретните български информационни условия, не само се преодоляват затрудненията, свързани с информационната осигуреност и оценителните процедури, но се създава и един ефективен инструмент, с чиято помощ биха могли да се вземат адекватни и коректни управленски решения. Поради универсалността му приложеният инструментариум може да бъде използван не само

при анализиране състоянието и промените в пазара на труда, но и процесите, засягащи останалите сфери на социално-икономическата действителност.

Литература

1. Найденов, Ал., Проблеми при статистическата оценка на безработицата в България за малки подсъвкупности, Сборник доклади от международна научна конференция „Статистиката и предизвикателствата на информационното общество през XXI век“, Рава, септември 2006 г., с. 191 – 197.
2. Найденов, Ал., Регионални аспекти в статистическото изследване на безработицата в България – настояще и бъдеще, сп. „Статистика“, С., бр. 1/2006, с. 51 – 64.
3. Съйкова, ИВ., А. Стойкова – Къналиева, Св. Съйкова, Статистическо изследване на връзки и зависимости, УИ „Стопанство“, С., 2002.
4. Ghosh, M., and J. N. K. Rao, Small area estimation: An appraisal, 1994.
5. Nicholas, T. L., Missing Data and Small-Area Estimation, 2005. **ИИ**