Quantifying ethnic populations without a census: methodological lessons from Ukraine

Patrik Tátrai
Geographical Institute,
HUN-REN Research Centre for Astronomy and Earth Sciences,
Budapest, Hungary
E-mail: tatrai.patrik@csfk.org

József Molnár
Department of Geography and Tourism,
Ferenc Rakoczi II Transcarpathian Hungarian University,
Berehove, Ukraine
E-mail: molnar.jozsef@kmf.org.ua

Ágnes Erőss
Faculty of Economics and Management,
Free University of Bozen-Bolzano,
Italy
E-mail: agnes.eross@unibz.it

Katalin Kovály
Geographical Institute,
HUN-REN Research Centre for Astronomy and Earth Sciences,
Budapest, Hungary,
Department of Geography and Tourism,
Ferenc Rakoczi II Transcarpathian Hungarian University,
Berehove, Ukraine
E-mail: kovaly.katalin@csfk.org

Keywords:
ethnic data collection,
ethnicity,
survey data,
categorization,
ethnic groups,
Ukraine

This study introduces a methodology that may be used to perform a head count of an ethnic group and collect detailed ethno-demographic data. Due to the lack of ethnicity questions in censuses or in the absence of a census, this is a real gap in the knowledge internationally. This study focuses on Ukraine, presenting a new methodology for estimating the basic demographics of ethnic Hungarians. The research was based on a large-scale questionnaire survey conducted in 2017, covering all municipalities in Western Ukraine, which hosts a significant Hungarian minority. According to our model, in 2017, the number of Hungarians in Transcarpathia most likely ranged between 125,000 and 135,000.

Two crucial methodological issues challenge the validity of data: the varying interpretations of temporary and permanent migration may lead to significant differences in the population size surveyed, while the situational self-identification of individuals, mostly of outlier categories such as the Hungarian-speaking Roma, may increase the uncertainty of results. Nevertheless, we argue that in the context where traditional censuses are being replaced by censuses that do not collect ethnic data, our model might fill the gap; however, its applicability to other minorities depends on the studied group's particular structural characteristics.

Introduction

The answer to "How can one headcount an ethnic group?" is usually simple: a population census. When it comes to "measuring ethnicity," most of the literature employs census methodology and data. However, what if no recent census data are available, and probably none will be in the near future? In this case, a research project that collects the missing data could fill the gap.

This is the case, for example, in Ukraine, where the last census was taken in 2001 and no reliable survey/census has been conducted to trace ethnic and demographic processes since then. Two revolutions and a war have occurred during this period, transforming everyday life and coping strategies and causing substantial population movements. Data sources are limited, not only on ethnic structure, and in the context of "high rates of migration, including refugee and labor migration, the accurate measurement of the population remains a challenge" (Chugaievska et al. 2023: p. 215).

The question "How many people are affiliated with an ethnic group?" may arise in several contexts, from the academic sphere through public policy, institutional use, linguistic rights, and direct political thematization to the self-image of the minority itself. Statistical data are an essential reference point in various discourses on the rights, education, and political and economic situation of any minority. This is, of course, not by chance: The statistical size of an ethnic group affects ethnic hierarchy and the possibilities for minority advocacy both locally and at a national level, and, in some cases, may affect the political autonomy and possession of various resources, i.e., power relations in general (Erőss et al. 2021). All this underlines the claim that ethnicity is not a simple statistical category but rather, an assigned status, a power position (Brubaker 1996).

The unequal power position of ethnic groups is generally justified by counting populations. The primary tool for this is the census, through which the nation-state categorizes people to legitimate the privileged position of the titular ethnic group and maintain ethnic hierarchy and asymmetry (Kertzer–Arel 2002). In addition to the purpose of counting, the subject of counting (i.e., ethnicity as a statistical category) must also be clarified. A large body of literature emphasizes that measuring ethnicity is rather controversial mainly due to its ambiguous conceptualization and related issues of categorization and ethnic boundary-making (e.g., Aspinall–Song 2013, Brown–Langer 2010, Elrick–Schwartzman 2015, Mateos 2007, Siegel 2018, Simon–Piché 2012, Supik–Spielhaus 2019, Surdu 2016, Williams–Husk 2013, Wimmer 2013).

Although there is no agreed-upon definition of ethnicity, most social scientists have adopted a constructivist approach, wherein ethnicity is taken as a social construct. As such, it cannot be described using simplistic statistical categories. Indeed, single-word ethnic affiliations ("German" or "Roma") are an invention of nation-states aimed at facilitating the exercise of control over otherwise heterogeneous societies (Kertzer–Arel 2002). This corresponds to the modern statistical imagination, which uses single, mutually exclusive categories to describe

cultural identity and people's ethnicity (Anderson 1991). Nevertheless, such categories do not always align with the situational, fluid, overlapping ethnic identities of people in daily practice and do not reflect the complex heterogeneity found within groups (Yao et al. 2022). Furthermore, they are always geographically and chronologically contextual (Jenkins 1994, Mateos 2007, Peach 2000, Simon–Piché 2012)–similar to official statistical concepts/definitions of ethnicity. To Brubaker, ethnic statistics and official ethnic classifications "*construct* and *constitute* the groups they ostensibly *describe*" (Brubaker 2009: p. 33. emphasis in the original). Consequently, as an instrument of power, the census not only reflects social realities but also participates in constructing them; thus, its use can be unambiguously considered a political act (Kertzer–Arel 2002).

In addition to fuzzy conceptualization, ethnicity is difficult to measure because of unclear boundaries that are often fluid and blurred at the micro-level (Barth 1998) but may be persistent at the macro-level (Wimmer 2013). Ethnic group membership can vary depending on which particular criteria are adopted as markers of ethnicity: ethnic nationality, political nationality, language, religion, etc. Moreover, measuring ethnicity is strongly influenced by the data collection method, i.e., whether it is based on self-definition or external ascription (Simon 2012).

However, despite the problems and inconsistent means of collecting ethnic data,[1] there is a demand for the recognition and statistical capture of the world's growing ethnic diversity (Simon–Piché 2012). All the more so because counting may be used not only to subordinate, exclude, and discriminate but also to serve the more socially just purposes of eliminating discrimination (Blum–Guérin-Pace 2008, Rallu et al. 2006), fighting actual or perceived inequality, or legitimizing an ethnic group (Williams–Husk 2013). Furthermore, in the context of national minorities in Eastern Europe, counting may be seen as a precondition of institutionalized power-sharing and autonomy (Kiss 2018). Nevertheless, similar to most European countries that collect data on ethnicity, Ukraine also counts to homogenize and dominate (Arel 2002), despite not having counted since 2001.

Of the countries that used to collect census ethnic data, Ukraine is not the only one lacking the respective statistics. Some Central and Eastern European (CEE) countries (such as Austria and Slovenia) have switched from traditional censuses to register-based censuses, which do not include ethnic data (Josipovič 2015). Several other countries are expected to either switch to a register-based census or no longer collect data on ethnic affiliation in their censuses (Skinner 2018). This results in a severe gap in knowledge about the demographic behavior and dynamics of ethnic

[1] Some argue that the statistical representation of ethnicity should be understood as a social practice in a social context that involves applying a well-defined and methodologically sound framework (Supik–Spielhaus 2019), whereas others are more interested in research than of reducing inconsistencies in data collection (Burton et al. 2010).

groups and ethnic populations,[2] thereby eroding the visibility and advocacy of such groups and their members.

Considering the above factors and the existing data gaps, the present study aims to present a method for potentially doing a head count of an ethnic group and collecting detailed ethno-demographic data comparable with previous censuses. Beyond developing a methodology and analysing its applicability, the research was designed to implement a process of voluntary and participatory counting as an alternative to the census. In particular, the research carried out within a pilot project called "Summa 2017" (launched in 2017) offers primary data on the number, population dynamics, and other demographic features of ethnic Hungarians in Transcarpathia. Although the study focuses on a specific ethnic group, we are convinced that the method we describe can also be applied to other ethnic groups.

Transcarpathia, the westernmost region of Ukraine bordering Romania, Hungary, Slovakia, and Poland, has experienced several border changes over the last century. Ethnic Hungarians, an autochthonous population, were entitled as the titular ethnic group until the end of World War I, when their number peaked at 184,287. Since then, in a minority position, their number declined to 151,516 (12.1% of the region's population) by 2001. Most live in a compact settlement area along the Ukrainian–Hungarian border. What makes the Transcarpathian Hungarians an ideal subject for such a quantitative survey is their relatively small number, their high level of institutionalization, and their relatively clear ethnic group membership, primarily based on cultural criteria (mainly language) coupled with a solid local identity (Veres 2015). According to the 2001 census, 97% of the ethnic Hungarians in Ukraine stated that Hungarian was their native language. Of the 163,000 people who self-identified as ethnic Hungarians and/or native speakers of Hungarian, 147,000 (90%) marked both of these categories, meaning that their ethnic and linguistic boundaries basically coincide.[3] Although individual ethnic boundary crossing has been common over recent decades, the rate of ethnic exogamy has remained quite low, and the intergenerational transmission of ethnic identity in mixed families has remained balanced (in Ukrainian–Hungarian relations) (Molnár 2016). This implies that, at the macro-level, ethnic boundaries are fairly stable. However, there are a few outlier categories such as the Hungarian-speaking Roma, descendants of ethnically mixed marriages mainly in urban areas, and the Ukrainian–Hungarian bilingual Greek Catholics in some rural localities.

[2] The lack of (up-to-date) ethnic statistics may open the door to speculation by decision-maker politicians and the local population in general. For example, in Transcarpathia, Ukraine, a survey conducted in 2016 reported that Hungarians in Transcarpathia estimated their population share to be, on average, twice as large as in the 2001 census, while Ukrainians estimated it to be one and a half times larger (Ferenc–Rákóczi 2018).

[3] The above values are particularly large compared to, for example, those for Poles, Belarusians, and Greeks in Ukraine, 85%–95% of whom claim a different native language and ethnicity, mainly due to the lack of mother-tongue education (Csernicskó–Kontra 2023).

In what follows, first, we review previous attempts at counting ethnic groups, focusing on methodological approaches that could be used in the present study. Then, we present a detailed methodology highlighting the key issues, such as sampling and surveying. We then give an overview of and contextualize the main results of the research. We conclude by discussing the main lessons learned from our survey methodology and, in general, how widespread ethno-demographic phenomena such as migration and situational ethnic self-identification challenge the counting of ethnic populations.

Previous attempts at counting ethnic populations

Ethnic data collection outside of censuses is usually linked to situations in which the size of a group is disputed for some reason. In Central and Eastern Europe, this is typical of the Roma. The number of ethnic Roma according to self-identification in censuses has always been much lower than that estimated by experts (Kemény–Janky 2005, Ladányi–Szelényi 2006, Messing 2014, Simon 2012) due to exclusion, racialization, stigmatization, and discrimination (Ahmed et al. 2007, Csata et al. 2020, Rughiniş 2011, Surdu 2016). As a consequence, the statistical size of the Roma provided by censuses was often considered unreliable. This gave rise to surveys that aimed to measure their number and socio-demographic characteristics since the nineteenth century. Most of these used a system of external categorization by local experts who interacted with the Roma (local government officials, teachers, social workers, etc.) or by interviewers (for example, Kemény–Janky 2005). Others combined the methods of external classification and self-identification by respondents (Ladányi–Szelényi 2006). However, only a few were able to cover all municipalities/settlements of a country or region (Horváth–Kiss 2018, Mušinka et al. 2014, Pénzes et al. 2019). All of these surveys collected data on the Roma from local governments using questionnaires.

Unlike the case of the Roma, the undercounting of other, "non-visible" ethnic populations are rarely suspected. As a result, there have been very few large-scale (non-census) efforts at ethnic data collection in CEE countries. One of these is related to the 1980 census in Hungary, when official data showed that the number of non-Hungarians had dropped dramatically. A supplementary survey was carried out to assess the number of people with minority ties, which employed external classification by local experts (Hoóz 1985). However, because the network of personal contacts of the informants was limited, the method could only be used in settings with < 2,000 inhabitants. In recent decades, similar research has been conducted on a local scale, focusing on ethnic classificatory systems and assimilation patterns rather than on actually measuring ethnic populations (Erőss et al. 2021, Keményfi 1998).

Due to the similarity of the study areas and partly due to the similarity of the research objectives, another study should be highlighted. Just before the 2001

Ukrainian census was implemented, Kocsis (2001) published an ethnic map of Transcarpathia based on the data (estimations) provided by local governments. Although this methodology can be used to estimate ethnic population numbers, it does not offer information on the socio-demographic composition of the population under study. Furthermore, after 2014, population mobility in Ukraine (and in Transcarpathia) had increased to such an extent (Malynovska 2021) that no reliable data on the population at the settlement level was even available. Therefore, our study could not adopt this method.

The literature offers some other approaches to counting ethnic group membership. Name-based ethnicity classification has been used relatively often, mainly in the case of immigrant groups in western countries (Mateos 2007). However, the prerequisites for this method are (1) the quite sharp ethnic boundary of the group under study, caused either by racial exclusion (such as the Asian populations in the UK) or religious segregation (e.g., Jews and Muslims) and (2) access to updated population registers listing the full names of individuals. In the case of Ukraine, neither of these conditions is satisfied: after hundreds of years of coexistence, permeable ethnic boundaries facilitate intermarriage and ethnic mixing; thus, surnames are not appropriate indicators of ethnicity, while forenames are often transcribed into the Ukrainian variants. With regard to the present research, the second criterion is perhaps less well satisfied as we could not identify an up-to-date register of the population of Transcarpathia.

In some cases, religion is used as an indicator of ethnicity. For example, in an attempt to map Jews in London, Waterman–Kosmin (1986) used both synagogue membership and names as proxies to count Jewish people. Unfortunately, we could not adopt this method because Transcarpathian Hungarians belong to three major and some smaller religious denominations, the denominational and ethnic boundaries in the region do not fully overlap, and church registers in larger towns only cover churchgoers.

National surveys can also serve as a source of estimates of minority ethnic populations. For instance, in the UK, labour force surveys are used for such purposes (Scott et al. 2001). However, such surveys are not available in Ukraine, and the Hungarian population would need to be larger to embody a representative sample within a national survey.

In our increasingly digitized world, there are new ways of counting. A recent study attempted to measure ethnicity in ethnically mixed settings in the Romanian–Hungarian border zone based on the language use of individuals with Instagram profiles (Tóth–Papp 2020). However, the results are not convincing due to the small number of cases available in small municipalities and problems with sample representativeness – e.g., younger individuals were strongly overrepresented.

The main conclusion from the literature review is that current established methods are not applicable in our case. Up-to-date census and survey data and basic statistics

such as population numbers by municipality have yet to be made available or reliable in post-Euromaidan Ukraine. What we did have access to, however, were the results of the 2001 census, which showed essentially the same patterns of ethnic group distribution as previous censuses in Transcarpathia. Consequently, when we developed our pilot project to estimate the number of Transcarpathian Hungarians and their demographic characteristics, we could only rely on the 2001 census as a starting point.

Methods

Sampling

Our research objectives were best achieved through a large-scale, representative questionnaire survey. The survey, covering the majority of Hungarian settlements,[4] was carried out in 2017 with the help of interviewers. When designing the survey, we drew primarily on the ethnic data by settlement from the 2001 Ukrainian census (Figure 1). The census confirmed that Hungarians still live in the settlements where they lived a hundred years ago. Considering this relative stability and the geographical concentration of Hungarians, we limited the survey to settlements with a Hungarian population of >100. We identified 113 such settlements, where 98.7% of Transcarpathian Hungarians had lived in 2001.

For the next step, 312 sampling sites were designated in the selected settlements and the number of sampling sites per settlement was determined according to the distribution of the Hungarian population based on the 2001 census. Accordingly, on average, one sampling site was selected per 500 Transcarpathian Hungarians (according to the 2001 data). Each sampling site comprised 20 households inhabited by at least one person self-identifying as Hungarian.

Due to the different settlement patterns of Hungarians within Transcarpathia, it was not possible to apply a uniform sampling method across the sampling sites. Therefore, three different methods were assigned to the three different types of Hungarian settlement, as follows:

(1) Most of the Hungarians were living in rural settings and small towns along the Hungarian border, forming the local majority population in 2001 (i.e., ethnic block settlements). In this area, cluster sampling was applied. Each cluster[5] comprised at least 20 geographically contiguous households.[6] The number of households surveyed

[4] In this study, we use the term "settlement" as the basic territorial level at which census data are published. In Ukraine, this does not fully correspond to the local administrative units. For instance, at the time of the 2001 census, there were 609 settlements in 330 local municipalities in Transcarpathia.

[5] In the case of cluster sampling, a sampling site is the same as a cluster.

[6] Since the number of Hungarians in 2017 was calculated based on the 2001 census (see later for details), the population dynamics of Hungarians between 2001 and 2017 could only be estimated by analysing the respective processes of the whole cluster, hence the need to select geographically contiguous households.

could vary; the main criterion was 20 households made up partly or fully by Hungarian inhabitants. Until this number was reached, all households (including uninhabited houses and vacant plots) had to be surveyed. The starting point (the first household) and the boundaries of the clusters were selected by the locally based interviewers and the research leaders (1) to avoid overlapping clusters, (2) to ensure that the cluster definitely included 20 Hungarian households, and (3) to avoid interviewers omitting households they did not want to survey, while still bearing in mind the basic criterion of selecting geographically contiguous households. Cluster sampling may sometimes lead to sampling bias, particularly in cases with high population heterogeneity and segregation. To the best of our knowledge, the social differences in the villages selected were not so great as to lead to significant intra-village segregation, thereby causing sampling bias. The only exception was the widespread residential segregation of the Roma population. In Transcarpathia, as in other CEE regions, the Roma face discrimination, stigmatization and marginalization; social rejection of the Roma is widespread among local populations (Stryapko 2016). As a result, several Roma-segregated settlements were found throughout the border region under study, mainly on the peripheries of villages and towns. Here, poverty is linked to ethnic exclusion; and despite the Hungarian mother tongue of most of the residents, ethnic identification is highly situational (but only few people claim Roma identity in official contexts to avoid the Roma stigma). When the time came to conduct the survey, we encountered the problem that our interviewers were distrustful of these settlements and their marginalized residents.[7] For this reason, we were constrained to exclude segregated settlements from the survey to avoid sampling biases. This does not mean that no Roma were included in our sample, as those who did not live in segregated settlements were eligible for inclusion in the study.

(2) Based on the 2001 census, a tenth of the Hungarians were residents of two middle-sized towns (cities on a Transcarpathian scale: Uzhhorod and Mukachevo), where they made up 7% and 9% of the total population, respectively. Surveying them was much more challenging. After a few unsuccessful attempts, we found a register comprising full names and addresses of individuals living in Ukraine. Although the essential attributes of this dataset (e.g., a reference date) were unknown, after checking the data in the list, we estimated the date of the register as the turn of the millennium; thus, it seemed to be an appropriate proxy for the population of the two cities at the time of the 2001 census. We compiled a list of typical Hungarian surnames from this dataset, and households for surveying were randomly selected from that list. This sampling method can result in two types of bias. (a) Respondents with Hungarian names but non-Hungarian identities are included, but this is not a problem as we do not take them into account in the calculations. (b) Residents with nontypical Hungarian names but Hungarian identity are omitted. This could cause problems in

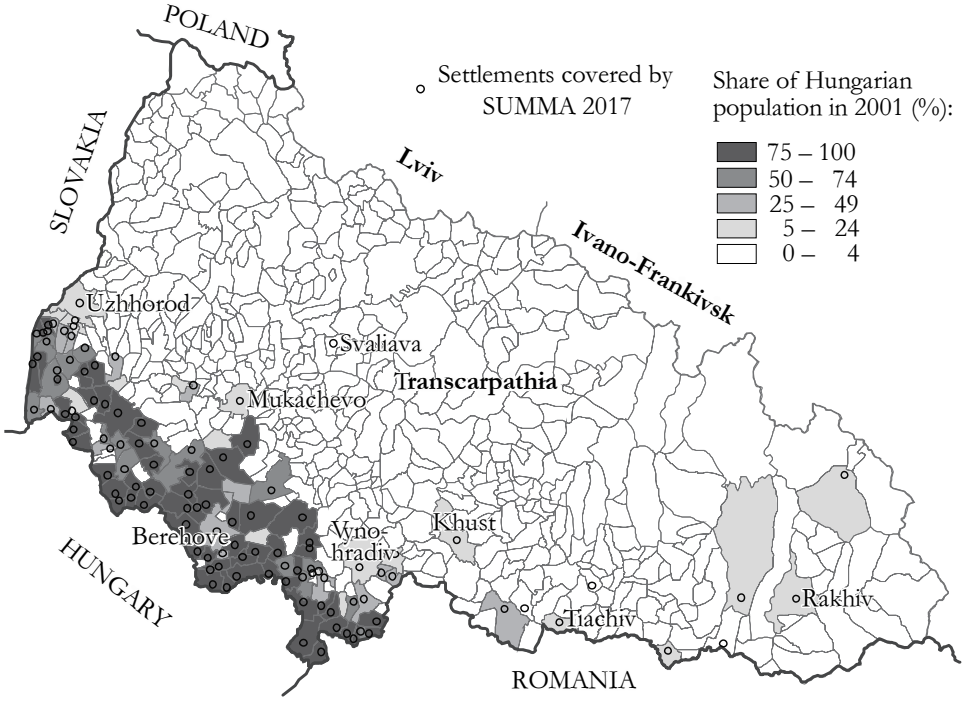
[7] Similar phenomena can be observed in other CEE contexts (Mirga-Wojtówicz–Fialkowska 2022).

two cases: first, if the demographic, migration, and assimilation behaviors of residents with nontypical Hungarian names differ significantly from those of residents with typical Hungarian names, but we have no reason to assume this. Second, if this method results in a smaller sample size but does not in itself lead to systematic bias, it will only slightly reduce the reliability of the estimate.

(3) The remaining 14% of Hungarians lived scattered outside this ethnic block settlement, and their local proportion of the total population was typically <10% in 2001. Their low settlement concentration did not allow us to use the cluster method, and due to the small number of local Hungarians (in many cases with non-Hungarian names), the second method, used in cities, was not practical either. Instead, we introduced a hybrid method that we call the "areal snowball." In this case, the "snowball" approach meant that the interviewers did not survey a predefined neighborhood but always the closest household inhabited by Hungarians.[8] Twenty Hungarian households, not necessarily geographically contiguous, constituted each sampling site.

Figure 1

Settlements included in the survey according to the local proportion of ethnic Hungarians in 2001



[8] To ensure comparability with the 2001 census, we defined Hungarian households as those that either comprised at least one Hungarian in 2001 or at the time of our survey. The location of the nearest Hungarian neighbor was obtained either from the interviewer's knowledge or from the respondent.

Questionnaire design

Due to the large number of households to be surveyed, the questionnaire had to be simple and concise. Two questionnaires were prepared: one for households with at least one Hungarian member ("Hungarian questionnaire") and one for uninhabited households and households with no Hungarian member (uninhabited/non-Hungarian questionnaire). The number of questions was limited to 15 and 8, respectively, to minimize the time spent per household and maximize the response rate. The Hungarian questionnaire included five groups of questions (Table 1).

Table 1

Structure of the Hungarian questionnaire

Topics	Questions				
Changes in individuals' life course between 2002 and 2017	No change/born/ died/moved in/moved out	Year of last move	Year of death	Place of residency on 1 Jan 2002	Current place of residence
Basic demographics	Kinship with informant	Year of birth	Sex		
Identity issues	Ethnic affiliation (two possible answers)	Mother tongue (two possible answers)	Religious affiliation		
Socio-economic status	Highest level of education completed	Current main occupation			
Temporary migration	Number of months spent abroad in 2016	Target country			

The first set of questions was designed to measure changes over the period since the last census, allowing us to link the current data with the status of individuals at the time of the 2001 census.[9] We decided to measure ethnicity through ethnic affiliation and mother tongue, thereby allowing the respondents to claim multiple identities (based on self-declaration). In many cases, multiple ties reflect everyday realities and are congruent with individuals' self-identification rather than imposed exclusive categories (Yao et al. 2022). However, it is also clear that the possibility of declaring multiple identities makes it difficult to make direct comparisons with data sources where this option was not available (e.g., the census). Socio-economic status (SES) was measured through education and occupation. However, it should be noted

[9] Indeed, we decided to ask about residency as of January 1, 2002, a date which is simpler to apply than the official date of the census (December 5, 2001) but still close to it. If the residency/status of any member of the household had changed between 2002 and 2017, it was required to be noted here, including births, deaths, immigrations, and any emigrations.

that in the Transcarpathian context, the correlation between SES and the main occupation is not necessarily strong, as a significant proportion of people have multiple jobs and the role of the informal sector and the shadow economy is very significant (Vinnychuk–Ziukov 2013). Partly as a consequence, temporary migration (primarily working and studying abroad) has been widespread among Transcarpathians (Çağlar 2013). Thus, two questions regarding migration were included in the questionnaire that addressed the length and place of stay abroad during 2016 (if this exceeded one month).

The questionnaire used to record data about uninhabited or non-Hungarian households in the sample area was even simpler. It collected data on the type of household (uninhabited houses/ethnic affiliation of current residents) if there were any Hungarian inhabitants in 2001, their number, and what had happened to them (died, moved away).

Survey execution

The survey was conducted in mid-2017. In total, 170 interviewers were recruited from the Transcarpathian Hungarian College of Higher Education in Berehove and the personal networks of the researchers. The interviewers were briefed in Berehove and Uzhhorod on May 29–30, 2017, when they were introduced to the purpose of the survey and the exact process of data collection. The data collection took place between June 1 and September 30. The interviewers asked available members of each household to self-report answers to the questions listed in Table 1. However, the respondents could also declare the data of other persons living in the household. Due to the widespread coping strategy of working abroad, in some cases, particularly in villages, the interviewers were allowed to collect data from neighbors if household members were found to be permanently unavailable. In the case of uninhabited households, data were provided by neighbors. While responses were voluntary, the response rate was exceptionally high due to the trust in the locally based interviewers, the simple and short questionnaire, and the importance of the issue to the community.

During (and partly after) the survey execution phase, we faced four main types of challenges and problems: sampling biases, a lack of available interviewers, the inadequate activity/reporting of interviewers, and inaccurate reporting by respondents. Sampling biases arose mainly in settings surveyed using the second and third methods. In the cities (second method), it soon became apparent that the typical Hungarian surnames selected from the register did not necessarily represent Hungarian individuals in the local context but after identifying the appropriate surnames with the help of the local interviewers, the method worked sufficiently. However, since the survey execution in cities required much more effort from the interviewers, the quality of the data depended largely on their own attitudes.

In towns with a small proportion of ethnic Hungarians, sampling bias also occurred. For instance, local interviewers in two towns deviated from the methodology ("areal snowball") and identified the sample through Hungarian institutional networks – such as families whose children attend Hungarian kindergarten or individuals who are members of the Association of Hungarian Entrepreneurs in Transcarpathia – resulting in systematic errors. Due to the late return of the questionnaires and the limited circle of available local interviewers in these settings, such sampling problems could not be resolved; thus, these two sites were not included in the remaining phases of the research. Other minor sampling problems also occurred in some locations, the most common of which was surveying < 20 Hungarian households per sampling site.

Beyond sampling problems, we were sometimes faced with the issue of incomplete or inaccurate questionnaires. For example, no deaths or migration events were registered in any of the 48 households surveyed in a town between 2002 and 2017. This may only be explained by the inappropriate activity of the interviewer, making the data unreliable. By contrast, other data on sex, age, etc., seemed precise and valuable.

Unlike the previous cases, measuring ethnicity and native language did not entail salient difficulties. However, we also experienced a phenomenon that several other studies (Telles–Lim 1998, Ladányi–Szelényi 2006, Ahmed et al. 2007) have identified: The person empowered to classify impacts the process and outcome of the categorization decisively. In our case, the Hungarian ethnicity of the interviewers may have influenced the ethnic self-classification of the respondents. For example, some interviewers reported that respondents with a Ukrainian ethnic background and poor Hungarian language skills (mainly in cities) claimed Hungarian (dual) ethnicity, presumably because they had already acquired preferential, nonresident Hungarian citizenship. However, these sporadic "reverse" identity shifts arose in a specific situation where the interviewer was Hungarian and the respondent perceived this as a formal Hungarian context, where citizenship and ethnicity are linked. These respondents would most likely not have self-identified as Hungarian in an official Ukrainian census, which is why we believe that these cases should not be interpreted as an inherent part of our demographic assimilation model.

We identified two other issues related to ethnicity when processing the questionnaires: the codes for individuals' ethnic identities and native languages were listed in ascending order in several questionnaires. Therefore, we were unable to identify the first and second ethnicity/native language in the case of multiple responses. This also reduced the comparability with census results. The second issue was related to terminology: Although the Hungarian term *nemzetiség* used in the questionnaire literally means "nationality," its meaning is better reflected by the term "ethnicity" or "national affiliation in the cultural sense." However, some interviewers understood it strictly as nationality (i.e., citizenship), reflecting a sometimes

unambiguous interpretation of this term. As a result, multiple (Hungarian and Ukrainian) ties were recorded in these predominantly Hungarian settlements, indicating that most locals may also have Hungarian citizenship.

Overall, the survey proved much easier to conduct in the Hungarian ethnic block settlement. However, in the cities and in the more scattered communities, where the second and third methods were used, we faced several difficulties. Most of them we managed to overcome, but some of them, mainly due to sampling errors and inaccurate data collection by the interviewers, we did not.

Data covering a total of 23,033 persons were recorded during the survey (further details in Appendix Table A1). At the beginning of 2002, the number of Hungarians living in the sample area was 19,993, representing >13% of the total of Transcarpathian Hungarians (as registered by the census in 2001). After 15 years, in mid-2017, there were 18,553 Hungarians in the sample, 6.6% of whom claimed multiple ethnic affiliations– in addition to the ethnic Hungarians and those with multiple ethnic affiliations, 658 Ukrainians and 125 other ethnicities (mainly Roma and Russians) were living in Hungarian households (Tátrai et al. 2018).

Calculation of the number of the ethnic Hungarian population

The changes in the population of an ethnic community are determined by three factors: natural changes, net migration, and net assimilation. These factors had to be quantified to estimate the population change of the Hungarian community in Transcarpathia for the period 2002–2017. Our method does not directly estimate the population of the Hungarian community based on the sample but rather calculates the change from a previous, well-known state, i.e., the 2001 census data, and calculates the number of the Hungarian population based on these changes.

For comparability with the 2001 census, we categorized dual-identity persons into one of the ethnic groups. Since local/district-level ethnic proportions are the most important factor that influences the self-definition of dual-identity persons in the event of a binary choice constraint (e.g., census), the data on dual-identity persons were weighted. The weights were set to reflect the direction and intensity of assimilation processes by districts. A good proxy indicator is the proportion of children from mixed marriages who declared Hungarian ethnicity in the 2001 census. For instance, in the Berehove district, the only district with a Hungarian majority in Ukraine (before administrative reform in 2020), 75.4% of those from mixed marriages self-identified as Hungarian in 2001. Accordingly, a weight of 0.754 was attributed to the data of individuals claiming dual identity in the survey.

Among the components of natural reproduction, the data on the number of births at most of the sampling points were found to be of good quality, so the birth rates derived from the survey results were used in our calculations. The few exceptions were those municipalities where the population was almost completely ethnically

homogeneous, and it was possible to obtain official birth data, which are – obviously – more accurate than from a sample survey due to their comprehensive nature.

However, our mortality data, in general, was found to be deficient compared to the birth data. This was due partly to subjective factors (the inaccuracy of interviewers' recording of data) and partly to objective issues (the complete absence of a family, thus the lack of a respondent to count). Therefore, when municipal mortality statistics were available, we preferred to use them – all the more so as the sporadically available mortality data by ethnicity showed no significant differences in death rates between the main ethnic groups.

Net migration had to be calculated based entirely on our survey, as the registration of migration in Ukraine – as in many other countries – is far from complete, being limited to officially reported movements. The migration rate was thus measured in the following two ways.

First, the difference between the number of people who moved out of the sample and the number who moved in was, as appropriate, proportional to the number of people in the sample. In theory, the number of people who had emigrated from one settlement in the sample to another and people who had immigrated from one settlement in the sample to another should have been the same. However, the estimation of the settlement number of emigrants is less accurate than the number of immigrants (think of complete families who left and thus may not be known to interviewers). In particular, we found that only 87% of out-migrants were recorded on average. To compensate for this systematic source of error, we divided the emigration data by 0.87, assuming that the unobserved migrants are similar to those recorded.

Second, out-migrants from one settlement are partly recorded as newcomers in other settlements, while some in-migrants to that settlement are recorded as out-migrants in other settlements. If these movements are not taken into account, valuable data are missed, which could improve the accuracy of the survey significantly. Therefore, the migration balance by settlement is also determined using the so-called "inverse" method. This was counted as the number of newcomers from other surveyed settlements in the one under analysis minus the number of out-migrants from the latter to other surveyed locations. Of course, it is not only from other Hungarian settlements in Transcarpathia that Hungarians could have arrived at the settlements of the sample but our analysis shows that 96% of them did arrive from the other settlements under study. Furthermore, they did not only move from the given settlements to other Transcarpathian Hungarian settlements; indeed, our analyses showed that only 44% of the migrants did so. Thus, to determine the actual migration balance using the inverse method, we divided the number of those who moved from other surveyed settlements to the given settlement by 0.96, and we divided the number of those who moved from the given settlement to other surveyed settlements by 0.44.

We estimated the migration balance in our model as the average of the two methods described.

The third factor affecting the size of ethnic populations is assimilation. Our survey was not designed to detect changes in ethnic identification over the individuals' life course (i.e., intragenerational assimilation). Although the other type of assimilation, intergenerational assimilation, can be measured through our survey, we did not include it in our model because the data on natural reproduction already contain it. Thus, regarding intergenerational assimilation, we drew on the literature, which, based on the 2001 census, estimated that although the degree and direction of assimilation varies across Transcarpathia, correlating strongly with the local/regional proportion of the Hungarian population, the overall assimilation balance is around zero (Molnár 2016). Considering the social changes and the intensifying kin-state politics of Hungary that have taken place since then, including the prestige of ethnic groups and languages (Tátrai et al. 2017), we assume that assimilation patterns remain unchanged, with possibly a slight increase in Hungarian assimilation gains.

Finally, we subtracted the mortality rates from the birth rates and added the migration rates. Thus, we obtained the population change rates of Hungarians by settlement for the period 2002–2017 and multiplied it by the number of Hungarians recorded in the 2001 census by settlement.

Control survey

How reliable were our processes of questionnaire-based data collection and calculations? This question is hard to answer; however, we tried to estimate it with a control survey carried out in July 2017. There were limited opportunities to do this due to the methodological constraints summarized in the literature review, but ultimately, we decided to conduct a control survey that applied external ethnic classification, similar to the survey that supplemented the 1980 Hungarian census (Hoóz 1985).[10] The control survey addressed the number and ethnic affiliation of the local population by household, as categorized by local experts (e.g., the mayor, local government employees, and parsons) who are in regular contact with local people. The two main criteria for selecting the localities were their geographical distribution in the Hungarian settlement area in Transcarpathia and population size, as this method only works reliably in small villages. However, this methodological constraint excluded locations where the most problems were encountered during the questionnaire survey, as these municipalities typically have large populations. Nevertheless, the comparison of the questionnaire survey and the control survey still provided valuable lessons.

As shown in Table 2, the results of the ethnic categorization of the questionnaire and control surveys were coherent for all selected sites even though they were

[10] For a more detailed methodological description, Erőss et al. 2021.

conducted using different methods (auto-identification vs. hetero-identification). However, this consistency is partly due to the fact that the control method was implemented in localities whose populations – with the exception of Hudia – were relatively homogeneous Hungarian, meaning that the proportion of multiple affiliations and outlier groups was much lower than in those settlements located along the Hungarian–Ukrainian ethnic boundary.

The main difference between the two surveys manifested in the number of inhabitants due to the high level of mobility of the population, mainly denoting labor migration. How these temporarily absent people are counted and interpreted may lead to significant differences in the numbers counted between surveys. In addition, the dates of the two surveys were at least one month apart, which further increases the chances of a migration event; as a consequence, the population numbers do not fully coincide.

Table 2

Number of Hungarians according to external classification (control) and self-identification (questionnaire) in the control survey sites in 2017

(persons)

Location	Survey type	Sampling site (cluster)			Locality total number of Hungarians
		population in the cluster	of which Hungarian	of which another ethnicity	
Perekhrestia/Tiszakeresztúr	control	55	55	0	533
	questionnaire	58	58	0	528
Tysauifalu/Tiszaújfalu	control	65	65	0	236
	questionnaire	65	65	0	214
Hudia/Gődényháza	control	43	38	5	144
	questionnaire	44	39	5	183
Pallo/Palló	control	68	65	3	217
	questionnaire	65	62	3	242
Mochola/Macsola	control	75	75	0	559
	questionnaire	66	66	0	516
Total	control	306	298	8	1,689
	questionnaire	298	290	8	1,683

While the population and ethnicity data for the sample sites differ only slightly in the two surveys, significant differences pertain between the data calculated and surveyed for the whole locality. Looking closely at the above data, two reasons may explain this phenomenon. First, due to sampling and/or data collection biases, the difference between the two data sets is inherently so significant for the sample site that the value calculated for the whole locality effectively multiplies this difference. Second, although sampling/data collection may be correct for the sampling site, the sampling site data do not represent the whole locality for some reason. In this case, the accuracy of the survey could be improved by increasing the number of households per sampling site.

Summary of results

By aggregating the birth, death, and migration rates by settlements, we estimated the number of Hungarians in the selected localities at 123,200 in 2017. This figure includes the temporarily absent population but excludes small, scattered Hungarian communities living outside the sampled settlements, which numbered 2,400 according to the 2001 census. For them, we applied the average proportion of such communities in the sample with similar ethnic structures. Thus, their number was estimated at 2,000, while the total number of Hungarians was roughly estimated at 125,300. Interestingly, the ethnic categorization of the mixed nationality population has little impact on this value. If individuals with dual ethnicity are excluded from the Hungarian population altogether, the number of Hungarians is barely lower (124,200). Similarly, if all dual-identity persons are included, the resulting value is only slightly higher (125,600).

However, even this figure is not directly comparable with the number in the last census, as our survey does not include the population of segregated Roma settlements claiming Hungarian identity. We do not have a methodologically substantiated method of determining their exact number. Based on the available literature, we only know that extrapolating from the number of individuals self-identifying as Hungarian in 2001, the size of this subpopulation was estimated at ~5,500 in 2016 (Molnár et al. 2016).[11] However, part of this group – those not living in segregated settlements – were included in our sample, and moreover, their ethnic identification appears to be particularly fluid. It also seems highly plausible that, amid the strengthening of Hungarian kin-state politics, and in the context of a survey conducted by Hungarian interviewers, a larger proportion of this population would have self-identified as Hungarian than during the 2001 census. Therefore, we assume that the census-comparable number of people considering themselves Hungarian in mid-2017 was most probably between 125,000 and 135,000, with an intermediate value of 130,000.[12]

[11] This figure is still far below the estimated 20,000 Hungarian-speaking Roma in Transcarpathia at that time (Molnár et al. 2016).

[12] The Hungarian population of 130,000 has declined by 21,500 compared to the 2001 census (a decrease of 14%). This is a slightly more favorable change than expected based on the immigration statistics in Hungary (Karácsonyi–Kincses 2020). In a broader context, this is also more favorable than the rate of decline of other significant minority Hungarian communities (in Romania, Serbia, and Slovakia) between 2001 and 2011.

Discussion and conclusion

In a context where the availability of reliable and up-to-date data is minimal, as in Ukraine, primary data collection by researchers can fill or at least reduce the gap. However, no methodology has been established for large-scale primary ethnic data collection. Thus, this study constitutes a search for a way forward.

We learned two critical lessons from the research methodology we developed regarding counting (sub-)populations (i.e., headcount) and measuring ethnicity. Among the theoretical considerations, (1) we found the enumeration of people to be the most challenging issue. The methodological classification of permanent and temporary migrants is crucial in this regard, and all the more so because both emigration (resettlement) and circular migration or dual residency in two countries (mostly cross-border labor and educational migration) have been essential livelihood strategies in Transcarpathia for decades, particularly since 2014 (Lipták–Kincses 2023). The impact of migration on accurate counting is, of course, underscored by other studies; for example, Skinner (2018: p. 50) identifies six main topics that challenge census-taking, two of which are related to migration ("individuals living in multiple locations," and "needs for more frequent data on changing patterns of internal and international migration"). In our study, we did accept the categorization of the respondents, even if, in reality, "temporary" migration can involve as long an absence as "permanent" migration. However, this methodological issue may impact the results profoundly (i.e., the number of Hungarians). If Hungarians spending more than six months a year abroad are classified as "permanent" migrants, the number of Hungarians would have to be reduced by 11,000−12,000. Consequently, we stress instead that considering the volatility of migration in our model, the number of Hungarians most likely ranged between 125,000 and 135,000 in 2017.

However, (2) some issues with measuring ethnicity, particularly the situational self-identification of individuals, can increase the uncertainty of the results regarding the number of Hungarians. It is primarily the Hungarian-speaking Roma who blur the otherwise stable ethnic boundary of Hungarians at the macro-level. However, our survey is not suitable for capturing this phenomenon more accurately, as we do not know how many individuals in our sample who self-identified (exclusively) as Hungarian may in actuality be of Roma origin (or self-identified as Roma previously), nor how many residents in the segregated Roma settlements would have self-identified as Hungarian. What we know is that, for example, the kin-state politics of Hungary around the turn of the millennium might have influenced the self-identification of the Roma in the 2001 census. The increased economic and political presence of the Hungarian state in Transcarpathia in the past decade (Tátrai et al. 2017) indicates that the declaration of Hungarian identity in (semi-)official situations may be more widespread among the Roma than before. With the growing prestige of the Hungarian language and Hungary (until 2022) and the benefits offered by Hungary (such as nonresident citizenship), an increasing number of Ukrainians also

claimed Hungarian (or multiple) identity in some contexts. However, fluid ethnic boundaries in the case of outlier categories such as the Roma and bilingual Ukrainians do not entail the blurring of the ethnic boundary between Hungarians and other categories at the macro-level, particularly in the Hungarian ethnic block settlement. This is confirmed by the fact that only 7% of the responding Hungarians claimed multiple identities and that in the public discourse and institutions in Transcarpathia, there is little room for dialogue about in-between categories (Hires-László 2017).

Other issues with measuring identity proved less complicated in the Transcarpathian context. It seems that the concept of ethnicity ("*nemzetiség*" in Hungarian) and mother tongue ("*anyanyelv*") have agreed-upon meanings shared by the majority of Hungarians. We experienced very few exceptions when ethnicity and citizenship were mixed up (Hires-László 2017); however, field research reports about sub-regions where the local understanding of ethnic terms is not equivalent to the standard meanings, and ethnic terms simultaneously function as religious terms and vice versa (Erőss et al. 2021). Nevertheless, we claim that most Hungarians in Transcarpathia accept and identify with the simplistic, one-word statistical ethnic categories. This is confirmed by a survey conducted in 2016, according to which the two crucial components of the identity construction of Hungarians in Transcarpathia are ethnic affiliation (Hungarianness: 70% of total mentions) and local/regional affiliation (Transcarpathianness: 60%) (Ferenc–Rákóczi 2018).

Beyond the theoretical considerations, a key question is how valid the survey results are. To this, we have three answers. First, although the control survey confirmed the data on ethnicity, it showed why the enumeration of the population may vary according to the different sources. The difference between the two datasets (questionnaire and control) regarding population numbers indicates that increasing the number of households surveyed at the local level ensured greater representativeness. Interestingly, despite the significant differences in the data at the local level, the aggregate data of the two datasets are similar. This indicates that the reliability of the aggregate data is better, although taking into account the small number of Hungarians in the control dataset (both per village and in total), not only the data of districts but also that of municipalities that include more than one sampling site are likely to be representative. However, it should be taken into account that four out of the five selected villages are located in the Hungarian ethnic block settlement, where ethnic identity and boundaries proved to be more stable than in the cities and the more scattered Hungarian communities, which were not included in the control survey for methodological reasons.

Second, survey results may be supported or refuted by external sources. We found only one study that is (partially) comparable to ours. Based on an analysis of immigration statistics in Hungary, Karácsonyi–Kincses (2020) found that the number of Transcarpathian Hungarians was ~120,000 in 2017. At first glance, the difference between the two values is considerable (11,000 persons). However, taking into account that migration statistics in Hungary are likely to include a significant

proportion of circular migrants whom we counted as residents in Transcarpathia, the two results may be much closer. Overall, we cannot claim that the two studies contradict or confirm each other due to the methodological differences.

Third, due to the nature of ethnicity, no control survey can claim validity in this regard. Self-identification may be situational and subject to change even within a short time, influenced by macro-level factors (such as kin-state politics) and micro-level factors (such as the person who collects the data). Thus, although data on the size of a (sub)population may be valid in the long term, those of an ethnic group are unlikely to be.

Another crucial issue is the extent to which the research method can be generalized, i.e., the potential for adapting it to other contexts. Although there are no context-dependent requirements that would make this method applicable only to Hungarians in Ukraine, the success of our model depends largely on a few structural factors. One such factor is that the group under study should be concentrated geographically (e.g., in borderlands). The greater the proportion of rural areas covered by the survey, the greater its effectiveness. Furthermore, it is advantageous if the group under study is characterized by low levels of assimilation, since our model cannot track changes in ethnic identification over the individuals' life course. Unless the survey provides full coverage, data from a relatively recent comprehensive population survey (preferably a census) are needed as a starting point for the survey. This may be a particularly relevant factor in countries where traditional censuses are being replaced by censuses that do not collect ethnic data. Accordingly, the model we described can be applied to other ethnic groups but only with certain limitations; thus, its methodological generalizability is constrained.

Overall, this study highlights two methodological and theoretical challenges associated with recent population and ethnic statistics. On the one hand, increasing mobility and multiple residencies make it difficult to keep track of the population and thereby identify them as residents of a single area; on the other hand, demographic and ethnic processes may diverge significantly due to the situational ethnic self-identification of parts of the population.

This has, unfortunately, been demonstrated by recent geopolitical events (i.e., the Russian invasion of Ukraine) that have fundamentally reshaped the ethno-demographic structure of the region, transforming our survey results into historical data.

Acknowledgement

The project "(Pre)retirement migration and its potential development impacts in rural Hungary" (K139127) was supported by the National Research, Development and Innovation Office. Ágnes Erőss was supported by the FemMinImm MSCA Seal of Excellence project financed by the Autonomous Province of Bolzano/Bozen. Katalin Kovály was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Appendix

Table A1

The key logistical information for the survey

	Cluster sampling method	Name register method	"Areal snowball" method	Total
Number of localities surveyed	103	2	8	113
Number of interviewers	152	9	9	170
Number of sampling sites with data of sufficient quality for calculations	269	16	11	296
Average number of surveyed households per sampling sites	19.9	21.0	17.1	19.9
The time frame of data collection	01.06.2017–15.09.2017	01.06.2017–31.08.2017	15.06.2017–30.09.2017	01.06.2017–30.09.2017

REFERENCES

AHMED, P.–FELICIANO, C.–EMIGH, R. J. (2007): Internal and external ethnic assessments in Eastern Europe *Social Forces* 86 (1): 231–255. https://doi.org/10.1353/sof.2007.0087

ANDERSON, B. (1991): *Imagined communities. Reflection on the origin and spread of nationalism* Verso, London.

AREL, D. (2002): Interpreting "nationality" and "language" in the 2001 Ukrainian census *Post-Soviet Affairs* 18 (3): 213–249. https://doi.org/10.2747/1060-586X.18.3.213

ASPINALL, P. J.–SONG, M. (2013): *Mixed race identities* Palgrave Macmillan, Basingstoke.

BARTH, F. (ed.) (1998): *Ethnic groups and boundaries: the social organization of culture difference* Waveland Press, Long Grove.

BLUM, A.–GUÉRIN-PACE, F. (2008): From measuring integration to fighting discrimination – the illusion of ethnic statistics *French Politics, Culture and Society* 26 (1): 45–61. https://doi.org/10.3167/fpcs.2008.260104

BROWN, G. K.–LANGER, A. (2010): Conceptualizing and measuring ethnicity *Oxford Development Studies* 38 (4): 411–436. https://doi.org/10.1080/13600818.2010.525629

BRUBAKER, R. (1996): *Nationalism reframed: nationhood and the national question in the New Europe* Cambridge University Press, Cambridge.

BRUBAKER, R. (2009): Ethnicity, race, and nationalism *Annual Review of Sociology* 35 (1): 21–42. https://doi.org/10.1146/annurev-soc-070308-115916

BURTON, J.–NANDI, A.–PLATT, L. (2010): Measuring ethnicity: challenges and opportunities for survey research *Ethnic and Racial Studies* 33 (8): 1332–1349. https://doi.org/10.1080/01419870903527801

ÇAĞLAR, A. (2013): Circular migration between Hungary and Ukraine: Historical legacies, the economic crisis, and the multidirectionality of 'circular' migration. In: TRIANDAFYLLIDOU, A. (ed.): *Circular migration between Europe and its neighbourhood: choice or necessity?* pp. 141–165., Oxford University Press, Oxford.

CHUGAIEVSKA, S.–DEHNEL, G.–TARGONSKII, A. (2023): Census administration in Ukraine: insight into the Polish experience in the context of international indicators analysis *Statistics in Transition* 24 (3): 213–239. https://doi.org/10.59170/stattrans-2023-045

CSATA, Z.–HLATKY, R.–LIU, A. H. (2020): How to head count ethnic minorities: validity of census surveys versus other identification strategies *East European Politics* 37 (3): 572–592. https://doi.org/10.1080/21599165.2020.1843439

CSERNICSKÓ, I.–KONTRA, M. (2023): The linguistic human rights plight of Hungarians in Ukraine. In: SKUTNABB-KANGAS, T.–PHILLIPSON, R. (eds.): *The handbook of linguistic human rights* pp. 373–382., Wiley, Oxford and Malden.

ELRICK, J.–SCHWARTZMAN, L. F. (2015): From statistical category to social category: organized politics and official categorizations of 'persons with a migration background' in Germany *Ethnic and Racial Studies* 38 (9): 1539–1556. https://doi.org/10.1080/01419870.2014.996240

ERŐSS, Á.–KOVÁLY, K.–TÁTRAI, P. (2021): Ethnic categorization practices and boundary (re)making in a multiethnic borderland of Ukraine *Border and Regional Studies* 9 (4): 173–198. https://doi.org/10.25167/brs4561

FERENC, V.–RÁKÓCZI, K. (2018): Ukrainian–Hungarian coexistence in Transcarpathia based on the 'TANDEM 2016' survey *Hungarian Journal of Minority Studies* 2: 137–161.

HIRES-LÁSZLÓ, K. (2017): Etnikai kategóriák a beregszásziak mindennapi diskurzusaiban. In: MÁRKU, A.–TÓTH, E. (eds.): *Többnyelvűség, regionalitás, nyelvoktatás* pp. 121−136., RIK-U, Uzhhorod.

HOÓZ, I. (1985): A népesség nemzetiség szerinti megoszlásának tanulmányozásáról. In: HOÓZ, I.–KEPECS, J.–KLINGER, A. (eds.): *A Baranyában élő nemzetiségek demográfiai helyzete 1980-ban* pp. 19–53., MTA RKK, Baranya Megyei Tanács, Állami Gorkij Könyvtár, Pécs and Budapest.

HORVÁTH, I.–KISS, T. (2018): *SocioRoMap. Compact Roma Communities in Romania: survey of local experts on social inclusion* ISPMN, Cluj-Napoca.

JENKINS, R. (1994): Rethinking ethnicity: identity, categorization and power *Ethnic and Racial Studies* 17 (2): 197–223. https://doi.org/10.1080/01419870.1994.9993821

JOSIPOVIČ, D. (2015): Slovenia and the census: from the 20. century Yugoslav counts to the register-based census of 2011 *Contemporary Southeastern Europe* 2 (2): 159–175.

KARÁCSONYI, D.–KINCSES, Á. (2020): Átrendeződés? Kárpátaljaiak Magyarországon, magyarok Kárpátalján, a 2017-ig terjedő adatok tükrében *Területi Statisztika* 60(3): 309–351. https://doi.org/10.15196/TS600302

KEMÉNY, I.–JANKY, B. (2005): Roma population of Hungary 1971–2003. In: KEMÉNY, I. (ed.): *Roma of Hungary* pp. 70–225., Atlantic Research and Publication, Highland Lakes.

KEMÉNYFI, R. (1998): Die Frage nach der ethnischen Grenze und der Multiethnizität am Beispiel des ehemaligen ungarisch–slowakischen Komitates Gömör und Kis-Hont *Ethnographica et Folkloristica Carpatica* 9–10: 50–70.

KERTZER, D. I.–AREL, D. (2002): Censuses, identity formation, and the struggle for political power. In: KERTZER, D. I.–AREL, D. (eds.): *Census and identity. The politics of race, ethnicity, and language in national censuses* pp. 1–42., Cambridge University Press, Cambridge.

KISS, T. (2018): Demographic dynamics and ethnic classification: an introduction to societal macro-processes. In: KISS, T.–SZÉKELY, I. G.–TORÓ, T.–BÁRDI, N.–HORVÁTH, I. (eds.): *Unequal accommodation of minority rights: Hungarians in Transylvania* pp. 383–417., Palgrave Macmillan, Cham.

KOCSIS, K. (2001): *Ethnic map of present day Transcarpathia 1941, 1999*. MTA FKI and MTA KKI., Budapest.

LADÁNYI, J.–SZELÉNYI, I. (2006): *Patterns of exclusion: constructing gypsy ethnicity and the making of an underclass in transitional societies of Europe* Columbia University Press, New York.
LIPTÁK, K.–KINCSES, Á. (2023): International migration of Ukrainian citizens to Central Europe before the Russo–Ukrainian wars *Regional Statistics* 13 (5): 781–796. https://doi.org/10.15196/RS130501
MATEOS, P. (2007): A review of name-based ethnicity classification methods and their potential in population studies *Population, Space and Place* 13 (4): 243–263. https://doi.org/10.1002/psp.457
MESSING, V. (2014): Methodological puzzles of surveying Roma/gypsy populations *Ethnicities* 14 (6): 811–829. https://doi.org/10.1177/1468796814542180
MOLNÁR, J. (2016): Національна ідентичність нащадків етнічно змішаних (українсько-угорських) закарпатських родин за даними перепису населення та анкетного опитування. In: BERGHAUER, S.–DNYISZTRJANSZKIJ, M.–FODOR, GY.–GÖNCZY, S.–IZSÁK, T.–JAKAB, N.–MOLNÁR, J.–MOLNÁR D., I.–PAPP, G.–SASS, E.–VINCE, T. (eds.): *Social geographical challenges and search for adequate answers in East-Central Europe of the 21st century* pp. 166–174., Ferenc Rákóczi II Transcarpathian Hungarian Institute, Berehove.
MOLNÁR, J.–CSERNICSKÓ, I.–BRAUN, L. (2016): Cigányok Kárpátalján. In: SZILÁGYI, F.–PÉNZES, J. (eds.): *Roma népesség Magyarország északkeleti határtérségében* pp. 91–108., Partium Kiadó, Nagyvárad.
MUŠINKA, A.–ŠKOBLA, D.–HURRLE, J.–MATLOVIČOVÁ, K.–KLING, J. (2014): *Atlas rómskych komunít na Slovensku 2013* UNDP, Bratislava.
PEACH, C. (2000): Discovering white ethnicity and parachuting plurality *Progress in Human Geography* 24 (4): 620–626. https://doi.org/10.1191/030913200100189148
PÉNZES, J.–PÁSZTOR, I. Z.–TÁTRAI, P.–KÓTI, T. (2019): Roma population in Hungary – spatial distribution and its temporal changes *Deturope* 11 (3): 138–159. https://doi.org/10.32725/det.2019.030
RALLU, J.-L.–PICHÉ, V.–SIMON, P. (2006): Demography and ethnicity: an ambiguous relationship. In: CASELLI, G.–VALLIN, J.–WUNSCH, G. (eds.): *Demography: analysis and synthesis. A treatise in population studies 3* pp. 531–549., Academic Press, London.
RUGHINIŞ, C. (2011): Quantitative tales of ethnic differentiation: measuring and using Roma/gypsy ethnicity in statistical analyses *Ethnic and Racial Studies* 34 (4): 594–619. https://doi.org/10.1080/01419870.2010.514055
SCOTT, A.–PEARCE, D.–GOLDBLATT, P. (2001): The sizes and characteristics of the minority ethnic populations of Great Britain – latest estimates *Population Trends* (105): 6–15.
SIEGEL, J. S. (2018): *Demographic and socioeconomic basis of ethnolinguistics* Springer, Cham.
SIMON, P. (2012): Collecting ethnic statistics in Europe: a review *Ethnic and Racial Studies* 35 (8): 1366–1391. https://doi.org/10.1080/01419870.2011.607507
SIMON, P.–PICHÉ, V. (2012): Accounting for ethnic and racial diversity: the challenge of enumeration *Ethnic and Racial Studies* 35 (8): 1357–1365. https://doi.org/10.1080/01419870.2011.634508
SKINNER, C. (2018): Issues and challenges in census taking *Annual Review of Statistics and its Application* 5: 49–63. https://doi.org/10.1146/annurev-statistics-041715-033713
STRYAPKO, I. (2016): The level of interethnic tolerance of the Ukrainians in the Transcarpathian region *Annales Scientia Politica* 5 (1): 42–47.
SUPIK, L.–SPIELHAUS, R. (2019): Introduction to special issue: matters of classification and representation: quantifying ethnicity, religion and migration introduction *Ethnicities* 19 (3): 455–468. https://doi.org/10.1177/1468796819833434

SURDU, M. (2016): *Those who count. Expert practices of Roma classification* Central European University Press, Budapest.
TÁTRAI, P.–ERŐSS, Á.–KOVÁLY, K. (2017): Kin-state politics stirred by a geopolitical conflict: Hungary's growing activity in post-Euromaidan Transcarpathia, Ukraine *Hungarian Geographical Bulletin* 66 (3): 203–218. https://doi.org/10.15201/hungeobull.66.3.2
TÁTRAI, P.–MOLNÁR., J.–KOVÁLY, K.–ERŐSS, Á. (2018): Changes in the number of Hungarians in Transcarpathia based on the survey 'SUMMA 2017' *Hungarian Journal of Minority Studies* 2: 103–135.
TELLES, E.–LIM, N. (1998): Does it matter who answers the race question? Racial classification and income inequality in Brazil *Demography* 35 (4): 465–474. https://doi.org/10.2307/3004014
TÓTH, B.–PAPP, S. (2020): A közösségi média társadalomföldrajzi alkalmazásának etnikai-földrajzi példája *Területi Statisztika* 60 (6): 714–741. https://doi.org/10.15196/TS600605
VERES, V. (2015): The minority identity and the idea of the 'unity' of the nation: the case of Hungarian minorities from Romania, Slovakia, Serbia and Ukraine *Identities: Global Studies in Culture and Power* 22 (1): 88–108. https://doi.org/10.1080/1070289X.2014.922885
VINNYCHUK, I.–ZIUKOV, S. (2013): Shadow economy in Ukraine: modelling and analysis *Business Systems and Economics* 3 (2): 141–152. https://doi.org/10.13165/VSE-13-3-2-01
WATERMAN, S.–KOSMIN, B. (1986): Mapping an unenumerated ethnic population: Jews in London *Ethnic and Racial Studies* 9 (4): 484–501. https://doi.org/10.1080/01419870.1986.9993547
WILLIAMS, M.–HUSK, K. (2013): Can we, should we, measure ethnicity? *International Journal of Social Research Methodology* 16 (4): 285–300. https://doi.org/10.1080/13645579.2012.682794
WIMMER, A. (2013): *Ethnic boundary making: institutions, power, networks* Oxford University Press, New York.
YAO, E. S.–MEISSEL, K.–BULLEN, P.–CLARK, T. C.–ATATOA CARR, P.–TIATIA-SEATH, J.–PEIRIS-JOHN, R.–MORTON, S. M. B. (2022): Demographic discrepancies between administrative-prioritisation and self-prioritisation of multiple ethnic identifications *Social Science Research* 103: 102648. https://doi.org/10.1016/j.ssresearch.2021.102648

INTERNET SOURCES

MALYNOVSKA, O. (2021): *Migration in Ukraine: facts and figures* IOM Ukraine, Kyiv. https://ukraine.iom.int/sites/g/files/tmzbdl1861/files/documents/migration_in_ukraine_facts_and_figures_2021-eng_web-1.pdf (downloaded: April 2025)
MIRGA-WÓJTOWICZ, E.–FIAŁKOWSKA, K. (2022): *"Be careful out there, in that Gypsy district" – anti-gypsyism in a war situation* Heinrich Böll Stiftung, Warsaw. https://pl.boell.org/en/2022/05/16/uwazajcie-tam-w-tej-cyganskiej-dzielnicy-antycyganizm-w-warunkach-wojny (downloaded: April 2025)