ETHNICITY NON-IDENTIFICATION IN 2011 CENSUS IN BULGARIA*

Kaloyan Haralampiev, Dimitar Blagoev

Abstract: In this paper we present results of secondary data analysis aimed at understanding the ethnic non-identification problem in the 2011 Census in Bulgaria. Using cluster analysis and analysis of associations we study the levels of dependency between specific sense of ethnic belonging and the absence of ethnic self-identification, expressed through refusal to answer the ethnicity question in Census 2011. Our analysis reveals that the relation between ethnic identity and lack of ethnic self-identification is mediated through the specific ethno-demographic structure within the respective municipality as well as, to a minor extent, by the administrative organization at district level. One of the factors of non-response to the ethnicity question in Census 2011 is the problematic ethnic identity among several ethno-demographic groups, including: the Bulgarian Muslims or 'pomaks' (almost entirely); other small compact (within the municipality) ethnic groups other than the three main ones in the country (partially); and those who had identified as Turks in municipalities with predominantly Turkish population in 2001 (a small portion of them).

Keywords: ethnicity non-identification; Census; cluster analysis; analysis of associations

JEL: C25, C38, C80, J11

1. INTRODUCTION

As official data from the 2011 Census carried out by the Bulgarian National Statistical Institute (NSI) show, 9.28 percent of the Bulgarian population did not answer the optional question on ethnicity. This share, compared with the corresponding share of 0.31 percent in 2001, presents a significant increase of people who did not respond.

* This paper is developed on the basis of a poster presentation at the European Population Conference (EPC 2014) held in Budapest, Hungary, 25-28 June 2014.

Is there, however, any scientific problem relevant to the large share of non-response to the ethnicity question? In our view, there is a research problem here on, at least two, levels. First, when there are significant shares of non-response to as important an ethno-demographic question as that of ethnic origin, the methodological question arises as to how analyses and extrapolations of trends are to be made as regards change of ethnic structure, based on a series of historical data on the indicator? A second important research question is what problematic processes in ethnic self-identification are revealed by the occurrence of such a large rate of refusal to declare one's ethnic group and what are the particular (sets of) factors leading to this ethnic non-identification?

The scientific importance of the problem is discussed in this paper against the background of data analyses revealing the 'history' of ethnic identification in the latest Bulgarian censuses. We argue that this problem deserves scientific attention as there is yet no scientific knowledge about its prerequisites and its consequences in the frame of the recent socio-demographic situation in Bulgaria.

Apart from research issues, the question discussed here also involves problems related to social practice. First, the absence of relevant data about the ethnic origin of one tenth of the national population can potentially create difficulties for the implementation of demographic policies (or the implementation of a variety of social policies indirectly linked to demographic ones) that require such data[1]. Second, given the existence of significant blanks in knowledge about important ethno-demographic characteristics of a national society (or of regional cross-sections of that society), it is possible that the self-knowledge of that society (and of specific groups of individuals in it) regarding its socially valid models of ethnic identification to be questioned or obscured, particularly when these models have a potential to become a reality through information made available to the public about the actual 'incomplete' ethno-demographic picture of society.

2. HISTORY OF THE PROBLEM

In all population censuses that took place in Bulgaria since 1881 till 1992, the question as to ethnic belonging, whatever its specific formulation, had always been a mandatory one to be answered. Answering this question became optional for the first time in Census 2001, when the possibility was provided to indicate one refuses to self-identify; and this condition was applied in Census 2011 as well. In the earlier of these two censuses the share of refusals to self-identify, and the lack of an answer to this question, were respectively 0.8 percent and 0.3 percent of the whole population. In the second census, however, they were respectively 0.7 percent and 9.3 percent and while the percentage of refusal to state one's ethnic identity has remained the same, the real problem is that the share of non-response to the ethnic identity question *has increased thirty times*. Since this is hardly coincidental, the task of our research effort has been to

[1] The need for both reliable and relevant data can be illustrated with the following example: Imagine only one third of the population answers the optional question about ethnic origin, then the data are reliable but they are not relevant as regards the needs of demographic and social policies.

seek the causes of this large increase. Some of these causes might be attainable through scientific methods, others, not.

In our previous publication on that matter (Blagoev and Haralampiev 2014) we sought to ascertain the preconditions and patterns of non-response based solely on publicly available data from Census 2011[2]. The results of this analysis pertain to highlighting differences by age and education (at national level) as well as regional differences (by districts, municipalities, types of settlement) in the separate ethnic groups and in the group of people who have not responded; however, significant specific correlations, with which to ascertain the socio-demographic preconditions of the problem, cannot be identified. The basic result, though, is that our analysis proved there is not just one single logic/explanatory model influencing non-response; there are multiple, interweaving and interacting, substantially heterogeneous social logics/explanatory models relevant to non-response.

3. DEFINING THE TASKS AND LIMITATIONS OF ANALYSIS. APPROACH.

In our previous analyses on this problem, we refuted the assumption that (a considerable share of) non-response is concentrated in a specific ethnic group, as well as the assumption that non-response is evenly (proportionately) distributed among all ethnic groups. A new hypothesis arose, that if there is some ethnic logic underlying non-response, it is primarily the result of a specific ethnic configuration *at municipality level*. This means, firstly, that, on the basis of cluster analysis of ethnic structures at municipality level, the municipalities should be distributed into typological groups based on the type of ethnic structure in them; and secondly, that an analysis should be made of the preconditions (factors) of non-response. The task is to have the obtained ethnic clusters (groups) correlated with the clusters of non-response (their structure), in order to ascertain whether there is an empirical linkage between non-response and ethnic structures at municipality level.

This approach, however, has the following specificity: non-response to the question about ethnic origin (i.e. no ethnic group is indicated) is part of the structure of ethnic groups, and the size of this share varies between a minimum of 0.3 percent and a maximum of 53.7 percent. Clearly, in municipalities in which this percentage is low, the structure of those who have answered the question is close to the structure of the respective ethnic structure. On the contrary, in municipalities with high shares of non-response, the current ethnic configuration can be deduced from the ethnic structure of non-response. In other words, this configuration should be used in correlation analysis, with the reservation that the used ethnic structure is extrapolated from the structure of those who did answer the question, and not from the structure of the whole local population.

[2] The scarcity and the higher level of data aggregation hindered to a greater extent any profound analysis.

Therefore, before carrying out this procedure with the data of Census 2011, we will first apply the above approach to the Census 2001 data, where for the first time the question as to ethnic affiliation became optional, but where, also, the mean percentage of non-response was very low, almost negligible – 0.3 percent. The purpose of seeking the correlation between ethnic structure and rate of non-response is to ascertain to what extent a clear ethnic picture in 2001 can serve as a 'prognostic' starting point in 'simulating' the shares of non-response in each municipality using data on lack of ethnic identification obtained ten years later. The use of this analytical framework can be considered an acceptable option, though with the reservation that it holds some uncertainty insofar as it does not take into account the socio-demographic changes that have taken place within this ten-year period. These differences, in fact, cannot be studied with sufficient reliability, for Bulgarian statistics lack data about a basic kind of change of population between two censuses, due to migration (both within the country and abroad)[3], and hence it is not possible to extrapolate regarding the prognostic share of ethnic groups on the basis solely of data on population change resulting from births, deaths and ethnic-specific fertility rates in the period 2001-2011[4]. In brief, in this case we will seek correlations between two independent variables, and the assumption will nevertheless be that the ethnic structure in 2001 is a factor in statistical terms[5].

After that we will apply the same procedure of analysis to the data on ethnic structures in municipalities obtained from the 2011 Census; the aim will be the same, but in that case the variables 'share of non-response' and 'type of ethnic structure' will be mutually dependent.

Finally, we will analyze the changes in the type of ethnic structure that occurred between 2001 and 2011, if any.

4. ANALYSIS

4.1. Cluster analysis of municipalities by the relative share of non-response to the question about ethnicity.

Using IBM SPSS Modeler software, we made an automatic clusterization of all 264 municipalities with respect to the share in each municipality of non-response to the question about ethnic identity in Census 2011. The results of this clusterization, which divides municipalities into typological groups, are the following:

[3] It is an important variable since the population in Bulgaria has decreased by 7.1 percent within this ten-year period, the main factor of this decline being emigration. But the extant data on emigration are not reliable.

[4] Cf. Arkadiev 2012, who deals only with this, and does not take into account the migration factor (internal migration and especially emigration).

[5] In other words, we will strive to show whether there is, or not, a correlation, without being able to claim there is a cause and effect relationship.

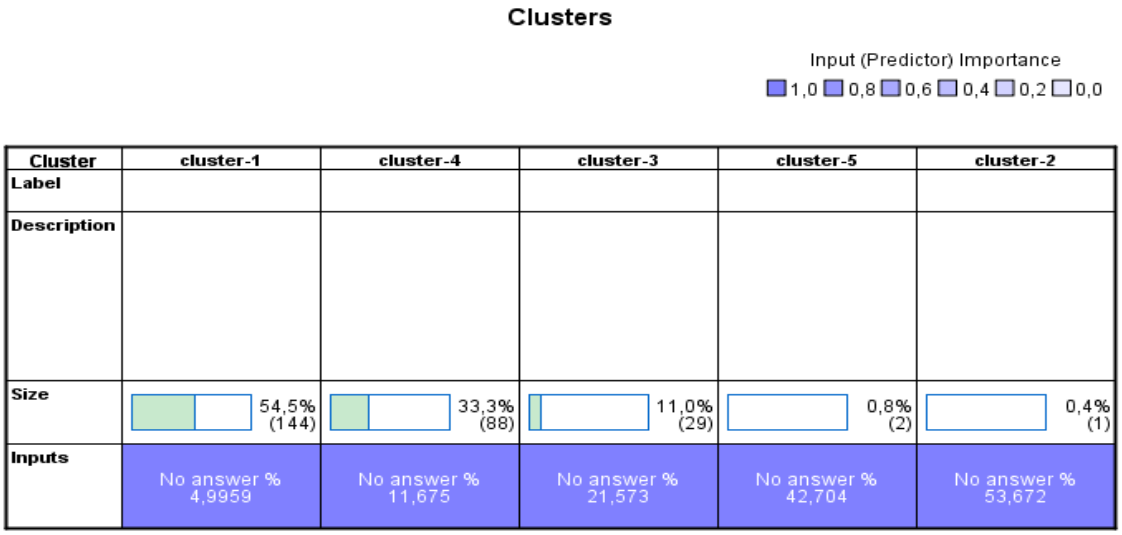


Chart 1. Clusters of non-response (shares and means), Census 2011

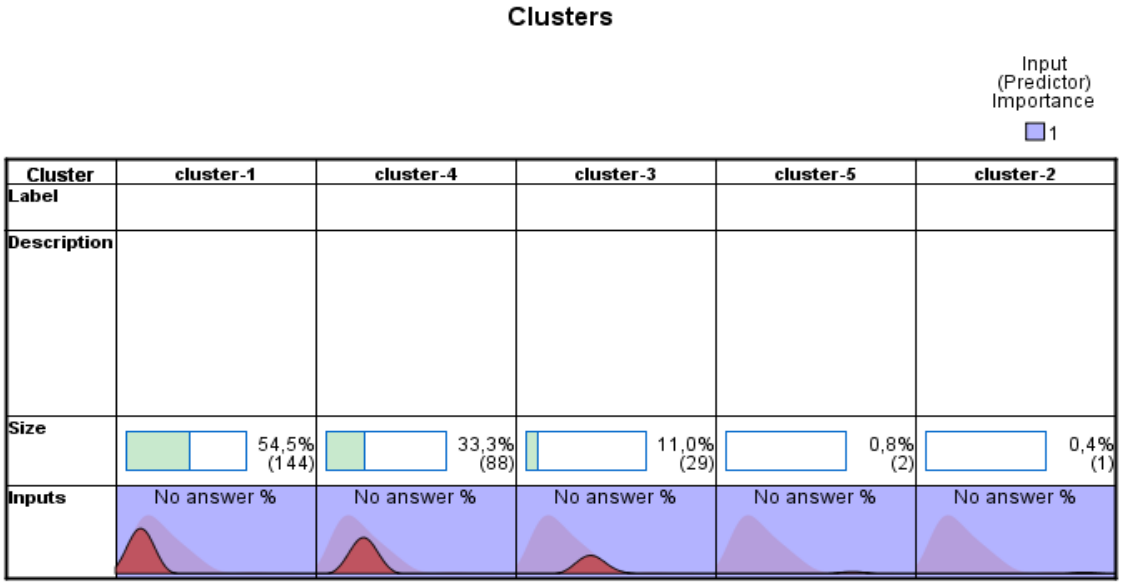


Chart 2. Clusters of non-response (shares and frequency distributions), Census 2011

Following the graph contents, we gave the resultant five clusters the following working names:

Cluster#1, „*Municipalities with small share of non-response to the question about ethnic identity*"; containing 144 municipalities (54.5% of all 264 municipalities);

Cluster#2, „*Municipality with largest share of non-response to the question about ethnic identity*", containing one municipality (0.4%);

Cluster#3, „*Municipalities with large share of non-response to the question about ethnic identity*", containing 29 municipalities (11.0%);

Cluster#4, „*Municipalities with medium large share of non-response to the question about ethnic identity*", containing 88 municipalities (33.3%);

Cluster#5, „*Municipalities with very large share of non-response to the question about ethnic identity*", containing 2 municipalities (0.8%).

As the clusterization results show, in most municipalities the share of non-response is small (below 8.3 percent in 144 (54.5 percent of) municipalities) or medium (between 8.5 percent and 16.6 percent in 88 (33.3 percent of) municipalities). In the remaining 32 (12.2 percent of) municipalities this share (above 17.3 percent) is large.

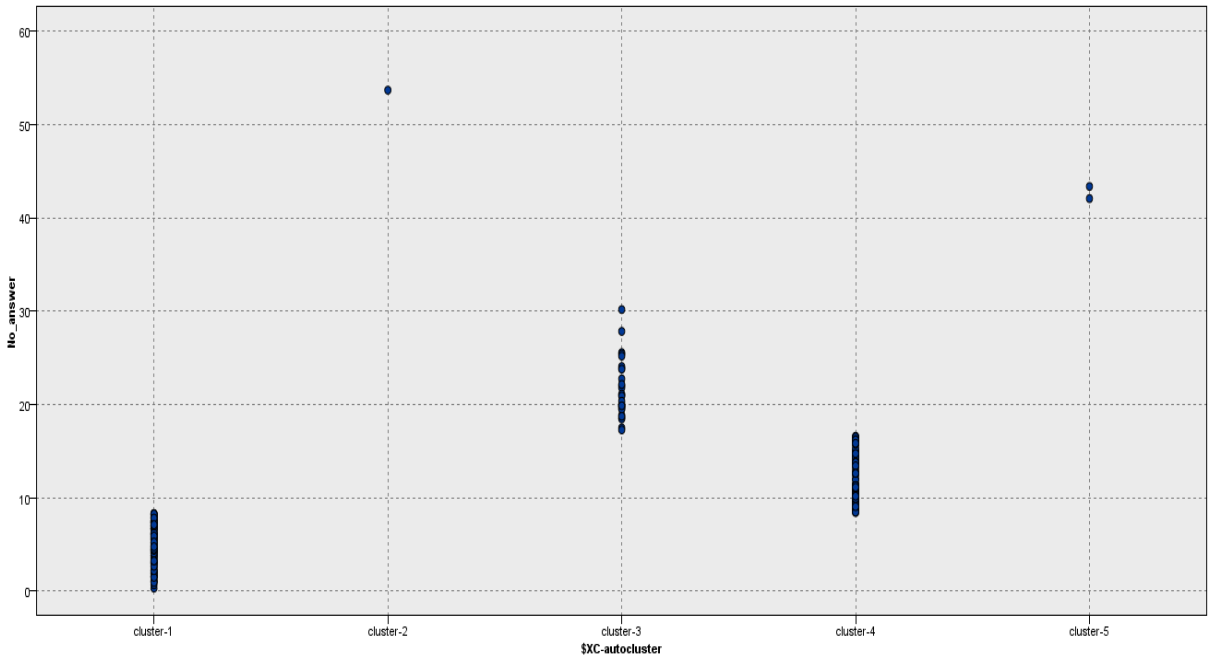


Chart 3. Distributions of municipalities by clusters of non-response, Census 2011

It is worth noting, however, that 'small', 'medium' and 'large' applied to the share of non-response are not defined here in absolute terms but relatively, for they are derived from the 'benchmarks' found in clusters. Other groupings, and consequently other scales, with different definitions of group distinctions, are also possible using the method chosen.

4.2. Cluster analysis of ethnic structures at municipality level in 2001, dividing the municipalities into typological groups.

Using IBM SPSS Modeler software, we established clusters of the ethnic structures at municipal level with the data used being derived from answers to the ethnic identity question in Census 2001 in all 262 municipalities[6]. The results of clusterization, dividing municipalities into typological groups, are the following:

Clusters

Input (Predictor) Importance
1,0 0,8 0,6 0,4 0,2 0,0

Cluster	cluster-1	cluster-5	cluster-2	cluster-4	cluster-3
Label					
Description					
Size	56,9% (149)	22,1% (58)	14,1% (37)	3,8% (10)	3,1% (8)
Inputs	Bulgarian % 90,214	Bulgarian % 76,258	Bulgarian % 30,472	Bulgarian % 77,387	Bulgarian % 84,421
	Turkish % 4,9999	Turkish % 8,4778	Turkish % 60,790	Turkish % 10,899	Turkish % 4,7856
	Other % 0,4583	Other % 0,5553	Other % 0,7971	Other % 6,2876	Other % 0,6335
	Non-identification % 0,5836	Non-identification % 0,6717	Non-identification % 1,6562	Non-identification % 0,9418	Non-identification % 9,3717
	Roma % 3,7446	Roma % 14,038	Roma % 6,2854	Roma % 4,4841	Roma % 0,7883

Chart 4. Clusters of ethnic structures (shares and means), Census 2001

Following the graph contents, we gave the following working titles to the five clusters:

Cluster#1, „*Municipalities with predominantly Bulgarian population*"; containing 149 municipalities (56.9% of all 262 municipalities);

Cluster#2, „*Municipalities with predominantly Turkish population*"; containing 37 municipalities (14.1%);

[6] The numerical difference between 262 municipalities (in 2001) and 264 (in 2011) is a result of administrative changes that took place in two of the municipalities after 2001; in both cases parts of territory were separated from a larger municipality (Karlovo and Rodopi) to form new municipalities (respectively Sopot and Kuklen).

Cluster#3, „*Municipalities with predominantly Bulgarian ethnic population and a high share of non-identification*"; containing 8 municipalities (3.1%).

Cluster#4, „*Municipalities with predominantly Bulgarian ethnic population and a high share of other ethnic groups*"; containing 10 municipalities (3.8%);

Cluster#5, „*Municipalities with predominantly Bulgarian ethnic population and high share of Roma*"; containing 58 municipalities (22.1%);

Before proceeding with the analytical timetable, a preparatory step was taken in listing all municipalities by their names according to the five clusters given above. The purpose of this measure was to compare the cluster membership of randomly selected municipalities with their ethnic structure according to the original data from the same Census 2001 data pool, as well as to verify the background information we already had about some ethnic particularities at local level. However, the comparison showed that the five clusters had two basic defects: first, they did not sufficiently differentiate the types of municipalities; second, they were not precise enough as regards the division of particular municipalities according to their ethnic configuration. One example of the first kind of defect is the lack of clusters including municipalities in which the sum total of all ethnic minorities[7] is larger than the ethnic Bulgarian population there; or of clusters in which the share of people who have not identified their ethnicity is distinctly larger than the mean for the country. One example of the second kind of defect is Cluster#3, which contains only 8 municipalities, although in fact municipalities with similar typological parameters are more in number.

To overcome these two defects and continue with the analysis, we had to use another approach.

4.3. Analysis (based on mean and standard deviations) of ethnic structures at municipal level in 2001, dividing the municipalities into typological groups.

In order to achieve greater differentiability and precision when dividing the municipalities into typological groups based on their ethnic structures, we applied the good old method of calculating means and standard deviation for each of the ethnic groups (including those who have not identified their ethnicity or that have not given a response); we took 262 municipalities as a base. The aim of this approach was to divide the shares of the respective ethnic groups in each municipality in intervals defined in relation to the mean values for the respective ethnic group and within the boundaries of one, two, three or more standard deviations from the means. The results (in percentage) are given in the following table:

[7] Of course, at national level they are minorities.

Table 1. Means and standard deviations of percentages of different ethnic groups within municipalities, Census 2001

	Bulgarian	**Turkish**	**Roma**	**Other**	**Non-identification**
Mean	78.0	14.8	6.6	0.8	1.1
Std. Deviation	22.5	21.9	5.6	1.4	2.0
Minimum	1.8	0.0	0.0	0.0	0.2
Maximum	100.0	97.3	27.5	12.1	21.6

Using these data, the intervals in which the ethnic group shares are situated in each municipality were determined as follows:

Table 2. Intervals of percentages of different ethnic groups within municipalities, Census 2001

		Intervals				
Relative share	**Range**	**Bulgarian**	**Turkish**	**Roma**	**Other**	**Non-identification**
Very low	2SD+ below Mean	0 – 33.0%	-	-	-	-
Low	1SD-2SD below Mean	33.0% – 55.5%	-	0 – 1.0%	-	-
Below the average	0-1SD below Mean	55.5% – 78.0%	0 – 14.8%	1% – 6.6%	0 – 0.8%	0 – 1.1%
Above the average	0-1SD above Mean	78.0% – 100.0%	14.8% – 36.7%	6.6% – 12.2%	0.8% – 2.2%	1.1% – 3.1%
High	1SD-2SD above Mean	-	36.7% – 58.6%	12.2% – 17.8%	2.2% – 3.6%	3.1% – 5.1%
Very high	2SD+ above Mean	-	58.6% – 97.3%	17.8% – 27.5%	3.6% – 12.1%	5.1% – 21.6%

We coded each interval for the respective ethnic group with numbers from 1 to 6, where ‘1’ means ‘*Very low share’*, ‘2’ is ‘*Low share’*, ‘3’ is ‘*Below average share’*, ‘4’ is ‘*Above average share’*, ‘5’ is ‘*High share’* and ‘6’ is ‘*Very high share’*; in this way we obtained 73 (seventy-three) particular combinations in the 262 municipalities. These combinations were processed to form several different clusters, and as a result they fall into eight distinct types of ethnic structures, as follows:

Table 3. Distribution of municipalities by ethnic clusters, Census 2001

Cluster#	**Type of ethnic structure of municipalities**	**Number of municipalities**
1	Predominantly Bulgarian population	105
2	Prevalence of Bulgarian population and a high share of Roma	66
3	Prevalence of Turkish population	31
4	Prevalence of Bulgarian population and high share of Turkish	28
5	Prevalence of Bulgarian population and high share of Other	4
6	Prevalence of Bulgarian population and high share of Non-identification	11
7	Prevalence of Bulgarian population and high share of all ethnic minorities	12
8	Bulgarian population less than the total of all ethnic minorities	5

By forming these eight new clusters (different typological groups with similar ethnic structures), both of the above-mentioned defects have been surmounted. In this way the ethnic picture of Bulgarian municipalities becomes more differentiated, more specific and more adequate to the ethno-demographic situation at the time of Census 2001 implementation.

4.4. Analysis of the relation between rate of non-response in 2011 and the type of ethnic structure of municipalities in 2001

The focus of this part of the analysis is the relation between clusters of non-response to the question about ethnic identity at the time of Census 2001 implementation and the already differentiated clusters (types) of ethnic structures.

The analysis shows that the correlation between the two is significant: the value of Cramer's V is 0.36. Within the framework of the ethnic structures themselves, the following more important relations are found to exist:

The relative weight of the high share of non-response to the ethnic identity question in Census 2011 in the municipalities with predominantly Bulgarian population in 2001 is twice smaller than the mean and for the municipalities with a prevalence of Bulgarian population and high share of Roma it is smaller by one fourth than the mean. In other words, *in municipalities with a high share of Bulgarian population, the high rate of missing ethnic identification occurs much less often than in all the other municipalities*. This indicates, most likely, that the connection between ‘ethnicity’ and ‘lack of problem regarding ethnic identification’ is not a direct connection between Bulgarian ethnicity and answering the ethnicity question, but is mediated by the specific ethnic structure at ‘municipality’ level in which the Bulgarian ethnic group is prevalent.

To the contrary, the relative weight of the high share of non-response to the ethnic identity question in Census 2011 for municipalities with a prevalent Turkish population in 2001 is two and a half times bigger than the mean. This indicates, most likely, that the

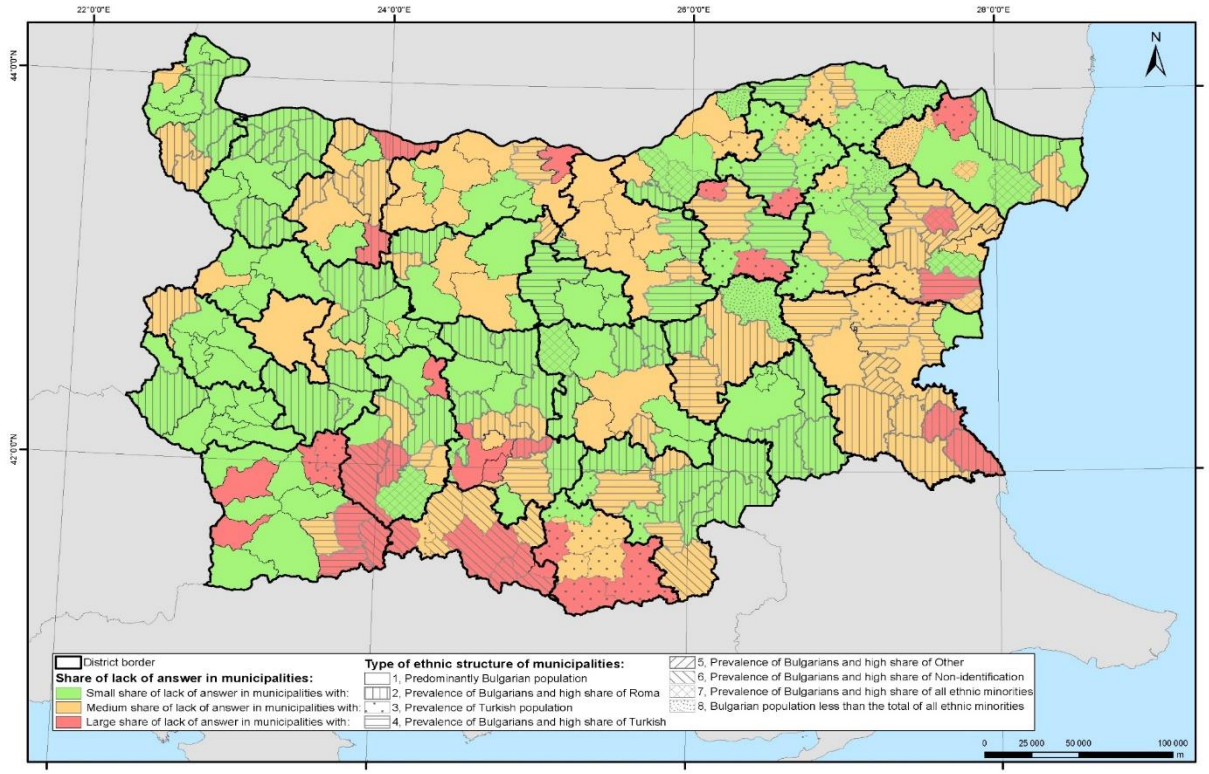
connection between ‘ethnicity’ and ‘lack of ethnic identification’ is not a direct connection between Turkish ethnicity and non-response to the ethnicity question, but is mediated by the specific ethnic structure at ‘municipality’ level in which the Turkish ethnic group is prevalent.

The third conclusion, also related to the high rate of non-response to the ethnicity question, specifically concerns the connection between ethnicity non-identification[8] in 2001 and non-response to the ethnicity question in 2011. The relative weight of the high share of non-response to the ethnic identity question in Census 2011 for municipalities with a prevalence of Bulgarian population and a high share of non-identification in 2001 is more than five times greater than the mean. In other words, the ethnic structure in which, in 2001, the Bulgarian ethnic group was the predominant one, in combination with a very high share of declared refusal to indicate ethnic self-identification, is a precondition for a high share of absence of ethnic identification expressed in the refusal to answer the question about one’s ethnic group ten years later.

To sum up these conclusions, we may say that the connection between ‘ethnicity’ and ‘lack of ethnic self-identification’ is not a direct connection between one’s ethnicity and non-response to the ethnic identity question but is mediated by the specific ethno-demographic structure at municipality level. There are two possible basic interpretations of this general conclusion, a combination of the two not being excluded. First, the visibility of the local ethno-demographic communities in certain municipalities (especially in small ones, of populations less than 20 000 people, which are more than two thirds of all municipalities) may be somehow related to the prevalence of a specific interpretation of one’s own ethnic identity; in this case the visibility of certain ethno-demographic constellations, more than of others, tends to work in favour of a higher rate of ethnic non-identification. Second, the administrative organization of municipality government (determined through local elections) corresponds to specific ethno-demographic constellations, and is thus somehow related to non-response to the ethnic identity question.

We may come closer to verifying these assumptions through a visual representation of the combination of the three elements: ‘Rate of non-response to the ethnic identity question in 2011’, ‘Ethno-demographic structure of the municipalities in 2001’ and ‘Administrative differentiation of municipalities in 2001’, as in the following map:

[8] Those who answered ‘I do not self-identify myself’.



Map 1. Geographic distribution of municipalities by share of no-response (Census 2011) and by type of their ethnic structure (Census 2001)

Analysis of the data visually presented on the map indicate that some new conclusions should be added to the ones already drawn, being directly or indirectly related to the upper level of administrative division – the district. At this level the differentiation of the non-response rates for the ethnic identity question is very distinctly visible when viewed in combination with differentiation by municipalities.

First, there is very clear division of districts into five different groups each of which has approximately the same number of districts in the group as regards the proportion of municipalities in each group with a small, medium and high share of ethnic non-identification[9]:

- The group of districts in which the share of non-response regarding ethnic identity is low in all municipalities (four districts: Montana, in Northwestern Bulgaria; Kyustendil in Southwestern Bulgaria, Gaborovo in North Central Bulgaria, and Yambol in Southeastern Bulgaria);

[9] The only exception is the particular case of the district of Sofia-capital, which also coincides with the administrative unit 'municipality' (Greater Capital Municipality): it had a predominantly Bulgarian population in 2001, but the share of non-response in it was medium.

- The group of districts in which a small number of municipalities (mostly one or two only) have a medium share of non-response to the ethnic identity question and the rest of the municipalities have a prevalent low share of non-response (five districts – Pernik and Sofia-region in Southwestern Bulgaria, Vidin in Northwestern Bulgaria, Silistra and Shumen in Northeastern Bulgaria);
- The group in which municipalities with low and medium non-response are approximately even in number (six districts – Lovech and Veliko Tarnovo in North Central Bulgaria, Ruse in Northeastern Bulgaria, Stara Zagora in South Central Bulgaria, Sliven and Haskovo in Southeastern Bulgaria);
- The group in which municipalities with a low share of non-response for their ethnic identity, on one hand, are comparatively equal in number with municipalities with a medium and high share of non-response, on the other hand (six districts – Razgrad, Targovishte and Dobrich in Northeastern Bulgaria, Blagoevgrad and Pazardzhik in Southwestern Bulgaria, Plovdiv in South Central Bulgaria);
- The group in which a small number of municipalities (mostly one or two only) have a low rate of non-response on ethnicity question and a large number of municipalities have medium and high rates (six districts – Vratsa in Northwestern Bulgaria, Pleven in North Central Bulgaria, Varna in Northeastern Bulgaria, Smolyan and Kardzhali in South Central Bulgaria, Burgas in Southeastern Bulgaria).

Secondly, these differences, besides being due to a factor already established through analysis (i.e., that the specific ethno-demographic structure at municipality level mediates the non-response to the ethnic identity question), are due to at least two (direct or indirect) additional preconditions.

The first additional precondition is administrative organization at district level, which has an impact on the non-response rate; the outstanding example of this are Montana and Vratsa, two neighboring districts in Northwestern Bulgaria, which are almost identical in size and in the ethno-demographic profile of the municipalities included in them[10]: In each district, three municipalities are with a predominantly ethnic Bulgarian population, and eight and six municipalities, respectively, have a prevalent Bulgarian ethnic population and a high share of Roma – yet the two *completely differ one another* with respect to the shares of non-response to the ethnicity question. The district of Montana is in the first of the five typological groups given above, having a low share of non-response in all its municipalities. The neighboring district of Vratsa, however, is at the other extreme in the group of districts in which only very few (two)

[10] They are also identical in their general socio-demographic and economic situation, a similarity that excludes these socio-demographic factors as a possible explanation of the difference in the ethnicity non-response rates.

municipalities have a low share of non-response while the others are with medium (six municipalities) and high (two municipalities) shares.

The second additional precondition is substantially different from the first and represents a projection of an intermediate conclusion stated above, regarding municipalities with a very high non-response rate to ethnicity question in 2001 (against the background of predominant ethnic Bulgarian population) which is a precondition of high non-response to the ethnic identity question in 2011. This second factor is relevant for eight out of the ten municipalities in the district of Smolyan in South Central Bulgaria and for two municipalities that neighbor with Smolyan but are included in two other proximate districts. What is common to them all is the correlation 'high share of non-identification with any ethnicity in 2001' and 'high share of non-response on ethnicity question in 2011', a correlation that is not determined or mediated by the administration factor. The explanation behind that high share of explicit refusal to declare one's ethnic identity in 2001 lies in a specific ethnic-cultural and religious group concentrated precisely in these municipalities of the district of Smolyan, the group of ethnic Bulgarians of Muslim religious affiliation (the so-called 'pomaks'[11]). The historical and cultural-religious background of this population in the last century and a half contains many preconditions for its ethnic identity to be problematic, and this attitude has been expressed in a widespread refusal to identify themselves with the main ethnic groups in the country, registered by the data of Census 2001.

4.5. Analysis of the ethnic structures at municipality level in 2011, dividing the municipalities into typological groups. Analysis of the relation between rate of non-response and the type of ethnic structure within the municipality.

The results of the analysis, through means and standard deviation, of the ethnic structures at municipality level in 2011, which divides municipalities into typological groups, are not significantly different compared to results for 2001, as evident in the following table:

Table 4. Means and standard deviations of percentages of different ethnic groups within municipalities, Census 2011 (in comparison to Census 2001 in parentheses)

	Bulgarian	**Turkish**	**Roma**	**Other**	**Non-identification**
Mean	77.3 (-0.7)	14.6 (-0.2)	7.2 (+0.6)	0.8 (0.0)	1.1 (0.0)
Std. Deviation	23.2 (+0.7)	22.4 (+0.5)	6.4 (+0.8)	1.2 (-0.2)	1.5 (-0.5)
Minimum	1.9 (+0.1)	0.01 (+0.01)	0.04 (+0.04)	0.04 (+0.04)	0.07 (-0.13)
Maximum	99.6 (-0.4)	97.1 (-0.2)	31.4 (+3.9)	9.6 (-2.5)	15.6 (-6.0)

[11] Although this designation is charged with ambiguity and negative connotations, its use in scientific discourse is justified. Cf. for instance, Ivanov 2012.

Based on the data of Census 2011, the intervals in which the percentages of the respective ethnic groups are situated have been defined as follows:

Table 5. Intervals of percentages of different ethnic groups within municipalities, 2011

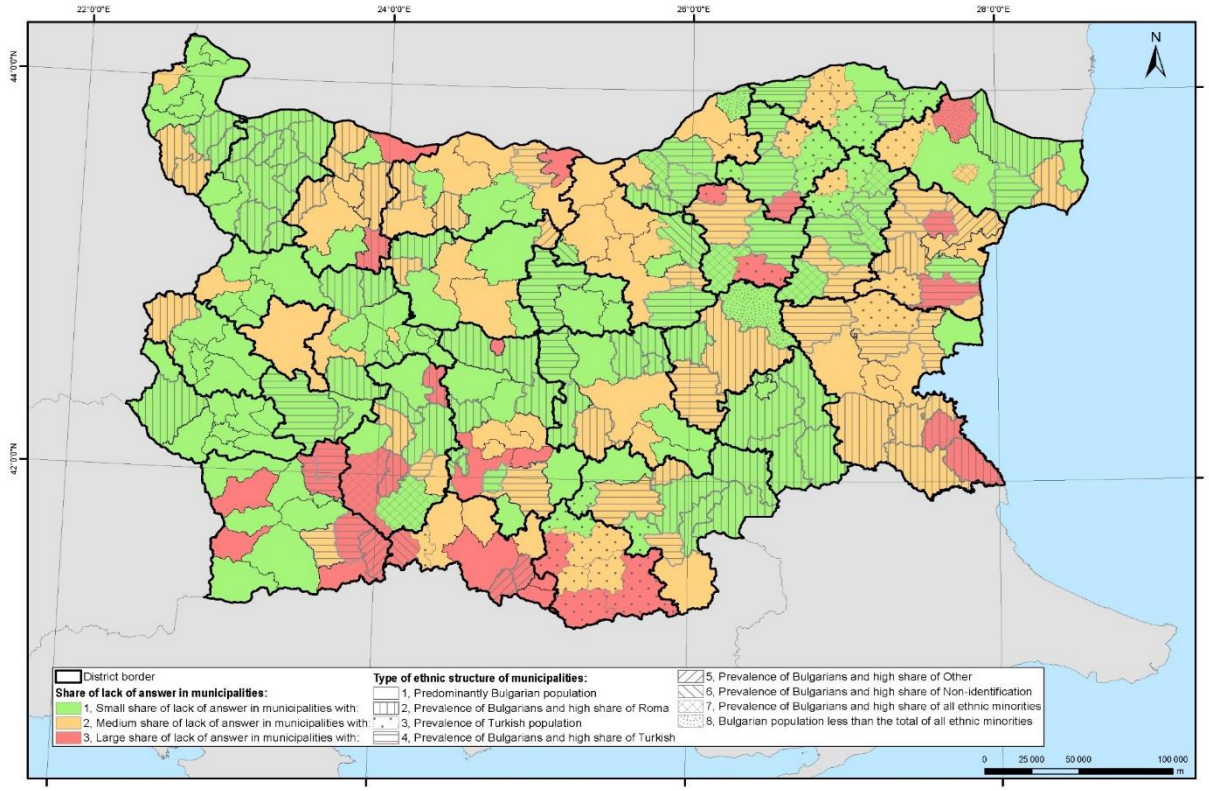
		Intervals				
Relative share	**Range**	**Bulgarian**	**Turkish**	**Roma**	**Other**	**Non-identification**
Very low	2SD+ below Mean	0 – 30.9%	-	-	-	-
Low	1SD-2SD below Mean	30.9% – 54.1%	-	0 – 0.8%	-	-
Below the average	0-1SD below Mean	54.1% – 77.3%	0 – 14.6%	0.8% – 7.2%	0 – 0.8%	0 – 1.1%
Above the average	0-1SD above Mean	77.3% – 100.0%	14.6% – 37.0%	7.2% – 13.6%	0.8% – 2.0%	1.1% – 2.6%
High	1SD-2SD above Mean	-	37.0% – 59.4%	13.6% – 20.0%	2.0% – 3.2%	2.6% – 4.1%
Very high	2SD+ above Mean	-	59.4% – 97.1%	20.0% – 31.4 %	3.2% – 9.6%	4.1% – 15.6%

The above table shows that neither the intervals nor their limits have significantly changed compared with 2001, although the mean values (base: 264 municipalities) of non-response to the ethnic identity question is 9.5 percent (SD – 7.1 percent), and the maximum is 53.7 percent. That is why the codes of intervals for the respective ethnic group are the same as those based on the Census 2001 data. The resulting 80 (eighty) particular combinations for all 264 municipalities were processed in order to allocate them into one or another of the eight distinct types of ethnic structures already defined:

Table 6. Distribution of municipalities by ethnic clusters, Census 2011 (in comparison with Census 2001 – in parentheses)

Cluster#	**Type of ethnic structure of municipalities**	**Number of municipalities**
1	Predominantly Bulgarian population	115 (+10)
2	Prevalence of Bulgarian population and a high share of Roma	63 (-3)
3	Prevalence of Turkish population	27 (-4)
4	Prevalence of Bulgarian population and a high share of Turkish	38 (+10)
5	Prevalence of Bulgarian population and a high share of Other	5 (+1)
6	Prevalence of Bulgarian population and a high share of Non-identification	5 (-6)
7	Prevalence of Bulgarian population and a high share of all ethnic minorities	8 (-4)
8	Bulgarian population less than the total of all ethnic minorities	3 (-2)

The changes in the numbers of municipalities falling within the frames of the separate clusters (numbers given in the parentheses) do not reflect the actual changes, as they are net values (municipalities both enter into and exit from a given cluster). The combination of ‘rate of non-response to the ethnicity question in 2011’, ‘ethno-demographic structure of the municipalities in 2011’ and ‘administrative differentiation of municipalities in 2011’ is visually presented in the following map; the map provides an idea of the substantial connection between ethno-demographic structure in 2011 and the scope of the ‘refusal of ethnic self-identification’ phenomenon in 2011.



Map 2. Geographic distribution of municipalities by share of lack of answer and by type of ethnic structure, Census 2011

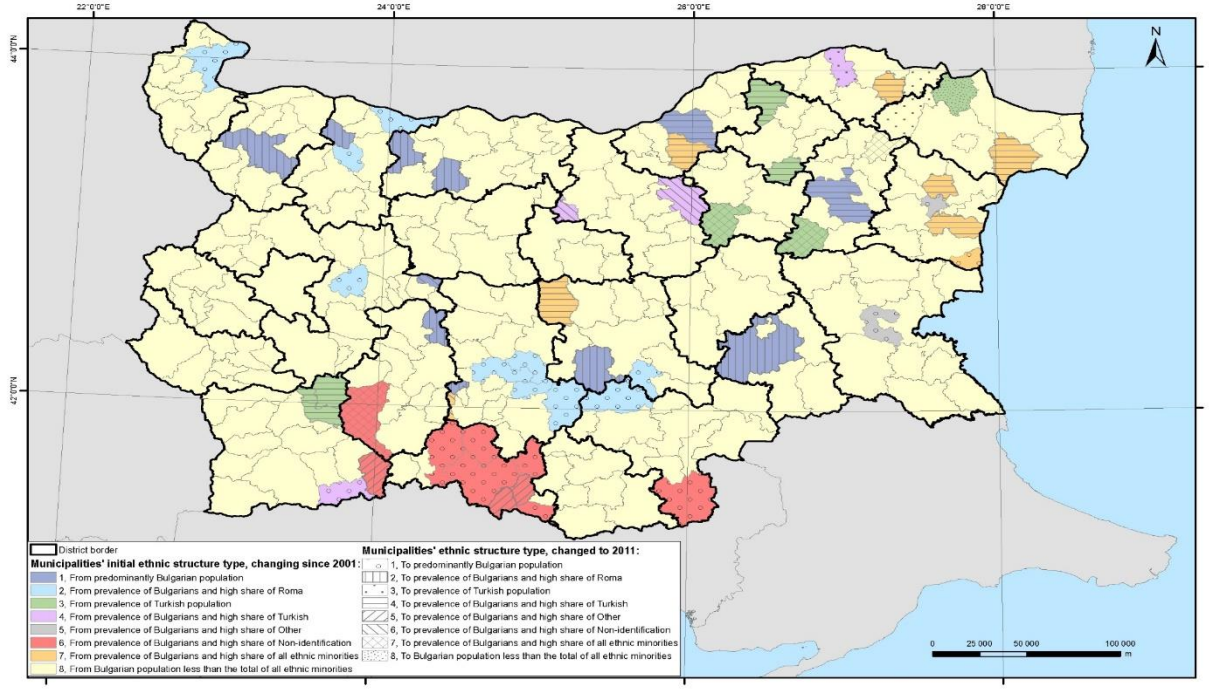
Since the changes that have taken place in the ethno-demographic structure in the municipalities merit special attention, in the following section of the text we will analyze those changes. Here, emphasizing the validity of the conclusions drawn so far, we should make the following corrections.

The correlation between clusters of non-response to the ethnicity question during Census 2011 and the eight clusters of ethno-demographic structures in the municipalities in that same year is of a weaker degree of significance than the correlation between ethnic configurations for 2001: Cramer's V is 0.23. Of course, the

particular dependencies within the frameworks of the ethnic structures themselves are also weaker. This is due to the mutual complementarity of the two variables in 2011, when 'ethno-demographic structure' is a projection only of those people who had identified themselves in the respective municipality, and not the 'real' picture, which remains unknown, at least in municipalities with a considerable (medium and high) percentage of non-response.

4.6. Analysis of changes of ethnic structure at municipality level between 2001 and 2011 and its relation with the lack of ethnicity identification in 2011.

The net amount of the number of municipalities that have their ethno-demographic structure changed between the two censuses is 56 (fifty six). Existing correlation between the amount of the shares (large, medium, small) of non-response on the ethnicity question and the change (presence or absence) in the ethnic configuration (Cramer's V: 0.23) is dependent only on the difference 'Large share of lack of answer – change in the ethnic structure'. With an average value of the high share of 12.5 percent of the non-respondents (in the aggregate), it is more than two times higher (26.8 percent) in municipalities in which we register a changed ethno-demographic structure. On the other hand, however, this means that in about 14 percent of them only (i.e., in, approximately, eight out of 56), the changed ethnic composition is a result of the effect of the distortion of ethnic picture due to the large size of the absent ethnic identity. The nature of this change is graphically depicted in the following map:



Map 3. Geographic distribution of municipalities by changes of ethnic structure type, Census 2011 compared to Census 2001

These changes can be basically described as a transition between two different ethno-demographic clusters, more specifically:

Eleven municipalities (out of 105) with predominantly Bulgarian population in 2001 changed as follows:

- Eight of them changed into municipalities with prevalence of Bulgarian population and a high share of Roma in 2011;
- Two of them changed into municipalities with a prevalence of Bulgarian population and a high share of Turkish population in 2011;
- One changed into a municipality with prevalent Bulgarian population and a high share of non-identification in 2011.

Eleven municipalities (out of 66) with predominantly Bulgarian population and a high share of Roma in 2001 changed as follows:

- Ten of them changed into municipalities with predominantly Bulgarian population in 2011;
- One changed to a municipality with predominantly Bulgarian population and a high share of Turkish population in 2011.

Seven municipalities (out of 31) that had had a prevalently Turkish population in 2001 changed as follows:

- Four of them changed into municipalities with prevalently Bulgarian population and a high share of Turkish population in 2011;
- Two of them changed into municipalities with prevalently Bulgarian population and high share of the sum of all ethnic minorities in 2011;
- One changed into a municipality with Bulgarian population amounting to less than the total of all other ethnic minorities in 2011.

Four municipalities (out of 28) with prevalently Bulgarian population and high share of Turkish population in 2001 changed as follows:

- Two of them changed into municipalities with prevalently Bulgarian population and a high share of non-identification in 2011;
- One changed into a municipality with predominantly Bulgarian population in 2011;
- One changed into a municipality with predominantly Turkish population in 2011.

Two municipalities (out of four) with predominantly Bulgarian population and a high share of other ethnic groups in 2001 changed into municipalities with predominantly Bulgarian population in 2011.

Ten municipalities (out of 11) with prevailingly Bulgarian population and a high share of non-identification in 2001 changed as follows:

- Six of them changed into municipalities with predominantly Bulgarian population in 2011.
- Three of them changed into municipalities with prevailingly Bulgarian population and a high share of other ethnic groups in 2011.
- One changed into a municipality with Bulgarian population and a high share of the sum of all ethnic minorities in 2011.

Eight municipalities (out of 12) with predominantly Bulgarian population and a high share of all ethnic minorities in 2001 changed as follows:

- Six of them changed into municipalities with prevailingly Bulgarian population and a high share of Turkish population in 2011.
- Three of them changed into municipalities with prevailingly Bulgarian population and a high share of other ethnic groups in 2011.
- One changed into a municipality with predominantly Bulgarian population in 2011.
- One changed into a municipality with prevailingly Bulgarian population and a high share of non-identification in 2011.

Three municipalities (out of five) with Bulgarian population amounting to less than the total of all ethnic minorities in 2001 changed as follows:

- Two of them changed into municipalities with prevailingly Turkish population in 2011.
- One changed into a municipality with prevailingly Bulgarian population and a high share of the sum of all ethnic minorities in 2011.

The most important of these changes relevant to medium and high percentages of non-response to the ethnicity question are concentrated in municipalities which in 2001 had an ethnic structure in which the Bulgarian ethnic group was numerically predominant, but with a high share of refusals to indicate ethnic identity (people who chose the answer option ‘I do not self-identify myself’). First, nearly all of these municipalities (ten out of eleven) proved to have a changed ethno-demographic structure in 2011, of which there is no comparable case among the changes in the other seven clusters. The most important finding is that six municipalities have a high share of non-response, and four have a medium share; this is quite different from the average proportion of small, medium and large shares of non-response (5:3:1) in the population under study.

In analyzing the patterns of change between 2001 and 2011, we may assert with certainty there is *a permanent trend of problematic self-identification regarding*

ethnic origin, a problem that is manifest in the declared refusal to indicate one's self-identification (in 2001) and the lack of an answer to the ethnicity question (in 2011). The 'escape' from ethnic identification of a considerable part of the population in these municipalities through non-response in 2011 has distorted the data on ethno-demographic structure of the municipalities, for the refusal *is not distributed between several ethno-demographic groups but is concentrated in only one of the groups*.

As supplementary analysis has shown, these are primarily those municipalities in the district of Smolyan that have a compact population with a specific ethno-cultural and religious background – ethnic Bulgarians of Muslim religious affiliation (the so-called 'pomaks', as mentioned above).

Our analysis has shown that a similar pattern of relation between 'change of ethnic configuration' and 'non-response' is also found in part of the municipalities with predominantly Bulgarian population and a high share of other ethnic groups in 2001. They changed into municipalities with predominantly Bulgarian population in 2011 due to the medium share of non-response in 2011, which corresponds to those respondents in 2001 who declared ethnicity other than the main three ethnic groups (Bulgarians, Turks and Roma). This pattern exists also within a small part of municipalities with prevalently Turkish population in 2001, but which changed into municipalities with prevalently Bulgarian population and a high share of Turkish population in 2011 due to the fact that many of those who in 2001 had declared themselves being of Turkish ethnicity gave no response in 2011.

This result once again confirms the validity and relevance of the conclusion indicated further above in this text, and complements it; the conclusion is that one of the factors of non-response to the ethnicity question in Census 2011 is the problematic ethnic identity of the Bulgarian Muslims or 'pomaks' (entirely) and of members of other small however compact within the municipalities ethnic groups (other than the three main ones), as well as of a small part of those who have self-identified themselves as Turks in municipalities that had a predominantly Turkish population in 2001.

5. GENERAL CONCLUSIONS ABOUT THE NEWLY DISCOVERED FACTORS AND DEPENDENCIES.

On the basis of the conducted analysis, we may sum up the achieved results by giving the following generalized conclusions:

There is a dependency between specific sense of ethnic belonging and the absence of ethnic self-identification, expressed through refusal to answer the ethnicity question in Census 2011, however the relation between ethnic identity and lack of

ethnic self-identification is not a direct one but is mediated through the specific ethno-demographic structure within the respective municipality.

There is a dependency between administrative organization at district level and the lack of ethnic self-identification (i.e. the existence of a particular municipality within a particular district has an impact on the non-response rate), but, on the basis of the available data, it is not possible for the analysis to ascertain a substantial significance of this dependency.

One of the factors of non-response to the ethnicity question in Census 2011 is the problematic ethnic identity among several ethno-demographic groups, including: the Bulgarian Muslims or 'pomaks' (almost entirely); other small compact (within the municipality) ethnic groups (partially) other than the three main ones in the country; and those who had identified as Turks in municipalities with predominantly Turkish population in 2001 (a small portion of them).

As in any research venture, in verifying the formulated hypotheses in the course of the study additional new research assumptions sprang up, the verification of which would go beyond the frame of the initially set tasks and/or the methods and data used in this study. In the course of the conducting of this analysis, new potential hypotheses were formulated, which could be the object of further research:

Since the municipality level proved to be a heuristic choice of research focus with regard to the problem we are analyzing, the following step would have be studying other socio-demographic parameters, for instance, conducting a comparative analysis of the age and education structure of those who have and those who have not answered the ethnicity question at municipality level.

A promising research question is whether there is a significant difference between ethnic self-identification and the lack of such self-identification at the personal individual level, on one hand (which we have followed here so far), and at household level, on the other hand. Such perspective is worth studying, for, in fact, the technique of the Census predetermines that the facts are registered through the head of the household for the entire household (including questions about ethno-cultural identification). Our assumption is that the different ethnic situations of the household make-up (size of household, number of children, number of elderly people) are characteristics that may possibly provide an opening to a deeper understanding of the dependency between ethnic belonging and lack of ethnic self-identification at municipality level.

A more extensive study of the factors of non-identification of ethnicity would have to include other elements of ethno-cultural identification too, such as mother language and religion; this is a matter for future research work that holds the potential of yielding even more heuristic results.

References

1. Arkadiev, D. (2012). Some problems in 2011 Census data of ethnic composition of Bulgarian population. *Statistics, Information Technology and Communications*. UNWE University Press, Sofia (in Bulgarian)

2. Blagoev, D., Haralampiev, K. (2014). In Search of Lost Ethnicity: Research Problems and Solutions. *Vanguard Scientific Instruments in Management 1*(8), 16-50. (in Bulgarian)

3. IBM Corp. (2010). *Clustering and Association Models with IBM SPSS Modeler*.

4. Ivanov, M. (2012). The 'Pomaks' According to Bulgarian Ethno-demographic Statistics. *Population*, *1-2*, 163-197 (in Bulgarian)

Acknowledgements

We express our gratitude to Mr. Ivan Balev and Mrs. Galya Stateva, directors in the National Statistical Institute, for the data they provided in electronic format. We express also our thanks to IBS Bulgaria, IBM Premier Business Partner in Bulgaria, for the software they provided. We extend our special thanks to Mrs. Lyudmila Boyadjieva and Mrs. Aglika Georgieva from GeoMarine Centre ltd., Sofia, for the preparation of the three maps used in the paper.