Using the Small Area Estimation Methods for the Local Unemployment Rates Identification

Assist. Prof. Alexandar Naidenov
UNWE, Statistics and Econometrics Department
e-mail: anaidenov@gmail.com

Summary: The article presents the results from the experimental estimation of municipality unemployment rates in Bulgaria, using the small area estimation methods. The need for this kind of estimates is based on the real needs for highly territorially disaggregated data in order to provide information for the correctly localized policies. The basic concepts concerning the specific methodology are clarified in the text and a generalized theoretical model for analysis of the phenomenon is presented too. A short review of the information sources is also available. The results from the realized survey are presented in graphics and tables for every method that is used and a short critical review is done in addition too. The formulas and the software products are also described in order to show the practical application of the methodology and calculation of the unemployment rate estimates.

Key words: Small area, municipality, unemployment rate, experimental results, estimation, issues, software.

JEL: C13.

Introduction

The need for a correct localization of the different types of policies leads to the revolving process of searching for highly territorially disaggregated data. Considering the labour market issues the need for detailed statistical information for the municipality unemployment rates arises. This is also based on the need for that kind of information for the realization of the Ministry of Labour and Social Policy activities and private investment projects. If we take closer look at the existing information about small area unemployment rates in Bulgaria, we could see that there are no reliable ones, which can provide both good methodology and highly detailed data [1]. Here it is important to mention that the last Census 2001 and the Registered unemployed persons are sources of information that could provide highly territorially disaggregated data, but they are not well methodologically grounded because the term 'unemployed' is not clearly defined in the both information sources (from the International Labour Organization requirements point of view). The other drawbacks of the Register are: it is created to serve the Register needs only; the register methodology is directly connected to the national employment legislation changes; it is not possible to include data for the non-registered unemployed persons; the unemployment rate cannot be calculated because of the missing data for the employed persons. In contrast to the mentioned information sources, the Labour Force Survey of the National Statistical

Institute of Bulgaria defines and measures the unemployment phenomenon correctly, but because of the resources restrictions and specific sample design it can provide data at higher aggregation levels e.g. districts.

At this moment the only possibility for the information lack overcoming is to use the specific methodology of small area[1] estimation. The latter consists of different kind of methods for small area characteristics estimation using data from a real survey. Most of the methods are already known from the statistical theory point of view ones, but adapted for the small area estimation purposes. More info can be found at [2].

Although lots of small area estimation methods are described in the statistical literature, at the concrete Bulgarian circumstances a small number of them can be used. This is caused by the insufficient information especially at low territorial level. Considering the latter, a need for adaptation of the existing methods to the specific needs have to be done and a combination of the different information sources could be reasonable too. The estimation process can be presented as follows:

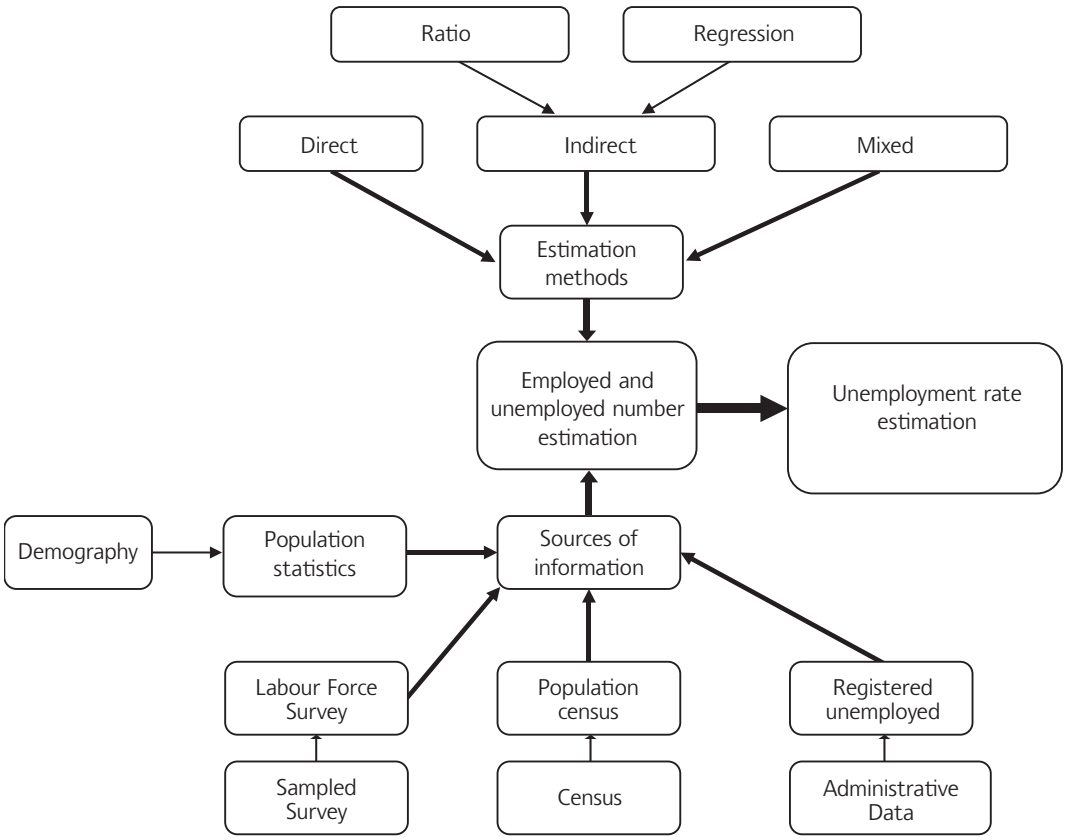

Figure 1. Methods and sources information for estimation of municipalities unemployment rates

[1] In theory 'small area' means an area for which the number of sampled elements, that are used for later analysis, is small comparing to the number of elements in the same population area[4]. Usually the areas are considered as small if the ratio sample/population is about 5 %.

As it is shown at the Figure 1, the only possible way to estimate the unemployment rates is to combine data from Labour Force Survey and the Population Statistics at municipality level, using direct, ratio, regression and mixed estimation methods.

Estimation process takes the following steps:

First – Employed and unemployed persons number estimation using the above mentioned methods;

Second – Municipalities unemployment rates calculation, using the estimated number of employed and unemployed persons by the formula:

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

where:
UR_d is unemployment rate for municipality d, $d = 1, 2, ..., k$;
UN_d – number of unemployed persons in municipality d;
EM_d – number of employed persons in municipality d;
k – number of municipalities in Bulgaria ($k = 264$).

In the present estimation process of the Bulgarian municipality unemployment rates, the data from the Labour Force Survey 2007 and the Population Statistics (31.12.2006) are used. By using annual data we can neglect the seasonal factor and we can provide enough number of sample elements in order to make proper estimation procedures.

The total number of sampled persons in the LFS for 2007 is 117 458 (15 years and over), and the Bulgarian population at the same age is 6 647 375.

1. Municipality Unemployment Rates Estimation – Experimental Results

Before we examine the concrete results we must emphasize on the fact that for the municipalities Malko Tarnovo, Madzharovo, Treklyano, Apriltci, Georgi Damyanovo, Koprivshtica, Anton, Chavdar, Alfatar, Gurkovo, Byala, Makresh and Zlataritca there are no data in the LFS sample. Also in 107 of 264 municipalities, the number of interviewed persons is less than 200, which sometimes leads to misleading results. These two facts generate difficulties in the employed and unemployed number estimation especially for the municipalities containing no data at all.

The results for experimental estimation, using the abovementioned methods and the information sources, can be presented in the following table:

The **direct estimates** of the employed and unemployed persons are calculated by weighting the individual data for the persons (Labour Force Survey data file) with the corresponding post-stratification weights, depending on the person's post-strata belonging, using the following formulas:

- For employed:

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

where:
EM_d^{dir} is a number of employed persons in municipality d, $d = 1, 2, ..., k$, using the direct estimation;
p – number of persons in the sample in municipality d;
y_{id} – dichotomy variable which takes value: 1 – if the person is employed and 0 – if the person is not employed, for municipality d;

Table 1. Results from the applied small area estimation of municipality unemployment rates

	Direct esimate		Regression estimate – linear		Mixed estimate – type 1		Regression estimate – loglinear		Mixed estimate – type 2		Regression estimate – with regional vars		Mixed estimate – type 3		Bayesian estimate		Mixed estimate – type 4	
	Municipality	UR (%)	Municipality	UR (%)	Municipality	UR (%)	Municipality	UR (%)	Municipality	UR (%)	Municipality	UR (%)	Municipality	UR (%)	Municipality	UR (%)	Municipality	UR (%)
Highest unempl.	Slivo pole	52.0	Treklyano	66.1	Treklyano	66.1	Mineralni bani	25.7	Nikola Kozlevo	26.1	Ruzhintci	34.4	Kaynardzha	34.1	Ruzhintci	44.5	Ruzhintci	43.9
	Kaolinovo	45.8	Boynitca	57.9	Boynitca	57.3	Garmen	24.9	Mineralni bani	24.6	Kaynardzha	33.6	Ruzhintci	34.0	Novo selo	43.6	Novo selo	43.3
	Glavinica	34.5	Chavdar	55.9	Chavdar	55.9	Svatovcha	24.2	Kaynardzha	23.8	Novo selo	30.7	Никола Козлево	31.3	Chuprene	43.3	Makresh	42.6
	Kotel	29.5	Makresh	52.4	Makresh	52.4	Ruen	23.4	Garmen	22.9	Borino	30.1	Novo selo	30.5	Makresh	42.6	Kaynardzha	41.7
	Isperih	25.8	Kovachevtci	51.8	Kovachevtci	51.6	Kaynardzha	22.4	Svatovcha	22.3	Tundzha	29.3	Samuil	29.9	Kaynardzha	41.8	Chuprene	41.3
	Devin	25.4	Anton	48.4	Anton	48.4	Maritca	22.2	Venetc	21.8	Mineralni bani	29.1	Borino	28.2	Boynitca	41.7	Boynitca	40.8
	Varbica	25.2	Madzharovo	46.9	Madzharovo	46.9	Nikola Kozlevo	22.2	Samuil	21.7	Samuil	28.9	Cherno-ochene	28.0	Georgi Damyanovo	39.7	Georgi Damyanovo	39.7
	Dobrich – rural	23.2	Chelopech	46.8	Chelopech	46.3	Cherno-ochene	21.9	Dobrich – rural	21.7	Cherno-ochene	28.9	Sitovo	28.0	Kovachevtci	39.6	Kovachevtci	39.3
	Razgrad	22.8	Chuprene	46.6	Chuprene	45.7	Avren	21.7	Ruzhintci	21.4	Chuprene	28.6	Mineralni bani	27.7	Borino	39.2	Nevestino	37.4
	Dolni chiflik	22.5	Georgi Damyanovo	43.2	Georgi Damyanovo	43.2	Dobrich – rural	21.6	Cherno-ochene	21.2	Yakimovo	28.3	Venetc	27.3	Yakimovo	38.1	Borino	37.1
Lowest unempl.	Gotce Delchev	1.4	Kyustendil	5.8	Kyustendil	5.8	Pleven	5.3	Kyustendil	5.4	Parvomay	4.1	Parvomay	4.1	Pleven	3.3	Pleven	3.3
	Dolna Mitropoliya	1.2	Ruse	5.4	Ruse	5.4	Stara Zagora	5.3	Pleven	5.3	Sopot	4.1	Sopot	4.0	Kazanlak	3.1	Ruse	3.2
	Belovo	1.1	Pleven	5.3	Pleven	5.3	Chiprovtci	5.2	Stara Zagora	5.3	Kazanlak	3.9	Kazanlak	3.9	Ruse	3.1	Kazanlak	3.1
	Peshtera	1.1	Gabrovo	5.2	Stara Zagora	5.2	Burgas	5.1	Pernik	5.2	Karlovo	3.4	Karlovo	3.4	Karlovo	2.8	Karlovo	2.8
	Karlovo	1.0	Stara Zagora	5.2	Pernik	5.2	Pernik	5.1	Burgas	5.1	Stara Zagora	3.2	Stara Zagora	3.3	Asenovgrad	2.5	Asenovgrad	2.5
	Elhovo	1.0	Pernik	5.1	Gabrovo	5.2	Gabrovo	5.0	Troyan	5.0	Asenovgrad	3.1	Asenovgrad	3.1	Stara Zagora	2.2	Stara Zagora	2.3
	Parvomay	0.9	Burgas	5.1	Burgas	5.1	Troyan	5.0	Plovdiv	5.0	Burgas	2.9	Burgas	2.9	Burgas	1.9	Burgas	1.9
	Chirpan	0.9	Plovdiv	5.0	Plovdiv	5.0	Plovdiv	5.0	Gabrovo	5.0	Varna	2.3	Varna	2.4	Varna	1.4	Varna	1.5
	Sevlievo	0.8	Stolichna	4.6	Stolichna	4.6	Varna	4.5	Varna	4.4	Plovdiv	2.3	Plovdiv	2.3	Plovdiv	1.3	Plovdiv	1.3
	Rakovski	0.4	Varna	4.4	Varna	4.4	Stolichna	4.2	Stolichna	4.2	Stolichna	0.9	Stolichna	0.9	Stolichna	0.4	Stolichna	0.4

Source: Author's calculations using data from the NSI of Bulgaria

w_{id} – weight for the person i in municipality d, which is the inverse value of the inclusion probability for that person;
k – number of municipalities in Bulgaria ($k = 264$).

- For unemployed:

The calculations are the same as these for employed persons and the formula is:

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

where most of the symbols are already known, except:
UN_d^{dir} is a number of unemployed persons in municipality d, $d = 1, 2, ..., k$, calculated using the direct estimation;
x_{id} – dichotomy variable which takes value: 1 – if the person is unemployed and 0 – if the person is not unemployed, for municipality d.

Taking a closer look to the results, it is noticeable that there are no results for 129 of 264 municipalities (about 50 %). This is caused by the small number of sampled persons (under 200) in 107 of 264 municipalities and the missing data in the LFS sample for unemployed persons (9) or for the population as a whole for those areas (13). It can be seen from Table 1 that the municipalities with the highest level of unemployment are Slivo pole (52.0 %), Kaolinovo (45.8 %), Glavinitca (34.5 %), Kotel (29.5 %), Isperih (25.8 %), Devin (25.4 %), Varbitca (25.2 %), Dobrichselska (23.2 %), Razgrad (22.8 %) and Dolni chiflik (22.5 %). The lowest unemployment rates are in Goce Delchev (1.4 %), Dolna Mitropoliya (1.2 %), Belovo (1.1 %), Peshtera (1.1 %), Karlovo (1 %), Elhovo (1 %), Parvomaj (0.9 %), Chirpan (0.9 %), Sevlievo (0.8 %) and Rakovski (0.4 %).

The results for the highest unemployment levels look realistic, but the lowest show some unknown issues because it can be foreseen that the big city's unemployment rates should be the lowest. May be this is caused by the relatively small sample size in these municipalities and the tiny number of persons identified as unemployed at the time of the survey. Also a possible reason for the low unemployment rates could be unreal post-stratification weights. The latter could distort the results because the small number of surveyed persons in some strata leads to large weights which can influence the final estimates for these municipalities.

Frequently the direct estimates are only used as an addendum but not as 'real' estimates. In practice the **indirect estimates** are more common. By combining the data from the Population statistics, concerning the persons aged 15 and over in the municipalities and for the country as a whole, and data for the total employed persons from the Labour Force Survey, we can estimate the number of employed by municipalities. Using the same approach we can calculate the unemployed persons too. The formulas for the **ratio estimates** are as follows:

- For the employed:

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

where:
EM_d^{ratio1} is a number of employed persons in municipality d, $d = 1, 2, ..., k$, using the total ratio estimation;
N_d – number of persons aged 15 and over in municipality d, from the Population statistics;
N – total number of persons aged 15 and over, from the Population statistics;
EM – total number of employed persons, from the Labour Force Survey.

- For the unemployed:

$$UN_d^{ratio1} = \frac{N_d}{N} \cdot UN, \qquad (5)$$

where:
UN_d^{ratio1} is a number of unemployed persons in municipality d, d = 1, 2, ..., k, using the total ratio estimation;
UN – total number of unemployed persons, from the LFS;

The results from the calculations are as shown in Table 2.

It is noticeable that even though the results for the number of employed and unemployed persons are realistic, the unemployment rate is the equal for all municipalities – 6.9 %. This is based on the fact that:

$$UR_d^{ratio1} = \frac{UN_d^{ratio1}}{EM_d^{ratio1} + UN_d^{ratio1}} \cdot 100 =$$

$$= \frac{\frac{N_d}{N} \cdot UN}{\frac{N_d}{N} \cdot EM + \frac{N_d}{N} \cdot UN} \cdot 100 = \quad (6)$$

$$= \frac{\frac{N_d}{N} \cdot UN}{\frac{N_d}{N} \cdot (EM + UN)} \cdot 100 = \frac{UN}{(EM + UN)} \cdot 100 = UR$$

Namely the unemployment rates in the municipalities are equal to the unemployment rate for the country as a whole. This is the main drawback of this method because there

Table 2. Employed, unemployed and unemployment rate, using the total ratio estimation

Municipality	Employed	Unemployed	Unemployment rate (%)
Bansko	5557.1	410.4	6.9
Belitsa	3809.3	281.3	6.9
Blagoevgrad	32565.3	2405.2	6.9
Gotse Delchev	13043.5	963.3	6.9
Garmen	5733.2	423.4	6.9
Kresna	2482.7	183.4	6.9
Petrich	23487.2	1734.7	6.9
Razlog	9029.7	666.9	6.9
Sandanski	17764.8	1312.0	6.9
Satovcha	7014.2	518.0	6.9
...	...	...	...
Nikola Kozlevo	2637.4	194.8	6.9
Novi pazar	7846.0	579.5	6.9
Smyadovo	3302.3	243.9	6.9
Hitrino	2860.0	211.2	6.9
Shumen	43217.1	3191.9	6.9
Bolyarovo	2067.8	152.7	6.9
Elhovo	7652.3	565.2	6.9
Straldzha	5910.8	436.6	6.9
Tundzha	12414.2	916.9	6.9
Yambol	33024.8	2439.1	6.9

Source: Author's calculations using NSI database.

is no way a comparison to be made. Therefore using these results no adequate decision can be made.

Using almost the same calculations we can estimate the unemployment rate by the use of **district (oblast) level information**. The estimates are as follows:

- For the employed:

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

where:
EM_d^{ratio2} is a number of employed persons in municipality d, d = 1, 2, ..., k, using the district level ratio estimation;
N_{ds} – number of persons aged 15 and over in municipality d, belonging to district s, based on the Population statistics data;
N_s – number of persons aged 15 and over in the district s, based on the Population statistics;
EM_s – number of employed persons in district s, based on the LFS database.

- For the unemployed:

Table 3. Number of employed, unemployed and unemployment rate, using the district ratio estimation

Municipality	Employed	Unemployed	Unemployment rate (%)
Bansko	6403.2	152.7	2.3
Belitsa	4389.3	104.7	2.3
Blagoevgrad	37523.9	894.8	2.3
Gotse Delchev	15029.5	358.4	2.3
Garmen	6606.2	157.5	2.3
Kresna	2860.8	68.2	2.3
Petrich	27063.5	645.4	2.3
Razlog	10404.6	248.1	2.3
Sandanski	20469.7	488.1	2.3
Satovcha	8082.2	192.7	2.3
...	...	...	...
Nikola Kozlevo	2424.0	512.2	17.4
Novi pazar	7211.3	1523.9	17.4
Smyadovo	3035.2	641.4	17.4
Hitrino	2628.6	555.5	17.4
Shumen	39721.0	8393.8	17.4
Bolyarovo	2142.6	165.2	7.2
Elhovo	7928.9	611.3	7.2
Straldzha	6124.5	472.2	7.2
Tundzha	12863.0	991.8	7.2
Yambol	34218.8	2638.3	7.2

Source: Author's calculations using NSI database

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

where:
UN_d^{ratio2} is a number of unemployed persons in municipality d, d = 1, 2, ..., k, using the district level ratio estimation;
UN_s – number of unemployed persons in district s, based on the LFS database.

After the estimation of the employed and unemployed number and unemployment rates, the results are shown in Table 3.

Using a thematic map we can present the results also as follows in Figure 2.

Here we can clearly see that the estimated coefficients (rates) of unemployment for the municipalities reproduce the rates from the district level ones – more aggregated. This is because:

$$UR_d^{ratio2} = \frac{UN_d^{ratio2}}{EM_d^{ratio2} + UN_d^{ratio2}} \cdot 100 =$$

$$= \frac{\frac{N_{ds}}{N_s} \cdot UN_s}{\frac{N_{ds}}{N_s} \cdot EM_s + \frac{N_{ds}}{N_s} \cdot UN_s} \cdot 100 = \qquad (9)$$

$$= \frac{\frac{N_{ds}}{N_s} \cdot UN_s}{\frac{N_{ds}}{N_s} \cdot (EM_s + UN_s)} \cdot 100 = \frac{UN_s}{(EM_s + UN_s)} \cdot 100 = UR_s$$

In contrast to abovementioned methods, the **regression estimates** give more realistic results. Here it is important to declare that in order to obtain reliable and robust regression coefficients of the estimated models, and therefore plausible estimates of the employed and unemployed number, the following *requirements* have to be observed concerning the LFS data:

- In the process of regression coefficient estimation the following criteria are observed:
  - The number of the sampled cases in given municipality must be at least 200;
  - No missing values for any of the variables used;
  - No outliers in the municipality data.

Figure 2. Unemployment rates in municipalities, using district ratio estimation (ratio estimate type 2)

• For the concrete estimation of the number of employed and unemployed only the models that fulfill the requirements of the Ordinary Least Squares method are used i.e. adequate models, significant regression coefficients, no multicollinearity, high R square and normal distribution of the residuals[2].

The first requirement is fulfilled only in 114 of 264 municipalities. Concerning the second one, only these that fulfill it are presented here. The rest models are out of our scope and are not mentioned thereafter.

The process of the employed and unemployed number estimation using regression estimates includes the following steps:

◊ Building a file using the Labour Force Survey data containing the variables needed for the regression estimation procedures over the 114 chosen municipalities;
◊ Executing stepwise regression analysis [3], "sifting out" the non-significant factor variables by the means of the partial F-criteria.
◊ Checking the regression estimation method requirements fulfillment;
◊ Building the regression models up using the estimated coefficients;
◊ Substituting the data for the factor variables with the data from Population statistics and estimation of the employed and unemployed number for all 264 municipalities;
◊ Calculation of the municipality unemployment rates using the employed and unemployed number estimates.
◊ Following the above described steps we obtain the **ordinary linear regression estimation** equation as follows:

• For the employed:

$$\hat{EM}_d^{regr} = \underset{(0.065)}{0.826}X_{1d} + \underset{(0.024)}{0.153}X_{2d} + \underset{(0.085)}{0.193}X_{3d}, \quad (10)$$

where:
X_1 is a number of persons aged 25-49;
X_2 – total number of urban population;
X_3 – males aged 50 and over[3].

• For the unemployed:

$$\hat{UN}_d^{regr} = \underset{(121.589)}{480.584} + \underset{(0.030)}{0.330}X_{1d}, \quad (11)$$

where:
X_1 is females aged 15-24.

Combining the data from the two sources (regression using the LFS data and the Population statistics factors' data), we estimate the number of employed and unemployed persons in all Bulgarian municipalities. By the use of these estimates we can calculate the municipalities' unemployment rates. The results are presented on the Figure 3.

As it can be seen from the Table 1 and the map above, the highest municipality unemployment rates are in Treklyano (66.1 %), Boynica (57.9 %), Chavdar (55.9 %), Makresh (52.4 %), Kovachevtsi (51.8 %), Anton (48.4 %), Madzharovo (46.9 %), Chelopech (46.8 %), Chuprene (46.6 %) and Georgi Damyanovo (43.2 %). The municipalities Varna (4.4 %), Stolichna (4.6 %), Plovdiv (5 %), Burgas (5.1 %), Pernik (5.1 %), Stara Zagora (5.2 %), Gabrovo (5.2 %), Pleven (5.3 %), Ruse (5.4 %) and Kyustendil (5.8 %) have the lowest unemployment rates. Taking a closer look at the results we can see that they are much realistic than those obtained by using direct and ration estimates – big urban municipalities have lower unemployment rates and rural ones have higher ones.

[2] SPSS ver.13 is used for estimation procedures
[3] Figures in brackets are coefficients errors.

Figure 3. Municipalities' unemployment rates, using the ordinary linear regression estimation[4]

In Table 1 we can also see that it is possible to calculate the so called **mixed estimates**. This can be made by using the formula:

$$Y_d^{comp} = \gamma . Y_d^{dir} + (1 - \gamma) . Y_d^{regr} \qquad (12)$$

where:
Y_d^{comp} is the mixed estimate (e.g. employed number) for the small area (municipality) d, calculated as a weighted average of the direct estimate Y_d^{dir} for the municipality d and Y_d^{regr} – the regression estimate for the small area d;
γ – special chosen weight usually estimated by:

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

where:
n_d is the total number of sampled cases in the small area d;
N_d – total number of individuals in the population in the same area.

If we take a look at the results concerning the mixed estimates, we can see that these estimates do not differ substantially from those ones that they refer to. For example the regression estimates for Boinica municipality is 57.9 % and the mixed one is 57.3 %. This is caused by the small number of sampled cases in that small area (municipality) – usually about 3 % from the total population in given area.

An improvement in the estimation process can be made by the use of **log-linear regression estimates**. Using logarithm over the data leads to suppression of the great data variation. The models for these kinds of estimates are as follows:

- For the employed:

$$\ln(\hat{E}M_d^{ln\_reg}) = \underset{(0.208)}{0.425} + \underset{(0.068)}{0.735}\ln(X_{1d}) + \underset{(0.042)}{0.156}\ln(X_{2d}) + \underset{(0.039)}{0.101}\ln(X_{3d}), \qquad (14)$$

where:
X_1 is the total population aged 25-49;
X_2 – total females number in urban areas;
X_3 – males aged 15-24.

[4] Mapping is made by using MapInfo ver 9.5 software.

- For the unemployed:

$$\ln(\hat{U}N_d^{ln\_reg}) = \underset{(0.011)}{0.893}\ln(X_{1d}) \quad (15)$$

where:
X_1 is the females aged 15-24.

By following the already known steps we obtain the following results mapped as shown on Figure 4.

From Figure 4 and Table 1 it is clear that the highest unemployment rates are in: Mineralni bani (25.7 %), Garmen (24.9 %), Satovcha (24.2 %), Ruen (23.4 %), Kaynardzha (22.4 %), Marica (22.2 %), Nikola Kozlevo (22.2 %), Chernoochene (21.9 %), Avren (21.7 %) and Dobrich-rural (21.6 %) municipalities. The least affected by the unemployment phenomenon are the Capital Sofia (4.2 %), Varna (4.5 %), Plovdiv (5 %), Troyan (5 %), Gabrovo (5 %), Pernik (5.1 %), Burgas (5.1 %), Chiprovci (5.2 %), Stara Zagora (5.3 %) and Pleven (5.3 %) municipalities.

Additional improvement in the estimates can be provided by the use of **regional variables** in the regression analysis. By that approach we can account for the different district features and characteristics that given municipalities belong to. In the case of regional data usage we can estimate the results by the ordinary regression or by the **Bayesian hierarchy approach**. The latter characterizes with the use of preliminary information (apriori) in the estimation process[5]. This, however, leads to the need for iteration procedures usage in the process of regression coefficients estimation[6].

Using the available data and the proper software we obtain the following models:

- For the employed:

$$\ln(\hat{E}M_d^{bayes}) = \underset{(0.284)}{0.330} + \underset{(0.068)}{0.738}\ln(X_{1d}) + \\ + \underset{(0.042)}{0.155}\ln(X_{2d}) + \underset{(0.040)}{0.100}\ln(X_{3d}) + \underset{(0.028)}{0.131}\ln(U_d) \quad (16)$$

where:

Figure 4. Municipalities' unemployment rates, using the log-linear regression estimates

[5] In the case of this study data for the distribution of the regression coefficients is used.
[6] By the use of WinBUGS software we bring a little comfort in the estimation procedures.

X_1 is the total population aged 25-49 г.;
X_2 – total number of females in urban areas;
X_3 – males aged 15-24;
U_i – GDP in million levs for the given district that municipality belongs to.

- For the unemployed:

$$\ln(\hat{UN}_d^{bayes}) = \underset{(1.293)}{5.583} + \underset{(0.094)}{0.614}\ln(X_{1d}) - \underset{(0.161)}{0.550}\ln(U_d), \quad (17)$$

where:
X_1 is the females aged 15-25;
U_i – GDP in million levs for the given district that municipality belongs to.

The results for the level of unemployment show that there is no significant difference between the two methods (see Table 1). The mapping of the results is as shown in the Figire 5.

Judging from the results from Table 1 and Fig. 5, the municipalities with the highest unemployment rates are Ruzhintci (44.5 %), Novo selo (43.6 %), Chuprene (43.3 %), Makresh (42.6 %), Kaynardzha (41.8 %), Boynica (41.7 %), Georgi Damyanovo (39.7 %), Kovachevci (39.6 %), Borino (39.2 %) and Iakimovo (38.1 %), and these with the lowest ones are: the Capital Sofia (0.4 %), Plovdiv (1.3 %), Varna (1.4 %), Burgas (1.9 %), Stara Zagora (2.2 %), Asenovgrad (2.5 %), Karlovo (2.8 %), Ruse (3.1 %), Kazanlak (3.1 %) and Pleven (3.3 %). Here we can see a great range between the lowest and the highest value that approximates 44.1 % but the results are logical and confirm the above mentioned ratiocinations.

Conclusion

The application of the small area methods for the municipality unemployment rates estimation, by the combination of the Labour Force Survey data with the Population statistics data, is unique for the Bulgaria statistical practice. The special use of three different software products at the estimation process, gave a chance to realize the purpose of this study, overcoming the problem with the lack of specific software for small area estimation application.

Figure 5. Municipalities' unemployment rates, using the log-linear regression with regional variables.

Considering the concrete results from the estimation process, we can mention that there is a slight difference in the various estimation approaches, probably based on the specific nature of the methods. However the employed and unemployed number estimates, and the derivative unemployment rates, are realistic and can be used in the decision making process. Generalizing the results we can conclude that the municipalities with the highes unemployment rates are Nikola Kozlevo, Kajnardzha, Mineralni bani, Novo selo, Iakimovo, Borino and Chernoochene. The lowest is the unemployment in the Capital (Sofia) and the big cities Plovdiv, Burgas, Varna, Stara Zagora and Pleven, which is reasonable.

Some of the methods gave some non-logical results, e.g. direct estimate and ratio estimates. The reasons for that root in fact that the direct estimate in the 50 % of the municipalities is based on a small number of sample cases and inadequate post-stratification weights. The ratio estimates suffer from the fact that they reproduce the results from higher territorial levels – districts and country.

The regression estimates are logical and adequate. Concerning the models the significant factors are: *for the employed* – total population aged 25-49, total urban population and male aged 50 and over; *for the unemployed* – females aged 15-24 and GDP at district level.

Finally it can be concluded that by the use of small area estimation techniques we can resolve one big issue concerning the missing reliable information for the unemployment rates at municipality level in Bulgaria. Adapting the specific methodology, accordingly to the concrete Bulgarian information circumstances, we not only overcame the difficulties concerning the information availability and estimation procedures, but we created an effective instrument for adequate and correct decision making. Because of the universality of instrument applied, it can be used not only for the analysis of the situation and changes at the labour market, but for the researching purposes in other scopes of the social-economic life.

Literature

1. Naidenov, Al., Problems in the process of the small area estimation of unemployment in Bulgaria, Collection of papers from the International scientific conference "Statistics and the challenges of the information society in 21st century", Ravda, September 2006, p. 191 – 197.

2. Naidenov, Al., Statistical survey of unemployment in Bulgaria – regional aspects, "Statistics", Sofia, 1/2006 issue, p. 51-64.

3. Suikova, Iv., A. Stoikova – Kunalieva, Sv. Suikova, Statistical correlation study, University publishing "Stopanstvo", Sofia, 2002.

4. Ghosh, M., and Rao, J.N.K., Small area estimation: An appraisal, 1994.

5. Nicholas, T. L., Missing Data and Small-Area Estimation, 2005.