



2017 Regional Roma Survey: Quantitative Data Collection of Socio-Economic Position of Marginalised Roma in Western Balkans

Technical report

2018

Prepared by :



Supported by:



Contents

Introduction	4
1. Developing the Survey	5
1.1. Survey goals and timeline.....	5
1.2. Project partners and their role.....	8
1.3. Project stakeholders.....	8
2. Questionnaire Development	10
2.1. Overview of Modules.....	10
2.2. Questionnaire Structure.....	11
2.3. Comparability needs (2004 and 2011, core indicators).....	11
2.4. Reasoning behind revisions.....	14
2.5. Translation.....	15
2.6. Constructing CAPI version.....	15
2.7. Interviewer training material.....	16
3. Interviewer selection and training	17
3.1. Project and survey briefing.....	17
3.2. Interviewer selection and training.....	18
4. Sampling	21
4.1. Target populations and sample requirements.....	21
4.2. Sample sizes: Roma and non-Roma.....	21
4.3. Sample frames and mapping of information available for sampling.....	23
Albania.....	23
Bosnia and Herzegovina.....	24
Kosovo.....	25
the former Yugoslav Republic of Macedonia.....	26
Montenegro.....	26
Serbia.....	27
4.4. Overview of sampling methodology.....	27
5. Piloting	31
5.1. Methodology.....	31
5.2. Key findings, recommendations and actions as a result of the pilot.....	32

6. Fieldwork	36
6.1. Fieldwork dates and progress (per country/territory)	36
6.2. Contacting respondents.....	37
6.3. Recruitment of respondents and interview format	37
6.4. Interview length.....	38
6.5. Fieldwork outcomes	39
6.6. Fieldwork control.....	40
6.7. Data processing and cleaning	40
7. Lessons learned	43
7.1. Questionnaire revisions	43
7.2. Fieldwork and data processing	43



Introduction

The **2017 Regional Roma Survey** on Socio-Economic position of Marginalised Roma¹ initially intended to include six countries/ territory of Western Balkans – Albania, Bosnia and Herzegovina, Kosovo*², the former Yugoslav Republic of Macedonia, Montenegro, Serbia, – and Turkey. During the course of preparation of survey design counterparts in Turkey opted out of participation.

The overall objective of the survey is to contribute to narrowing the gap in multi-dimensional poverty and access to socio-economic rights between marginalised Roma and non-Roma populations in Western Balkans, by contributing to fact-based decision making and to more informed and focused policy debate on Roma inclusion.

The need to conduct a survey about socio-economic position of marginalised Roma was a result of a consensus among EU (DG NEAR), World Bank (WB) and UNDP that well-targeted and impactful Roma inclusion policies and programmes are only possible if there is up-to-date and reliable data to inform and monitor them. The last regional Roma survey was carried out by a coalition of EU, WB and UNDP in 2011, and the resulting data have been widely used and consulted by national and international bodies. However, it was established that there is a need to update the 2011 results, in order to monitor progress in different countries, evaluate the impact of different policies and adjust them accordingly, guide budget allocations, and to monitor trends in different dimensions of Roma exclusion across time.

The need for ‘fresh’ data became more apparent as the ‘new generation’ of strategies and action plans were further aligned with accession requirements of Western Balkan candidate and potential candidate countries, and reflect the need for further convergence with the European Union policy on Roma inclusion, embedded in the EU Framework for National Roma Integration Strategies up to 2020³. The framework highlights the need for: (i) an integrated approach to socio-economic inclusion of Roma; (ii) incorporation of national Roma integration strategies within the broader framework of EU 2020; (iii) the use of different EU funding mechanisms to implement strategies; (iv) a focus on disadvantaged micro-regions and segregated areas; and (v) the need for monitoring and evaluation of different interventions and overall national strategies. In line with this overarching framework, National Roma Integration/Inclusion strategies, including in enlargement countries should serve as pillars for targeted and cross-sectorial interventions.

In light of this background, an agreement between DG NEAR, UNDP and the World Bank has been reached to initiate third survey about socio-economic position of marginalised Roma in Western Balkans and in Turkey in order to provide data for purposes of comparison over time and across countries.

The initiation of the survey on Socio- economic position of marginalised Roma in Western Balkans was formally announced in March 2017 with the fieldwork to be conducted from September to November 2017. This technical report presents the process of preparation of the fielding of survey: questionnaire

¹ In line with the terminology of European institutions and international organisations, the term ‘Roma’ is used here to refer to a number of different groups (e.g. Roma, Sinti, Kale, Gypsies, Romanichels, Boyash, Ashkali, Egyptians, Yenish, Dom, Lom, Rom, Abdal) and includes travellers, without denying the specificities of these groups.

² * For the United Nations: All references to Kosovo shall be understood to be in the context of UN Security Council Resolution 1244 (1999)

*For the European Union: This designation is without prejudice to position on status, and is in line with UNSCR 1244/1999 and the ICJ Opinion on the Kosovo declaration of independence.

³ Council of the European Union (2011)

design, sampling, piloting and fieldwork, as well as different partnerships crucial for the process and main lessons learned.

1. Developing the Survey

1.1. Survey goals and timeline

The Regional Survey on Socio-Economic position of Marginalised Roma initially intended to include six countries/ territory of Western Balkans – Albania, Bosnia and Herzegovina, Kosovo, the former Yugoslav Republic of Macedonia, Montenegro, Serbia, – and Turkey. During the course of preparation of survey design counterparts in Turkey opted out of participation.

The overall objective of the survey is to contribute to narrowing the gap in multi-dimensional poverty and access to socio-economic rights between marginalised Roma and non-Roma populations in Western Balkans, by contributing to fact-based decision making and to more informed and focused policy debate on Roma inclusion.

The need to conduct a survey about socio-economic position of marginalised Roma was a result of a consensus among EU (DGNEAR), World Bank (WB) and UNDP that well-targeted and impactful Roma inclusion policies and programmes are only possible if there is up-to-date and reliable data to inform and monitor them. The last regional Roma survey was carried out by a coalition of EU, WB and UNDP in 2011, and the resulting data have been widely used and consulted by national and international bodies. However, it was established that there is a need to update the 2011 results, in order to monitor progress in different countries, evaluate the impact of different policies and adjust them accordingly, guide budget allocations, and to monitor trends in different dimensions of Roma exclusion across time.

Socio-economic position of estimated one million Roma living in Western Balkans has been the subject of intense debate in the Western Balkans, particularly since 2005, when the Governments of 10 countries from Central and South-Eastern Europe agreed to join forces to address the most pressing problems facing the Roma population through the creation of a multi-national platform - the Decade of Roma Inclusion 2005-2015. (Kosovo and Turkey were not part of this multi-country initiative.) Throughout the Decade there emerged a growing consensus that solid analysis, based on reliable data, is a prerequisite for documenting and addressing the multidimensional exclusion experienced by Roma.

The Decade, and subsequent supranational policy frameworks put significant emphasis on the need to improve the position of Roma in the areas of education, employment, health and housing, as well as to address cross-cutting issues of poverty reduction, gender equality and anti-discrimination. This led to recognition of a need to determine the size of population whose needs should be met. Government efforts to collect quantitative data on ethnicity are usually confined to the national census. However, it is widely accepted that where Census data allows for minority/ethnicity self-identification, the reported minority/ethnic affiliation for all minorities, and in particular the Roma minority, is significantly underestimated. Alternative estimates abound, but the most cited are those of the Council of Europe⁴.

⁴ <http://www.coe.int/t/dg3/romatravellers/Source/documents/stats.xls>

These estimates suggest that the number of Roma can be 2.7 to 4 times higher than that recorded in national census data.

In other regular and specific quantitative surveys governments have been (and remain) constrained by the standard interpretation of legal regulations restricting the collection of official data based on ethnicity, and by lack of resources to implement comprehensive surveys of this type⁵. In addition, as Roma may opt to avoid self-identification for numerous reasons. This is evident when comparing officially recorded numbers of those who self-identified as Roma in national censuses (that allow for self-identification based on ethnicity) with estimates provided through a combination of administrative and data collected by (pro) Roma non-governmental organizations.

In order to address these gaps UNDP and its international partners have been working, since 2004 with independent survey firms, national statistical offices and community-based organizations to show that it is possible and necessary to collect socio-economic data from Roma households. This approach has shown that: i) there are ethically and legally correct methodologies for collecting data based on ethnicity, making it possible to collect data on the socio-economic position of Roma; and that ii) problems of Roma exclusion can be quantitatively measured and monitored – and therefore better addressed – by governments and the international community.

Importantly, UNDP acknowledges that Roma are a highly heterogeneous group, and are perceived as a unified group only by outsiders. There is no straightforward answer to the very simple question ‘who is Roma?’, because ‘Roma identity’ combines elements of ethnic, cultural, linguistic and socio-economic characteristics, in various combinations depending on the specific context. Roma are also reluctant to self-identify as Roma to ‘outsiders or researchers’, in order to avoid ghetto stigma and because of fear of discrimination or being treated as not (full) citizens if they disclose their identity.

Thus, UNDP in previous survey approached to the identification of the ‘Roma universe’ that combines subjective self-identification with territorial markers and external identification. In 2017 survey the approach was modified to include self-identification only. Importantly, the underlying message of all survey since 2004 is that they are not ‘counting the Roma’ exercise, but one which outlines the major characteristics that are specific to the group and that correlate with – as well determine – various aspects of exclusion. The purpose of data collection is to identify the specific drivers of Roma exclusion, so that relevant and targeted approaches can be developed. In other words, UNDP has invested time and resources in building reliable data profiles of Roma communities that can be – and are – the basis for evidence-based policy making for promoting Roma inclusion.

The need for ‘fresh’ data became more apparent as the ‘new generation’ of strategies and action plans were further aligned with accession requirements of Western Balkan candidate and potential candidate countries, and reflect the need for further convergence with the European Union policy on Roma inclusion, embedded in the EU Framework for National Roma Integration Strategies up to 2020⁶. The framework highlights the need for: (i) an integrated approach to socio-economic inclusion of Roma; (ii) incorporation of national Roma integration strategies within the broader framework of EU 2020; (iii) the use of different EU funding mechanisms to implement strategies; (iv) a focus on disadvantaged micro-regions and segregated areas; and (v) the need for monitoring and evaluation of different interventions

⁵ For example, when national statistical offices are engaged in ethnic data collection, as in case of UNICEF’s MICS, external funding is available to them.

⁶ Council of the European Union (2011)

and overall national strategies. In line with this overarching framework, National Roma Integration/Inclusion strategies, including in enlargement countries should serve as pillars for targeted and cross-sectorial interventions.

Importantly, the 2014 Enlargement Strategy and Main Challenges 2013-2014⁷ identified Roma inclusion as one of the ‘fundamentals’ in the negotiation process. The Strategy asserts that ‘the situation of most Roma communities is a matter of serious concern’⁸. This requires integration of policies, funding mechanisms and focused attention on interventions enabling access to rights, capacities to deliver rights and support for implementation processes. Data collection and adequate indicators are needed to help (re)define and monitor processes and their outcomes at national levels, and can support the European Commission’s efforts to better target funds and strengthen strategic cooperation with different stakeholders, including international development partners, building on their comparative advantages. In addition, the thematic evaluation undertaken by the EU to evaluate support for Roma inclusion provided under the Instrument for Pre-accession Assistance /IPA I⁹ drew attention once more to the importance of reliable data for monitoring existing policies/ programmes and informing the design of new ones.

In light of these background, an agreement between DG NEAR, UNDP and the World Bank has been reached to initiate third survey about socio-economic position of marginalised Roma in Western Balkans and in Turkey in order to provide data for purposes of comparison over time and across countries.

The survey was formally announced in March 2017 with the fieldwork to be conducted from September to November 2017. The tender for services of survey companies was announced on 21st April 2017 with an original deadline date of Monday, 22nd which then had been extended until 12th June 2017 in order to ensure that sufficient numbers of bidders apply. A total of 6 bidders submitted their proposals which were then evaluated based on technical and financial criteria. Ipsos Statagic Marketing was selected as the best bidder and contracted in August 2017.

Given that in the initial phase of the survey design, as early as summer 2016, it was established that there is a need to revise the questionnaire whilst maintaining the comparability with core indicators (Annex 1) set in 2011 survey, discussions were initiated with a number of partners to examine possibilities to introduce changes based on needs of different users of survey data. World Bank, as the main partner was also the main interlocutor in introducing changes which mostly focused on incomes, poverty and early child development related set of questions. Other interlocutors included Fundamental Rights Agency, UNICEF and UNHCR. Questionnaire revisions proved to be cumbersome and time consuming as they interfered with the philosophy of multi-indicator household surveys with ‘hard to reach’ population. Ultimately, this had an impact on the start of the fieldwork which was postponed for November 2017. More details about questionnaire revisions are provided in the dedicated section in this report.

⁷ European Commission (2013)

⁸ European Commission (2013)

⁹ European Commission DG NEAR (2016).

1.2. Project partners and their role

The Survey has been designed in close cooperation with the World Bank (WB), as both UNDP and WB are implementing Component 2 of the IPA-funded action “Technical Assistance for Roma Integration at the Community Level” (TARO). Division of tasks meant that UNDP implemented the survey, while the World Bank works on in-depth analysis of all the data and outputs produced by the survey. For this reason, technical staff from the UNDP and World Bank have formed a Technical Working Group (TWG) to ensure coordination on all aspects of survey design and implementation. The TWG also included representatives of the EU’s Fundamental Rights Agency (FRA): who provided technical advice at different stages of questionnaire development and discussions about sampling approaches, but were not involved in any field or survey work in the non-EU member states; UN Agencies, namely UNICEF, WHO, ILO, UN Women; and potential users of the survey data, such as National Statistical Offices (NSOs) and other relevant national government bodies.

As the main project partner tasked by DG NEAR to analyse data, World Bank was involved in all stages of the project: from consultations in survey design stage, discussions on approach to sampling, questionnaire revisions, fieldwork monitoring and data checking. Expertise in questionnaire design and data processing was the value added of this collaboration to UNDP and the survey company conducting fieldwork.

Fundamental Rights Agency’s expertise with surveying disadvantaged and hard to reach population in EU Member States was of particular value, in particular in light of their 2015 EU MIDIS II survey experience. This experience included revisions of questionnaire as well as experimenting with different sampling approaches such as adaptive cluster sampling¹⁰.

1.3. Project stakeholders

The project stakeholders ranged from (pro) Roma civil society organization, relevant national government bodies and national statistical offices to international organisations such as Regional Cooperation Council.

Survey design and structure of questionnaires were presented to central and local level stakeholders in Western Balkans during inception workshops held in the period from 7 March - 20 May. In Turkey, the survey was presented in May 2017 during the coordination meeting of the actors implementing National Roma Inclusion Strategy, and at later stages to relevant ministries

Workshop participants included representatives of EUDs in each country, national Roma focal points, line ministries in charge of implementation of sectoral level policies (education, employment, health, housing, social policies); national statistical offices representatives of (pro) Roma civil society

¹⁰ For discussion on sampling approaches, please consult Discussion paper Sampling for Roma Regional Survey 2017: Possible approaches (UNDP)

organisations; think tanks and independent research organisations; and international organisations. The number of participants varied from 27 to 36 per workshop.

Inception workshops resulted in useful inputs for TWG :

- 2011 questionnaire was too long. There is a need to strike a good balance between general/broad questions and focused ones.
- Interest in more detailed questions on views of Roma about government interventions supporting employment; identification of non-formal skills which could help Roma enter labour markets; the trends in terms of employment in public institutions.
- The position of Roma being returned from European Union and which services they have access to upon return.
- Exploring possibility to interview same households as in 2011 which could result in creation of longitudinal household panel data.
- Stronger message should be sent to National Statistical Offices (NSOs) to include Roma sub-samples into cyclical surveys.
- Launching country specific data analysis based on survey results was a good practice in 2011. This approach should be maintained in this round survey.

Most of these inputs were incorporated, save for the comment on the length of questionnaire (as presented in the dedicated section later in the report). **National statistical offices** were additionally consulted at the **workshop Methods and Tools for Welfare Measurement and Diagnostics and the Roma Regional Survey** organised by UNDP and the World Bank in Vienna, 6-7 July 2017. Based on these discussions a final sampling approach was decided and used to construct the sample in cooperation with the survey company conducting the fieldwork. In addition, the workshop was used to raise awareness among dedicated technical staff of the national statistical offices about the need to regularly collect ethnically disaggregated data.

Cooperation with Roma Integration 2020 technical team of the Regional Cooperation Council resulted in the presentation of the survey design at a workshop¹¹ with National Roma Focal Points in charge of monitoring and reporting about the progress of National Roma Inclusion policies. This was the opportunity to present links between survey data and specific requirements of the monitoring framework.

Technical Assistance for Promoting Social Inclusion in Densely Roma Populated Areas (SIROMA) project team and National Focal Point for Roma National Action Plan in Turkey were regularly consulted with regards to possibilities to implement the survey in Turkey. Survey design and questionnaire structure were presented to national stakeholders during a workshop dedicated to progress monitoring¹².

¹¹ Vienna, December 2016.

¹² Ankara, May 2017.

2. Questionnaire Development

2.1. Overview of Modules

The questionnaire was developed, based on 2011 regional Roma survey questionnaire, in close cooperation between United Nations Development Programme and the World Bank. Fundamental Rights Agency, UNICEF, UNHCR, WHO, central level authorities from Western Balkans and (pro)Roma civil society organisations were consulted at various stages of revisions. The final questionnaire consisted of three core Modules and a Management Module. The number of core Modules was reduced from four to three in comparison with 2011 survey since Early Childhood Development Module was revised to the extent that it no longer functioned as separate module. Instead six questions about parenting techniques were included into Module Household Members Profiles.

Questionnaire followed the philosophy of integrated household surveys.

Management Module was filled in by the interviewer before and after the interview.

Module Household Members Profile was completed for all members of the household from the account of the household head or other knowledgeable person in household, while participation of other members was acceptable. It contains information about: demographic profile; education; health profile; current employment status, wages and salaried income; household members currently out of the country and returnees; and parenting techniques.

Module Status of the Household was recorded from the account of the household head. The Module collected general information about household such as status of the neighbourhood over time, use of language at home and others; livelihoods and self-employment/entrepreneurship; self-employment income; pensions and social benefits in the last 12 months; social assistance; transfers between households, financial transfers and agricultural activity; and levels of living standards and income security.

Module Individual Status and Attitudes was administered to randomly selected respondents and included sets of questions about migration and mobility; access to health services; values and norms; time use, rights awareness and general discrimination; active citizenship and trust; financial decision making in household.



2.2. Questionnaire Structure

The questionnaire covered following topics:

MODULE	TOPICS COVERED
Management section	<ul style="list-style-type: none"> ● ethnic background: self-reported ● external assessment of dwelling and neighbourhood
Household Profile	<ul style="list-style-type: none"> ● demographic structure of household ● education ● employment ● health ● wages and salaried income ● personal documentation ● household members currently abroad and returnees
Status of the Household	<ul style="list-style-type: none"> ● status of the household ● regularity of waste collection ● use of language at home ● participation in civic organisations ● livelihoods and self-employment/entrepreneurship ● self-employment income ● pensions and social benefits ● social assistance ● transfers between households, financial transfers and agricultural activity ● living standards and economic security
Individual Status and Attitudes	<ul style="list-style-type: none"> ● migration ● self-reported access to health services ● values and norms ● general discrimination ● active citizenship and trust ● time use ● financial decision making in the household

2.3. Comparability needs (2004 and 2011, core indicators)

To monitor progress in Roma inclusion it was necessary to maintain comparability with previous and on-going surveys, in particular:

- **The UNDP/WB/EC regional Roma survey 2011** which interviewed 750 Roma and 350 non-Roma households living in or close to Roma communities in 12 countries of Central and South-Eastern Europe— Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Hungary, Moldova, Montenegro, Romania, Serbia, Slovakia, the former Yugoslav Republic of Macedonia.
- **The EU FRA 2011 Survey** carried out by the European Union Agency for Fundamental Rights and UNDP (2011) interviewed a total of 22,203 Roma and non-Roma providing information on 84,287 household members in Bulgaria, Czech Republic, France, Greece, Hungary, Italy,

Poland, Portugal, Romania, Slovakia and Spain. Face-to-face interviews were carried out in Roma and non-Roma households, living within close proximity of each other.

- **The EU-MIDIS II: European Union minorities and discrimination survey 2015.** The second wave of the European Union Minorities and Discrimination Survey (EU-MIDIS) aimed to assess progress made over the past five years, since the first EU-MIDIS survey in 2008. It collects comparable data in all 28 EU Member States (aims to cover around 25,000 people) to support the EU in protecting the fundamental rights of people with an immigrant or ethnic minority background, including Roma. It also aims to supply data for the calculation of core indicators for measuring progress in the implementation of the EU Framework for National Roma Integration Strategies, as well as selected indicators on migrant integration.

The questionnaire has been revised in such a manner that it retained the following core indicators ensuring compatibility with 2011 regional Roma survey and EU FRA 2012 survey, as well as with EU MIDIS II survey (see Table 1).

Table 1 : List of indicators

Indicator name
Education
Literacy rates
Highest completed education
Pre-school enrolment rate
Gross enrolment rate in compulsory education (7-15)
Gross enrolment rate in upper secondary education (16-19)
Number of years in education
Educational expectations boys*
Educational expectations girls*
Employment
Employment rate
Unemployment rate
Activity rate (Labour force participation rate)
Last employment experience
No employment experience rate
Self-employment rate
Informal employment incidence
Minimum monthly wage for which a person is willing to work full time*
Preferences - employment regularity*
Preferences - employment security*
Health
Health self-assessment
Access to medical insurance
Incidence of specific medical checks*
Shares of the population not having access to essential drugs ¹³

¹³ RRS2010 and 2017 use slightly different question formulation:
 2010 Q2.3 Were there any instances in the past 12 months when your household could not afford purchasing medicines prescribed to, needed for a member of your household?
 2017 Q-4.8 Can you afford [...]? 5. Purchase medicines prescribed to or needed for a member of your household



Indicator name
Access to health services*
Housing
Neighbourhood change
Regularity of waste collection
Rooms per HH member
Square meters per household member
Shares of the population not having access to secure housing
Shares of the population not having access to improved water source
Share of the population not having access to drinking water inside their dwelling
Shares of the population not having access to improved sanitation
Access to electricity
Size of households
Source of energy for heating
Source of energy for cooking
Access to various HH amenities
EU material deprivation index (adjusted) ¹⁴
Dwelling ownership
Preferences - source of income*
Economic situation
Absolute poverty rate PPP\$ 5.50 income based
Absolute poverty rate PPP\$ 4.30 expenditures based
Absolute poverty rate PPP\$ 2.15 income based †
Absolute poverty rate PPP\$ 2.15 expenditures based †
Relative poverty rate (60% equalized median income)
Poverty gap PPP\$ 4.30 income based
Poverty gap PPP\$ 2.15 income based †
Poverty gap PPP\$ 60% equalized median income
Gini coefficient
Ratio of poorest 20% v. richest 20%
Source of income
Structure of HHs expenditures 1
Structure of HHs expenditures 2
Financial security
Outstanding payments
Outstanding payments – type
Share of outstanding payments in HHs monthly income
Share of outstanding payments in HHs monthly income – type
Home production
Malnutrition
Migration
Household migration history

¹⁴ EU Severe material deprivation refers the population that cannot afford at least three of the following nine items:

(1) to pay their rent, mortgage or utility bills; (2) to keep their home adequately warm; (3) to face unexpected expenses; (4) to eat meat or proteins regularly; (5) to go on holiday; (6) a television set; (7) a washing machine; (8) a car; (9) a telephone. For more details see <http://bit.ly/2B0UW9B>

Indicator name
Support from abroad
Migration intention*
Migration targets*
Migration timing*
Lifestyle and awareness
Acceptability of bribing*
Acceptability of getting bribes*
Acceptability of tax evasion*
Acceptability of stealing food*
Acceptability of mixed marriages*
Acceptability of positive discrimination - language support*
Acceptability of positive discrimination - scholarships*
Acceptability of positive discrimination - support to employers*
Acceptability of positive discrimination - public employment*

* Based on random respondent questionnaire (Module 4)

2.4. Reasoning behind revisions

The revisions of the questionnaire took into account experience with use of Early Child Development Module findings from 2011; the need for further alignment of socio-economic data about poverty and sources of income with data collected during the accession process based on by Eurostat methodology; and the need to collect more information about differences between genders in terms of time use and domestic/care work and intra-household financial decision making. In addition, a set of questions about possession of personal documentation and household members who recently went abroad or returned from EU member states were added based on request of DG NEAR.

The experience showed that data from 2011 Early Child Development Module was not fully utilized for analytical purposes. In addition, UNICEF’s Multi Cluster Indicator Survey (MICS) provides more comprehensive insights about situation of children and families with children, including in Roma households. Given that UNICEF is preparing next round of MICS surveys in Western Balkans, a decision was made to retain only six questions from 2011 survey which were deemed to be crucial for the analysis of interactions between socio-economic status of a household, parenting techniques and educational outcomes for children.

The set of questions about sources of incomes was significantly expanded compared to 2011 survey based on an attempt to include relevant questions from EU-SILC. However, given the level of detail of EU –SILC questionnaire about incomes and their sources and the fact that regional Roma survey is designed as multi-indicators survey not all EU-SILC questions could be incorporated. In addition, pilot testing showed that even abridged version of EU-SILC income set of questions, including social assistance was complex to manage for Roma respondents: there were too many questions requiring too long recall period; and some were not applicable. Thus, a pragmatic approach was taken to adapt these questions to realities of respondents.

2.5. Translation

Questionnaires and additional fieldwork materials (introduction letter and interviewers' manual) were translated into the relevant languages for use in each country and in Kosovo. Languages used in each country and in Kosovo are presented in Table 2.

Table 2 : Languages used

COUNTRY/TERRITORY	Language(s)
Albania	Albanian
Bosnia and Herzegovina	Bosnian
the former Yugoslav Republic of Macedonia	Macedonian, Albanian
Montenegro	Montenegrin
Serbia	Serbian
Kosovo	Serbian, Albanian

Pilot questionnaires were translated into local languages by translators from local teams upon approval of UNDP and the World Bank. All translators were experienced and native speakers of the respective language(s). Ipsos Coordination Team checked local version of the questionnaire in terms of consistency of questions and codes with Master English questionnaire. Once when the translation process was completed, local versions of the questionnaire were submitted to UNDP for review.

Due to limited time, back-translation was performed on pilot questionnaires. All local versions of the pilot questionnaire were back-translated to English. This task was performed by translator(s) who was not involved in translation to local language. Members of Ipsos Coordination Team reviewed both translated and back-translated version of the questionnaire.

Following post-pilot finalization of questionnaire, local version of questionnaires were modified in line with modifications in master English questionnaire.

Single translation was used for additional fieldwork materials: interviewers' manual and introduction letter.

2.6. Constructing CAPI version

Fieldwork in all countries are conducted by CAPI (Computer Assisted Personal Interview).

CAPI was used as specified in the Terms of References, due to the possible complex routing and the overall length of the questionnaire. All routing in the questionnaire was pre-programmed which ensured that all relevant interview questions were asked. In addition, it also prevented human error on the part of the interviewer and allowed interviewers to concentrate on interviewing and interacting with the

respondent. An additional benefit of use of CAPI is the improved quality of data, given that it eliminates the need for manual data entry (as required for PAPI) at which point human error could occur.

CAPI interviewing had immense impact on the duration of the interview. It was calculated that the average duration of interview using PAPI method would last around 90 minutes. Using CAPI significantly reduced the time needed to conduct the interview – average duration was 57, 6 minutes (calculated at the end of the fieldwork).

Using single CAPI software platform (SM-S) in all countries provided several benefits. The questionnaire was scripted centrally (and overlaid with translations for use in each country). This provided for greater quality control of the CAPI script and ensured that a harmonized version with same checks built in was used in each country and territory. Centralized CAPI software allowed logical and consistency checks to be built in to the script. The programme incorporated all filters and routings, logical consistency checks and range checks (i.e. minimum/maximum values) for numerical questions.

Upon finalization, members of Ipsos Coordination team performed several rounds of checks making sure routing is working properly and all needed logical checks are implemented. This was done by checking both ‘dummy’ data and through ‘manual’ check of master English script. Once Master English script was finalized, local versions of the questionnaire were incorporated, making script available in all languages used in the survey. Upon creation of local versions of the script, country project managers checked local version(s) ensuring that CAPI script is in line with final local versions of the questionnaire and signed off local versions before the start of the fieldwork.

2.7. Interviewer training material

Draft version of interviewer manual was produced by Ipsos. World Bank provided instructions for questions related to Income Module. Review of draft version of manual was done by UNDP and World Bank. Final version of the manual was produced based on the proposed revisions and approval by UNDP and the World Bank.

Interviewer manual covered the following topics:

- Introduction and project background
- Definition of target groups: defining Roma settlements and non-Roma living in proximity
- Description of fieldwork procedures: random route, selection of household procedure, procedure related to working in team
- General guidelines for the interview
- Questionnaire instructions: instruction for all questions within questionnaire
- Tips for making contact
- Tips for gaining participation
- Tips for Improving Respondent Co-operation

Once finalized, interviewers’ manual was translated to local languages.



3. Interviewer selection and training

3.1. Project and survey briefing

Prior to main fieldwork, one-day briefing of national project managers was held in Belgrade on October 27th. Briefing was attended by national project managers, members of Central coordination team, UNDP and World Bank representatives.

All training sessions covered topics related to: i) aims and objectives of the study; ii) the sampling method (with the opportunity to discuss the implementation of adjustments per country/territory specific circumstances), and iii) detailed instructions on how the data collection should operate on the ground with special attention to process of contacting households and quality control procedures. The focus of the training was also on the random route technique for respondent selection and how to deal with different situations in the field (procedures of respondent selection; finding the primary sampling units; household selection). Full list of topics for project manager briefing is presented in Table 4.

Table 4. Briefing coverage

Background to the survey	<ul style="list-style-type: none"> • Brief description of the research • Purpose and importance of the survey
Project overview	<ul style="list-style-type: none"> • Requirements from each country/territory • Brief description of fieldwork procedures
Sampling and fieldwork procedures	<ul style="list-style-type: none"> • Sampling design, method of selecting households in the PSUs and respondents within households • Sampling procedures including review of contact sheet outcome codes and contact data that should be delivered • Description of the random selection of households and individuals • Rules about the number of visits to be made and their timing • Management of failed contacts: recall procedure, call appointment rules
Questionnaire	<ul style="list-style-type: none"> • Overall structure of the questionnaire • Structure of the questionnaire and definition of respondent for each Module
Quality control	<ul style="list-style-type: none"> • Reporting during fieldwork • Details of quality procedures and processes to be followed • Fieldwork control • Overview of deliverables

3.2. Interviewer selection and training

All Ipsos agencies have experienced panels of trained interviewers from which they selected interviewers for this survey. Candidate’s prior experience with random probability surveys and surveys on sensitive topics, in particular with Roma and other disadvantaged groups was considered as asset in the selection interviewer selection process

During the course of survey 47 – 68 interviewers teams per country/territory were recruited. Interviewers from Roma communities were engaged wherever possible. Roma interviewers were engaged by local NGOs and through personal contacts.

The number of interviewer’s teams recruited in each country/territory and number of engaged Roma interviewers is presented in the Table 5.

Table 5. Number of recruited interviewers’ teams

COUNTRY	No of interviewer teams recruited	No of Roma engaged interviewers
Albania	47	14
Bosnia and Herzegovina	58	/
the former Yugoslav Republic of Macedonia	55	28
Montenegro	54	5
Serbia	68	5
Kosovo	55	36

The main training sessions for interviewers were held in each country/territory participating in the survey in a period from November 23rd to December 1st.

Prior to fieldwork, the national project managers provided formal training sessions for all interviewers involved in the survey implementation. Interviewer training included all topics discussed in the project manager training. The interviewer training manual was updated to incorporate feedback from the briefing and pilot finding. Revised briefing material was distributed to the national teams. Ipsos provided training materials, which were tailored to the needs of countries/territory, ensuring that consistent instruction is given to all who participated in the survey.

As mentioned, Ipsos developed training programme for interviewers, to be delivered by national project managers. The training materials included:

- Training package (in English), which the national project managers used in organizing and delivering interviewer training in each country/territory. This package consisted of a standardized agenda for an interviewer training and presentations (in PowerPoint format) on the key issues to be covered.
- An interviewer's manual containing all relevant information for the interviewers to be able to successfully carry out the survey interviews.

The interviewer training included information on the following topics:

- Background and objectives of the study
- Contacting respondents: random route, how to contact respondents, rescheduling the interview;
- Collecting outcome information (e.g. non-response);
- The structure and content of the questionnaire: topics covered in the survey, the logic of the questionnaire, filters and routing, the purpose of specific questions;
- Interviewer instructions and skills: neutrality and tact, following the exact wording of the questions, not to rush, make assumptions or pre-empt answers;
- General advice on respondent and interviewer safety, including working in a team of two interviewers
- Fieldwork administrative issues

Although most of the interviewers had experience with random probability surveys, due to the fact that sampling requirements for this survey were complex it was important to organize a dedicated sampling session, where the approach was explained in detail including the review of contact sheets. This session included a couple of practical exercises where interviewers had to complete the contact sheet as they would if in the field.

Another dedicated session was devoted to the logic and content of the questionnaire. While the use of CAPI meant that interviewers did not need to be familiar with the complex routing, it was essential that they had a thorough understanding of the questionnaire and were confident with the flow of the questionnaire so that they could collect accurate data and focus on the respondent during the interview.

All interviewers were given a copy of the training material so that they were able use them as live documents to refer to during fieldwork.

Engaging Roma organizations

In projects like this, participation of Roma organizations is very valuable not only in sampling process, but in fieldwork implementation as well. In all countries/territory certain level of cooperation with Roma NGOs was established.



Albania

Local agency in Albania contacted 8 Roma NGOs and established cooperation with only 2 of them - Roma Active & IRCA. These organizations provided local agency with draft list of people interested in engaging in the fieldwork.

Bosnia and Herzegovina

Local agency in Bosnia and Herzegovina contacted 3 organizations and 2 individuals. Two organizations (Kali Sara and Bolja buducnost) never replied to contact attempts, while third one, Romale Kakanj requested considerable amount of money for being included in this survey. From 2 individuals, assistance in location of Roma settlements in process of maps preparation was given by Roma representative in Bosnia and Herzegovina.

Kosovo

Local agency in Kosovo established cooperation with 1 Roma NGO working with students from RAE community – Roma Versitas. They supported local agency by recommending a number of persons of Roma, Ashkali and Egyptian background as interviewers, who were later engaged during fieldwork.

Voice of Roma Ashkali and Egyptian NGO was also contacted but cooperation was not established.

Montenegro

Local agency in Montenegro established cooperation with 2 Roma NGOs – HELP and Mladi Romi. They supported local agency by recommending persons interested in participation of data collection. 4 more organizations were contacted, but cooperation was not established.

Serbia

In Serbia, cooperation was established with 3 NGOs - Edukativni centar Roma, Udruženje romskih studenata and Vojvođanski romski centar za demokratiju. These NGOs provided assistance in process of identification and mapping of Roma settlements.

the former Yugoslav Republic of Macedonia

Local agency in the former Yugoslav Republic of Macedonia established cooperation with 8 NGOs : CDRIM /Romski Resursen Centar, Sumnal, Ambrela, Roma Versitas Stip, NVO KHAM, Bairska Svetlina, Mesecina and Romano Vila. Beside engaging interviewers these organizations recommended, premises of NGO Mesecina and Roma Veritas were used for additional interviewers trainings.



4. Sampling

This section provides a description of the target populations and sampling methodologies used in Western Balkans.

The initial phases included identifying the most appropriate sample frames that could be used for developing an appropriate sample plan for each country which could best meet the survey requirements given the available sources.

4.1. Target populations and sample requirements

As in previous, 2011 survey, this survey included both Roma and non-Roma households.

As main target population, survey sampled Roma households in Roma settlements and areas of concentrated Roma population and in settlements where share of Roma is 10%-40% ('lower density' areas) who identify themselves as Roma. This meant that, within Roma sample, sampling methodology didn't cover assimilated Roma, Roma who are dispersed and integrated in other communities where they represent less than 10% of total population, as well as Roma who do not self-identify as Roma.

Beside Roma sample, survey sampled non-Roma population living in close proximity to Roma settlements or 'lower density areas'.

Roma settlement is defined as a part of administrative settlement where Roma population represent more than 40% of the population (in the respective section of the locality).

The non-Roma population living in close proximity to Roma was determined based on the Roma sample. The non-Roma living in close proximity to Roma were defined as the non-Roma living in close geographic proximity (down the road to the center of municipality) to primary sampling units where the interviews with Roma were conducted. For purpose of the survey, non-Roma population living in close proximity to Roma was defined as non-Roma population which lives in the range of 300 meters from sampled Roma settlement.

For the purpose of this survey, the term 'Roma' refers to persons who self-identify as 'Roma' or as one of the other groups, which are subsumed under the term 'Roma' (Ashkali, Egyptian, Lom, Dom, Abdal, Gypsy). Household was classified as Roma or non-Roma based on explicit self-declaration of ethnicity given by head of household.

Random probability samples - multistage clustered random probability sample design was used in all surveyed countries/territories (described within section 4.4).

4.2. Sample sizes: Roma and non-Roma

In total, 4 592 Roma households and 2 168 non-Roma households from Western Balkans participated in the Regional Roma Survey. As in 2011 survey, 750 Roma and 350 non-Roma households (as minimum) were sampled in each country/territory that participated in the survey.

The total number of completed interviews per country/territory and per sample is presented in Table 6.

Table 6: Total number of completed and accepted interviews per country/territory and sample

COUNTRY	No of interviews (total)	No of interviews in Roma sample	No of interviews in non-Roma sample
Albania	1116	766	350
Bosnia and Herzegovina	1102	751	351
the former Yugoslav Republic of Macedonia	1118	756	362
Montenegro	1139	780	359
Serbia	1152	764	388
Kosovo	1133	775	358

In the 2017 survey, a subsample of Roma households living in areas with lower density of Roma population was included. Areas of lower density are defined as those where the share of Roma is from 10 % to 40 % in total population. Table 7 presents breakdown of sample with number of households living in areas higher and lower density of Roma population.

Table 7: Roma and non-Roma sample disaggregated by density areas.

COUNTRY/ Territory	Roma			Non-Roma		
	No of households in Roma settlements (>40%)	No of households in lower density areas (10-40%)	Total number of households	No of households in Roma settlements (>40%)	No of households in lower density areas (10-40%)	Total number of households
Albania	383	383	766	169	181	350
Bosnia and Herzegovina	325	426	751	144	207	351
the former Yugoslav Republic of Macedonia	413	343	756	182	180	362
Montenegro	436	344	780	195	164	359
Serbia	368	396	764	184	204	388
Kosovo	484	291	775	223	135	358



4.3. Sample frames and mapping of information available for sampling

Given that there is a lack of official data on size of Roma population in most of the countries/territory included in the survey, determining a suitable sample frame in each country that could identify the target groups within foreseen timeframe and resources was a critical part of the survey preparatory work. Section below specifies possible sampling approaches that were taken into account in a process of sample preparation along with selected sample approach.

Albania

Data available:

- Census 2011 data at the level of census enumeration areas (CEAs)
- 2011 Roma NGO data at the level of 'Roma settlements'

Sampling approach: After reviewing the Census 2011 data on CEA level, discrepancies between this data and the NGO data on 'Roma settlements' were noticed. Some of the large settlements on the NGO list seem not to be covered in Census. So, if we sampled CEAs from the Census list only, we would have excluded a large section of population living in 'Roma settlements' according to the Roma NGO. Therefore, sampling based on 2011 Census data on CEA level was deemed as not adequate. Instead, two possible solutions explained below were considered

Sampling approach 1: The Census 2011 data at the level of administrative settlements, rather than CEA level, can be used. This allows covering the most of the 'Roma settlements' from the NGO list. As the settlements can get quite large, the indicator showing the share of Roma within total population could not help in identifying exact areas where Roma live. Hence, the settlements need to be divided into smaller units suitable for use as PSUs. Therefore, the settlements where Roma live can be divided into units of approximately equal size in terms of Roma population. The aim is to have approximately 30 Roma households in each unit. For the settlements selected in the sample the local teams will need to identify areas where Roma live (in concentrations higher or lower than 40%) and divide them into units of approximately 30 Roma households. In this way both 'Roma settlements' from the NGO list, and areas where Roma live in lower concentrations have chance to be selected. Once these units are identified and numbered, the central sampling team informs the local team which units are selected in the sample. The same approach was followed in the 2011 survey (with the NGO list used as the sampling frame). Within each unit (PSU) households are to be selected randomly via random walk approach.

Coverage (approach 1): Settlements with less than 80 Roma are excluded due to practical reasons. This gives the coverage of 84% of Roma enumerated in Census 2011. At the same time 75% of population who live in 'Roma settlements' on the NGO list would be covered.

Sampling approach 2: The 2011 Roma NGO data at the level of ‘Roma settlements’ can be used. This would assume following the same approach used in the 2011 survey. The ‘Roma settlements’ would be divided into smaller units suitable for use as PSUs – units of approximately equal size in terms of Roma population. The aim is to have approximately 30 Roma households in each unit. For the ‘Roma settlements’ selected in the sample the local teams will divide them into units of approximately 30 Roma households. Once these units are identified and numbered, the central sampling team informs the local team which units are selected in the sample. Within each unit (PSU) households will be selected randomly via random walk approach.

The drawback of this method is potentially not covering Roma who live integrated in lower concentration areas. However, a consistency between the sample allocation and the population covered on the field would be a benefit.

Coverage (approach 2): 97,4% of Roma listed in the 2011 Roma NGO data

Selected sample frame: 2011 Roma NGO data at the level of ‘Roma settlements’

Selected sampling approach: The 2011 Roma NGO data at the level of ‘Roma settlements’ was used. This assumed following the same approach used in the 2011 survey. The ‘Roma settlements’ were divided into smaller units suitable for use as PSUs – units of approximately equal size in terms of Roma population. The aim was to have approximately 30 Roma households in each unit. For the ‘Roma settlements’ selected in the sample the local teams divided them into units of approximately 30 Roma households. Once these units were identified and numbered, the central sampling team informed the local team which units are selected in the sample. Within each unit (PSU) households were selected randomly via random walk approach.

Coverage within selected sampling approach: 97.4% of Roma listed in the 2011 Roma NGO data

Bosnia and Herzegovina

Data available:

- Census 2013 data at the level of administrative settlements
- (UNDP was exploring possibility of obtaining more precise data on where Roma live from the Ministry of Human Rights and Refugees – not taken into account in the proposed approach)

Sampling approach: The Census 2013 data at the level of administrative settlements was used. As the settlements can get quite large, the indicator showing the share of Roma within total population could not help in identifying exact areas where Roma live. Hence, the settlements needed to be divided into smaller units suitable for use as PSUs. Therefore, the settlements where Roma live were divided into units of approximately equal size in terms of Roma population. The aim was to have approximately 30 Roma households in each unit. For the settlements selected in the sample the local teams needed to identify areas where Roma live (in concentrations higher or lower than 40%) and divide them into units of approximately 30 Roma households. Once these units were identified and numbered, the central sampling team informed the local team which units are selected in the sample. The same approach was

followed in the 2011 survey. Within each unit (PSU) households was selected randomly via random walk approach.

Coverage: Settlements with less than 80 Roma was excluded due to practical reasons. This gives the coverage of 77% of Roma enumerated in Census 2013.

Kosovo

Data available:

- Census 2011 data at the level of administrative settlements
- Census 2011 data at the level of census enumeration areas (CEAs)

Sampling approach: After reviewing the Census 2011 data on CEA level, it was noticed that RAE population registered in Census is quite dispersed across CEAs. Hence, these resulted in considering two possible approaches in Kosovo, as explained below.

Sampling approach 1: The Census 2011 data at the level of administrative settlements. As the settlements can get quite large, the indicator showing the share of Roma within total population could not help in identifying exact areas where Roma live. Hence, the settlements needed to be divided into smaller units suitable for use as PSUs. Therefore, the settlements where Roma live would have to be divided into units of approximately equal size in terms of Roma population. The aim was have approximately 30 Roma households in each unit. For the settlements selected in the sample the local teams would need to identify areas where Roma live (in concentrations higher or lower than 40%) and divide them into units of approximately 30 Roma households. Once these units are identified and numbered, the central sampling team would inform the local team which units are selected in the sample. The same approach was followed in the 2011 survey. Within each unit (PSU) households would be selected randomly via random walk approach.

Coverage 1: Settlements with less than 120 Roma would be excluded due to practical reasons. This gives the coverage of 82% of Roma enumerated in Census 2011

Sampling approach 2: The census enumeration areas to be used as Primary Sampling Units (PSUs). The PSUs will be selected with probabilities proportional to number of Roma in each PSU. Within each PSU households would be selected randomly via random walk approach.

However, no geographical data connected to the CEA list was available at the time – only non-unique IDs of CEAs were available along with the number of Roma households and total number of households in each CEA. To be able to proceed with sampling, this list needed to include codes of administrative settlements (as provided in separated list shared by the NSO). Then, the NSO would need to provide maps or descriptions of borders for each CEA that will be selected in the sample.

Coverage (approach 2): Census enumeration areas with less than 15 Roma households would be excluded due to practical reasons. This gives the coverage of 66% of Roma enumerated in Census 2011.

Selected sampling frame: Census 2011 data at the level of administrative settlements.

Selected sampling approach: The Census 2011 data at the level of administrative settlements was used. As the settlements can get quite large, the indicator showing the share of Roma within total population could not help in identifying exact areas where Roma live. Hence, the settlements needed to be divided into smaller units suitable for use as PSUs. Therefore, the settlements where Roma live were divided into units of approximately equal size in terms of Roma population. The aim was to have approximately 30 Roma households in each unit. For the settlements selected in the sample the local teams identified areas where Roma live (in concentrations higher or lower than 40%) and divided them into units of approximately 30 Roma households. Once these units are identified and numbered, the central sampling team informed the local team which units are selected in the sample. The same approach was followed in the 2011 survey. Within each unit (PSU) households are selected randomly via random walk approach.

Coverage within selected sampling approach: Settlements with less than 120 Roma were excluded due to practical reasons. This gives the coverage of 82% of Roma enumerated in Census 2011

the former Yugoslav Republic of Macedonia

Data available:

- Census 2002 data at the level of administrative settlements
- (World Bank was exploring possibility of obtaining updated data from the Ministry of Education – not obtained)

Sampling approach: The Census 2002 data at the level of administrative settlements was used. As the settlements can get quite large, the indicator showing the share of Roma within total population could not help in identifying exact areas where Roma live. Hence, the settlements needed to be divided into smaller units suitable for use as PSUs. Therefore, the settlements where Roma live were divided into units of approximately equal size in terms of Roma population. The aim was to have approximately 30 Roma households in each unit. For the settlements selected in the sample the local teams identified areas where Roma live (in concentrations higher or lower than 40%) and divided them into units of approximately 30 Roma households. Once these units are identified and numbered, the central sampling team informed the local team which units are selected in the sample. The same approach was followed in the 2011 survey. Within each unit (PSU) households are selected randomly via random walk approach.

Coverage: Settlements with less than 120 Roma was excluded due to practical reasons. This gives the coverage of 95% of Roma enumerated in Census 2002

Montenegro

Data available:

- Census 2011 data at the level of administrative settlements
- Census of RAE population 2008 at the level of municipalities

Sampling approach: As Census of RAE population is considered to be more reliable source and was used as basis for allocation of the sample per municipalities. As the municipalities can get quite large, the

indicator showing the share of Roma within total population could not help in identifying exact areas where Roma live. To be able to narrow down the locations where Roma live, we used the settlement level data from Census 2011. However, the settlements are also quite large, so these needed to be divided into smaller units suitable for use as PSUs. Therefore, the settlements where Roma live were divided into units of approximately equal size in terms of Roma population. The aim was have approximately 30 Roma households in each unit. For the settlements selected in the sample the local teams identified areas where Roma live (in concentrations higher or lower than 40%) and divided them into units of approximately 30 Roma households. Once these units were identified and numbered, the central sampling team informed the local team which units are selected in the sample. The same approach was followed in the 2011 survey. Within each unit (PSU) households are selected randomly via random walk approach.

Coverage: Settlements with less than 20 Roma was excluded due to practical reasons. This gives the coverage of 95% of Roma enumerated in Census 2011.

Serbia

Data available: Census 2011 data at the level of statistical sectors

Sampling approach: The list of statistical sectors was used. Statistical sectors were selected with probabilities proportional to number of Roma persons in each sector. Large statistical sectors were divided into clusters of approximate size of 30 Roma households. Sample of PSUs is selected from that list of clusters and within each PSU households were selected randomly via random walk approach.

Coverage: Settlements with less than 80 Roma were excluded due to practical reasons. Within settlements with more than 80 Roma, statistical sectors with less than 30 Roma were excluded, as less than 7 Roma households are expected in the PSU. This gives the final coverage of 77% of Roma enumerated in Census 2011.

4.4. Overview of sampling methodology

The sample selection included following stages, and each of these is described further below:

- **Stage 1:** Stratification and selection of settlements
- **Stage 2:** Selection of Primary Sampling Units (PSUs)
- **Stage 2:** Selection of households
- **Stage 3:** Selection of respondents in the household

Stratification

Sample is stratified proportionately. Following variables were used to stratify the PSUs:

1. **Regions:** The PSUs was stratified by the geographical regions in a given country. This ensured that the sample had proportional representation of the geographical distribution of the Roma population within country.

2. Rurality/urbanity or other indicator of population density: At the second stage, PSUs are ordered by their rurality/urbanity indicator within each geographical region (This was used in countries where information was available).

Sample Strata Estimation and allocation of PSUs

Based on planned total sample size and number of interviews per PSU, number of PSU was estimated using following formula:

$$\text{No_of_PSUs} = \text{No_of_interviews_in_strata_cell} / \text{No_of_interviews_per_PSU}$$

After creating initial strata and estimating their size (i.e. their proportion in population and in sample), clusters were distributed to initial strata using proportional criterion – percentage of clusters per stratum is proportional to its size, that is share of Roma population within stratum in total Roma population. Number of cluster in cell (strata=region*type_of_settlement) is calculated using following formula:

$$\text{No_of_PSUs_in_strata_cell} = (\text{Population_in_strata_cell} / \text{Roma_population_in_total}) * \text{No_of_PSUs}$$

Selection of settlements and Primary Sampling Units (PSUs)

It was necessary to create list of clusters for selecting Primary Sampling Units (PSUs) in all countries/territory. Firstly, list of settlements was selected with PPS algorithm (where size is number of Roma persons). This implies that number of allocated clusters per settlement was proportional to the number of Roma persons. Primary Sampling Units (PSUs) were defined as clusters within chosen settlements consisting of approximately 30 Roma households.

In the first stage, list of clusters for sampling was developed in each country/territory. Within each settlement included in sample, all areas where Roma live in concentrations above 10% were identified. Out of these areas, the specified number of clusters of approximately 30 Roma households were defined. Number of clusters that should be created in each settlement was calculated based on available data on the number of Roma in each settlement. If an identified Roma settlement / area where Roma live had more than 30 Roma households, it was divided on the map into more clusters. After mapping clusters, number of clusters and selection of cluster per settlement was made. PSUs within settlements were selected randomly with equal probability.

Sample allocation is presented in the Table 8:

Table 8: Sample allocation

		PSU/cluster distribution
Country/territory	Geographical regions/municipalities	Total
Albania	Berat	6
	Durres	14
	Elbasan	18
	Fier	20
	Gjirokaster	3



	Korce	16
	Lezhe	4
	Shkoder	1
	Tirane	26
	Vlore	6
	TOTAL	114
Bosnia and Herzegovina	Herzegovina-Neretva Canton	4
	Sarajevo Canton	17
	Central Bosnia Canton	12
	Tuzla Canton	26
	Una-Sana Canton	4
	Zenica-Doboj Canton	28
	Brcko Distrikt	4
	Republika Srpska	19
	TOTAL	114
the former Yugoslav Republic of Macedonia	East	21
	NorthWest	25
	Skopje	50
	SouthWest	18
	TOTAL	114
Montenegro	Bar	4
	Berane	8
	Bijelo Polje	4
	Budva	2
	Cetinje	1
	Herceg Novi	3
	Kotor	1
	Niksic	12
	Podgorica	67
	Rozaje	1
	Tivat	5
	Ulcinj	6
	TOTAL	114
	Serbia	Vojvodina
Belgrade		21
West		7
Central		9
East		9
South		35
TOTAL		114
Kosovo	North	4
	Northwest	44
	Pristina (central part)	46
	Southeast	2
	Southwest	19
	TOTAL	113



For each cluster selected in Roma sample, cluster in close proximity inhabited by the non-Roma population was selected for non-Roma sample. Non-Roma populations living in close proximity to Roma were defined as non-Roma population that lives in the range of 300 meters from our Roma settlement.

Similar to UNDP Roma survey in 2011, within each PSU approx. 10 interviews were completed (7 to 8 interviews with Roma population and 3 interviews with non-Roma living in close proximity to Roma). In order to have required number of interviews within non-Roma sample, in certain percentage of PSUs 4 interviews in non-Roma sample were administered, while in all other PSUs 3 interviews in non-Roma sample were administered.



5. Piloting

5.1. Methodology

Before to launching the main fieldwork, the initial pilot interviews were conducted in all target countries/territory. Preliminary testing was introduced in order to:

- Check the usage of each module and the overall questionnaire flow (for example: Are the question skips correct and overall logic of the questionnaire? Is there a continuity between the questions and the modules?)
- Check for the flaws in the administration of the questionnaire (such as routing)
- Determine the cross-country stability of the questionnaire language and test the quality of translation.

Pre-test was conducted in all countries/territory participating in the survey. Pre-test included 10 Roma and 10 non-Roma households in each country/territory. The fieldwork timeline for each country is shown in the Table 9. Since pilot was planned in the period when there was no final decision on involvement of Turkey in research, pre-test survey was completed in Turkey as well.

Table 9: Start and ending dates of the pre-test surveys

COUNTRY/ territory	Start Date	End date
Albania	13.10.2017	15.10.2017
Bosnia and Herzegovina	13.10.2017	15.10.2017
the former Yugoslav Republic Macedonia	13.10.2017	16.10.2017
Montenegro*	14.10.2017	16.10.2017
Serbia	13.10.2017	15.10.2017
Kosovo	13.10.2017	16 .10.2017
Turkey	13.10.2017	15.10.2017

*Due to the local elections taking place on 15th October and 14th being election silence day, interviewers worked only on 13th and 16th October.

The pre-test surveys were conducted on the majority language in each country. In each country 6-10 interviewers were engaged in pre-test survey.

5.2. Key findings, recommendations and actions as a result of the pilot

Several issues were reported during pre-test. The most important of them were:

- Length of the questionnaire and interviewing several household members caused respondents impatience and lack of concentration during interviewing process.
- In Roma households, not all dates of birth were familiar to members, so the usage of the birthday method to randomly select respondents for particular modules needed to be reconsidered.
- Sometimes it was difficult to separate household members from each other. Reasons for this were various: in some cases this was not possible due to lack of space within dwelling unit, but in some cases interview was often interrupted by head of household or head of household didn't allow randomly selected respondent to be interviewed in privacy.
- Questions about money and earnings induced distrust and lead to false answering and further refusal.
- The skip conditions were sometimes hard to apply on PAPI format, since interviewers made pauses in order to check the consistency of the answers.
- Interviewing several household members sometimes caused impatience and nervousness. The same happened in case one household member was answering more than one module - after completing one module they were becoming nervous and in general – sometimes it seemed like they were giving answers without really thinking about their meaning.

The remaining sections of this chapter presented in Table 10 provide the key findings, recommendations and actions taken with regards to questionnaire and separate modules.

Table 10: Key findings, recommendations and actions taken as the result of the pilot

Key findings	Recommendation	Actions taken
Dates of birth are often not familiar to members, so the usage of the birthday method should be reconsidered	Changing method of selection respondent to random selection done within CAPI script based on household grid from Module 1.	Implemented as necessary during main stage fieldwork.
Module 1		



<p>Overall, the feedback is that there are too many questions in the Module 1 and that respondents are starting to lose attention in this section.</p> <p><u>E5a, F1</u> – A lot of the respondents declined to answer these questions, mostly because some of them were engaged in undeclared work and were afraid that the interviewers will report them to authorities, which may cause them to lose their job. Also, having in mind that most of them are working on-and-off jobs, it was really difficult for them to recall how many months were they paid or even put it in terms of months when it might have really been just a few days in each month.</p> <p><u>Wages and salaried income section</u> – The greatest difficulties were observed within this part of the questionnaire. The whole section is too long and detailed, provokes additional questions from the respondent because most of them are not familiar with exact terms, so additional clarifications are needed, questions are considered as sensitive and too detailed. After several questions for only one household member, respondents are getting nervous and start losing their attention and patience and give answers that aren't well thought out, only to finish interviewing faster.</p> <p><u>Social insurance contributions section</u> – respondent were often not familiar with the fact if employer payed social insurance contributions. In case they knew that these were payed, they were not aware of the amount.</p> <p><u>Pension and social benefits</u> – the same comment as for section Wages and</p>	<p>Removing sections Wages and salaried income, self-employment income and pension and social benefits from Module 1. In case these questions are mandatory – they can be asked within Module 3 – but only grouped in categories and in total amount for whole household. Estimation was that, in this way, length of the questionnaire will be reduced by at least 15 minutes.</p>	<p>Wages and Salaried Income section reorganized and shortened Added instructions and modified routing where needed</p>
--	---	---



Key findings	Recommendation	Actions taken
<p>salaried income. In addition, respondents were not able to distinguish certain allowances on question F8 from others – allowances 1 and 4 are considered as part of regular income</p>		
Module 2a		
<p>In great majority, head of household insisted on answering all the questions Kindergarten / Pre-School Characteristics subsection – exact age of kindergarten/pre-school age should be defined in order to make selection of child more appropriate. EC7 & EC8 – it was difficult for interviewer to ask those questions, considering the housing condition of some Roma households.</p>	<p>Removing Module 2b – head of household is requesting to provide answers within this section, so interviewing primary care taker is often not possible. This way, there are no benefits of including this Module as a separate one. Instead of having one separate module, we suggested to reorganize and shorten questionnaire the following way :</p> <ul style="list-style-type: none"> • Moving questions EC2, EC5, EC10 and EC11 into Module 1 – within Education Profile Card (with revising skip patterns within this Module) • Removing questions EC6, E12 and section Parenting Techniques • Moving and reformulating question EC13 into Module 3 – with costs provided for all children of kindergarten and preschool age – costs can be separated into 1. Monthly fee for kindergarten/preschool, 2.Costs related to kindergarten/preschool (including transportation, books, pencils, and other supplies and special clothes/uniform), 3. Food (in kindergarten/preschool – snack, lunch...) 	<p>Module 2a removed as separate module, while needed questions are incorporated into Module 1 :</p> <ul style="list-style-type: none"> • Questions EC2, EC5, EC6, EC10 and EC11 moved within Module 1 to be asked households members aged 0-6 years • Questions EC7, EC8, EC9, EC10, EC12, EC13 and EC14 moved at the end of Module 1 to be administered to the primary care taker of randomly selected child among those who are attending daycare center, nursery, crèche, preschool or kindergarten (answer 2 on EC2). In case primary care taker is not available or refused to be interviewed, answers were provided by the head of household



Key findings	Recommendation	Actions taken
Module 2b		
<p>In Roma households, women (especially younger) were not allowed to answer.</p> <p>G1:</p> <ul style="list-style-type: none"> • Too complicated, respondents could not remember the length of the activities’ performed. • The feelings of privacy violations were also reported on this question. • Not all activities that respondents reported could have been classified under the given code scheme 	<p>Removing Module 2b – having in mind different issues and time needed for completing this module by all women aged 16+ Ipsos suggested to keep this Module only as a part of Module 4 while adding detailed instruction concerning codes and coding primary and secondary activity. In that case ‘Financial decision-making power in the household’ section would be administered to women only.</p>	<p>Removed Module 2b as separate module. Section ‘Financial decision-making power in the household’ section moved to module 4 and administered to women only.</p>
Module 3		
<p>Some questions regarding income and living expenses discouraged the respondents (both Roma and non-Roma), one of them even felt humiliated and commented that: “It seems as the questionnaire is retrieving the worst things about myself”. Generally, there was a tendency to present the situation better than it is.</p> <p>Section G : Household income – similar to income section, respondents were hesitating to provide answers; In addition, respondents were not able to distinguish different forms of social assistance from others including those already within Module 1</p>	<ul style="list-style-type: none"> • Modifying questions G1, G2 from Module 3 – these can be included on the level of household together with questions to be revised from Module 1 • Removing question G5 – not applicable and understandable for great majority of both Roma and non-Roma households • Modifying questions G6 and G7 – to be included in grouped categories of income 	<ul style="list-style-type: none"> • Section G. Self-employment Income added (reorganized previous version from Module 1) • Section H: Pension and social benefits in the last 12 months (reorganized previous version from Module 1) • Section J: Social assistance (reorganized previous version) • Section K: Transfers between households, financial transactions, leasing and agricultural activity (reorganized previous version)
Module 4		
<p>No specific issues</p>	<ul style="list-style-type: none"> • Cutting some questions due to overall length of the questionnaire 	<ul style="list-style-type: none"> • Questionnaire modified removing some questions

6. Fieldwork

6.1. Fieldwork dates and progress (per country/territory)

Table 11 provides details of the final sample size achieved in each country by target group and the fieldwork dates.

Table 11: Final achieved sample and fieldwork dates

COUNTRY / territory	No of interviews (total)	No of interviews in Roma sample	No of interviews in non-Roma sample	Fieldwork Dates		Dates fieldwork paused	
				Start	End	Start	End
Albania	1116	766	350	26.11.2017	4.1.2018	1.12.2017	7.12.2017
Bosnia and Herzegovina	1102	751	351	2.12.2017	31.12.2017	NA	NA
The former Yugoslav Republic of Macedonia	1118	756	362	30.11.2017	31.12.2017	NA	NA
Montenegro	1139	780	359	2.12.2017	26.12.2017	NA	NA
Serbia	1152	764	388	26.11.2017	31.12.2017	NA	NA
Kosovo	1133	775	358	28.11.2017	30.12.2017	NA	NA

Table 12 provides details of how fieldwork progressed on a weekly basis.

Table 12: Weekly progress per country

COUNTRY / territory	No of completed interviews					
	03/12/2017	10/12/2017	17/12/2017	24/12/2017	31/12/2017	04/01/2018
Albania	67	196	433	789	1112	1116
Bosnia and Herzegovina	5	94	413	949	1102	
The former Yugoslav Republic of Macedonia	113	477	819	987	1118	
Montenegro	99	589	998	1138	1139	
Serbia	174	444	828	1115	1152	
Kosovo	63	234	563	1028	1133	

6.2. Contacting respondents

Each interviewers' team was assigned with the appropriate number of primary sampling units (PSU) in which they should identify and contact a certain number of Roma and non-Roma households. Households were contacted using random route procedure in each PSU. Random route was performed to ensure that every household in a PSU has equal chance to be selected and that every household is randomly selected.

Contacts was made only at **residential addresses**. Office buildings, businesses, schools, hospitals, public buildings, factories, workshops, supermarkets and shops were excluded from the count along the route.

First address was in the centre of PSU, place that was roughly equally distant to the borders of the PSU. First address was selected by the field coordinator. Selection of next addresses was done by selecting every 3rd house or flat in towns and every 3rd in rural areas. Flats in apartment buildings were counted as separated addresses.

6.3. Recruitment of respondents and interview format

In all countries, interviewers worked in teams. Each team consisted of one male and one female interviewer. Due to the requirement that interviewers must visit household in teams, script was programmed in a way that allows only working as a team so interviewers were not able to enter separate questionnaire if they are working alone.

Within each team, each interviewer had a role of 'main' or 'second' interviewer. Difference between these two was reflected in a number of information to be collected by each of them. This practically means that 'main' interviewer was the one who was completing contact sheet for each visited address, completing administrative module and post-interview assessment of the household. Only after 'main' interviewer completed contact sheet for the household that agreed to be interviewed, 'second' interviewer was allowed to enter separate questionnaires for household members.

After they approached selected address, 'main' interviewer entered address and description of selected address/household. If contact with selected address/household was not achieved, this was marked within contact sheet and information on main reason for not making contact was entered. This household was visited again. If contact with selected address/household is achieved, interviewers introduced themselves and the survey and asked to talk with head of household.

When they achieved contact with head of household, interviewers introduced purpose and aim of the survey and provided respondent with cover letter.

- In case head of household weren't agree to be interviewed, this was marked in contact sheet along with reasons for this. In case head of household refused to be interviewed, interviewers marked reason, gender and age of the person who refused to take part in the survey in contact sheet. After completing this, they continued random walk until next selected address/household.
- In case head of household agreed to be interviewed, interviewers asked following questions: Which ethnicity do you feel you belong to? and Do you feel belonging to another ethnicity in

addition to what you stated above? Respondent provided spontaneous answers, without reading questions by the interviewers' team.

- In case answer on any of these two questions, head of household answered Roma/Ashkali/Egyptian/Lom/Dom/Abdal – this household was considered as Roma household
- In all other cases, household is considered as part of non-Roma sample.
- Screening criteria was included in CAPI script, so Roma/non-Roma sample was automatically calculated.
- After script provided interviewers with information if household is considered as Roma/Non-Roma, they needed to check the number of interviews they have already completed with Roma and non-Roma households. If sufficient number of interviews was already met in either Roma or Non-Roma sample, they thanked respondent for willingness to cooperate and left.

After head of household agreed to be interviewed and it was determined within which sample this household should be included, interviewers proceeded with interview. At this point, 'main' interviewer needed to complete quick demographic card of household members – collecting information on their age, gender and country in which they were born. Meanwhile 'second' interviewer, entered only household number he/she was told by 'main' interviewer (based on number script gave to this household), after which they needed to complete the same quick demographic card of household members – collecting information on their age, gender and country in which they were born. 'Main' and 'second' interviewer entered demographic card of household members simultaneously – at the same time and in the same manner and order – only this way they were able to proceed with rest of the questionnaire.

After both 'main' and 'second' interviewer completed the same demographic card, list of household members and their assigned questionnaires were shown on the screen.

After list of household members and their assigned questionnaires were shown on the screen, interviewers needed to agree which one will complete which interview. After they made agreement on questionnaire(s) to be completed, they marked boxes next to the selected respondents/modules. This led them to the main screen, where they were able to enter only questionnaires they previously selected.

The interview for Module 1. Household members profile with the head of household or other knowledgeable member had to be performed first. Other modules could be done simultaneously if different persons were respondents.

6.4. Interview length

Recorded average duration of the CAPI questionnaire is 57,6 minutes.

Table 13 provides the average interview lengths across all survey countries and for all modules.



Table 13: Average interview lengths

COUNTRY / territory	Module 0	Module 1	Module 3	Module 4	Total
Albania	3,6	19,7	19,5	14,1	56,9
Bosnia and Herzegovina	3,8	20,1	19,6	14,5	58
the former Yugoslav Republic of Macedonia	3,5	19,9	19,8	14,7	57,9
Montenegro	3,4	20,2	19,7	14,4	57,7
Serbia	3,5	20,3	20,1	14,3	58,2
Kosovo	3,8	19,9	19,5	13,9	57,1

6.5. Fieldwork outcomes

Table 14 provides a summary of total number of contacted households and the associated response rate.

Table 14: Number of contacted households and the associated response rate

COUNTRY	No of contacted households	No of completed interviews	RR%
Albania	1625	1116	69%
Bosnia and Herzegovina	1624	1102	68%
the former Yugoslav Republic of Macedonia	2087	1118	54%
Montenegro	1798	1139	63%
Serbia	1897	1152	61%
Kosovo	1861	1133	61%



6.6. Fieldwork control

The fieldwork and project managers was responsible for coordinating and managing the back-checks, which was carried out on at least 20% of interviews for this survey and was conducted on an ongoing basis throughout fieldwork.

Fieldwork control involved re-contacting the respondent to check that the interview was conducted correctly. These were conducted either by telephone or face-to-face by a separate quality control team. Using standardized questionnaire that was approved by UNDP, quality control team completed control asking about the interview, such as the date of the interview, length of the interview, gender of the interviewer and key factual questions to check the data recorded is accurate.

Table 15 provides summary of fieldwork control with number of interviews excluded after control.

Table 15: Number of controlled interviews and number of interviews excluded after control

COUNTRY	No of controlled interviews	No of interviews excluded after control
Albania	240	17
Bosnia and Herzegovina	230	18
the former Yugoslav Republic of Macedonia	250	14
Montenegro	230	15
Serbia	235	20
Kosovo	240	18

6.7. Data processing and cleaning

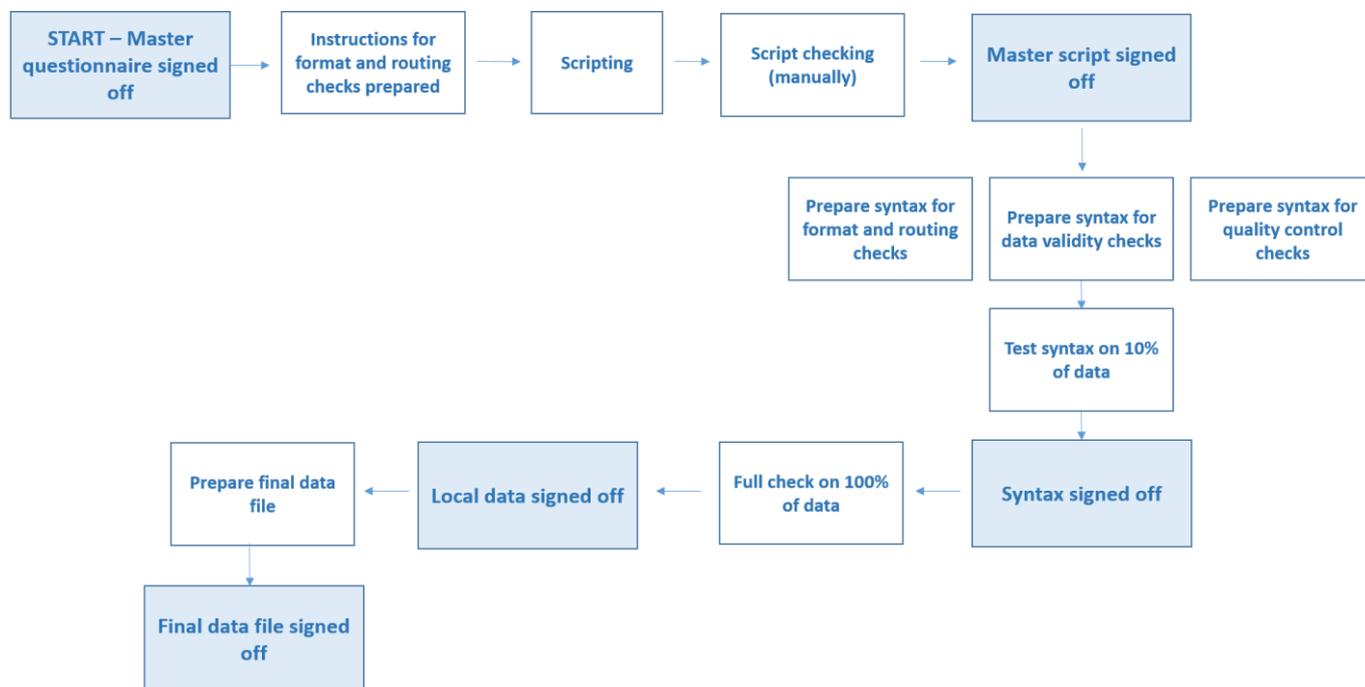
Due to the fact that two interviewers were completing questionnaire in the same households and at the same time (without previous synchronization) scripting process was complex. However, questionnaire with incorporated all skips and the most important checks lowered number of checks done after fieldwork which was important in short period of time dedicated to data processing.

All tasks related to data processing were completed by central data processing team.

Using a single data collection platform had great benefit since data from all countries was available in the same format, so there was no requirement to recode and merge data files from individual sources. This reduced back-end data processing time.

Data checking procedures included conducting detailed checks on different stages of the project lifecycle. Data processing workflow is shown in Figure below.

Figure 1: Data processing workflow



There are several key steps in the data scripting and processing workflow:

- Based on the signed off source questionnaire, the DP Manager defined the data format and structure (variable names and variable labels, values and value labels, permitted values, routings, etc.). In addition to this, she set up instructions for scripting, ensuring that the questionnaire specification is clear/all routing unambiguously specified. Based on this, the DP team scripted the questionnaire in SM-S.
- One of the most important steps to ensure data consistency and the proper structure of data sets is to check the scripts before data collection starts. This was conducted both by the CCT as well as by local teams for all the other languages after translations to local languages are imported into the main script. Master script was signed off by the national project managers once error free.
- Once the master script was signed off (and so fieldwork can start), the Central Data Processing Team set up the data checking process by designing an SPSS syntax for checking the data format (variable names and labels, values and value labels, permitted values), structure (check of all routing and contingencies) and quality.
- The first test on real data was performed after 10% of the data has been collected. This stage was very important for two reasons: it is an opportunity to spot and amend any issues that may have been uncovered with the data / questionnaires at an early stage before most the fieldwork has been completed, and to validate the checking approach.

- If the final datasets passed all checks, the data was regarded as clean and was signed off by the national project managers.
- All clean and signed off local data sets was delivered to UNDP.

Set up of data checking process

The Ipsos Central Coordination Team prepared SPSS syntax to run all the checks and produce data checking reports. The checks were prepared once the master questionnaire was signed off.

Based on this protocol, the Central Data Processing Team in close cooperation with DP Manager set up a macro (syntax) that was used to run a series of data checks. Where a check indicates an interviewer-related issue the report was sent to the applicable partner agency to investigate/feed back to the interviewer.

Ongoing data validation and consistency checks

The CCT was able to download the questionnaire whenever required. Data validation and consistency checks were built around three pillars:

1. Routing - As with all data-checking, one of the most important priorities was to check the total base at each question, checking that it was correct in line with the base description. Ipsos produced automated syntax in place that checked dummy data from each mode script prior to fieldwork, and early mainstage data to ensure that any issues with the routing or validation were picked up and could be rectified before they affected the quality of the final dataset. Errors were flagged and appropriate measures were taken to correct the scripts.

2. Permitted values - Every question had a list or range of permitted values. This could include permitted responses from a code list (e.g. 1 – Male, 2 – Female) or a range (e.g. age is restricted to a maximum of 100), and includes ‘Don’t Know’, ‘Refused’ and ‘Not Applicable’ (where applicable). Every question was analyzed at the stages described above so that we were able to correct the script in advance of (or early in) the fieldwork.

3. Item-nonresponse - High levels of item non-response (don’t know or refused) are not desirable as they usually suggest a problem, either with the question or response wording, respondent disinterest or interviewer poor work. Therefore, analysis was undertaken:

- On the respondent level in order to check if any interviews should be removed due to high level of disinterest
- On questionnaire level in order to check if there were interviews with significantly higher proportion of non-responses than in average.

After performed data checking and cleaning, datafiles (separated by Modules) was delivered to UNDP in format of SPSS datafiles.

7. Lessons learned

7.1. Questionnaire revisions

Major lessons learned from the lengthy process of questionnaire revisions are that: i) multi-indicator survey as this one has to have balanced set of questions with different subsections treated in a same manner ; ii) ensuring comparability with previous surveys (2004 and 2011), as well as with other similar surveys are key to measuring progress within the country and between countries ; iii) any revisions introduced can have significant impact on the coherence and flow of questionnaire and have to be carefully and consistently managed .

Importantly, the content of the questionnaire, formulation of questions and choosing respondents for different modules need to take into account background of surveyed population. The latter was in particular the case with income set of questions. On one hand, it was important to learn in detail amounts and sources of income of marginalised Roma population in order to be able to compare their situation with the non-Roma neighbours and possibly with national averages. On the other hand, using complex EU-SILC methodology to make such comparisons would have contributed to disbalance among modules, it seemed overambitious in terms of collecting accurate response, it would have significantly prolonged duration of the response and resulted in overall refusal of participants to respond. These initial assumptions were confirmed during the pilot testing. Thus, a compressed version of income set of questions was used, yielding more comprehensive information about sources and levels of incomes of marginalised Roma and their non-Roma neighbours than in 2011 survey.

The introduction of a new set of income questions had impact not only on the flow coherence and balance of questionnaire, but it also meant that some other questions from 2011 survey had to be omitted in order keep the length of questionnaire within reasonable time of 60 minutes maximum. Thus, a number of questions, which were not associated with core indicators had to be removed.

Further, when constructing the questionnaire the population background and past experiences need to be taken into account. Marginalised Roma are not an 'average' population, rather they are often times targeted by different actors with different types of surveys. They may also be reluctant to disclose some information, or the information they provided can result in self-reflection and feedback, such as the following 'This questions seem to be showing me in the worst light possible'. Similar comments to some of the sections about economic position and living standards of households were received not only by marginalised Roma respondents, but also by their neighbours. This was handled by interviewers providing further explanation about the nature of the survey and particular questions. Nevertheless, this should be taken into account in future surveys, in analytical phase and during presentation of results.

7.2. Fieldwork and data processing

Due to late start of the project and delay in finalization of the questionnaire, final scripting, fieldwork and sampling preparation was done in November, which is not optimal time for fieldwork in general

given the conditions of settlements and dwelling units where Roma population live. Having this in mind, recommendations for future waves would be:

- Leaving adequate time for sample preparation – in most countries, official data was used. This data was not only underestimating number of Roma population in general, but is not updated as well. Having in mind that Roma population change a place of residence more often than general population, official data often doesn't allow obtaining accurate picture of Roma settlements. Having more time for sample preparation would allow obtaining and combining data from different 'unofficial' sources such as data that often exists in Roma NGOs.
- Foreseeing more time for translation and back-translation in order to avoid 'translation' issues especially in complex questionnaire such as one used in this survey
- Avoiding fieldwork in period November-February. Conducting fieldwork in late autumn/winter time has a lot of disadvantages:
 - Time spent daily on fieldwork location is significantly less comparing to the time in case fieldwork is done in spring/summer. This means that interviewers can perform interviewing fieldwork until 4 pm at the latest – given that Roma settlements often don't have street lighting, interviewers don't feel secure when conducting interviews during night.
 - At the same time, since significant number of interviews were conducted in front of dwelling unit, time spent on fieldwork location is considerably less due to the cold.
 - Given that dwelling unit is sometimes composed of one room only (or only one room is used during winter time due to the heating), it was hard to secure privacy needed for conducting interview with randomly selected respondent
 - Roma population due to bad dwelling conditions and lack of proper heating often change their place of residence during winter times and sometimes they are hard to reach in sampling areas (where is their main place of residence)
- Fieldwork time should not be less than 6 weeks. Only fieldwork that lasts at least 6 weeks allows proper implementation of all rules related to repeated visits

Due to the length, complex routing and the fact that two interviewers are completing interview in one household at the same time, the questionnaire was challenging to script.

During the data processing and after queries from World Bank on the data delivered, it became clear that some additional soft checks could have been added to the script to improve data quality. However, it should be noted that creating too many checks within script is not always recommended, having in mind that illogical answers sometimes can indicate that some questions are not clear to the target population or translation was not done in appropriate manner. Therefore, future waves will need to balance the inclusion of extra checks against further complicating the script, making it difficult and more time-intensive to programme and check (and increasing the chances of errors) and making it difficult for interviewers to administer (by asking them to frequently re-confirm respondents' answers).

In addition, future waves should allow proper time for all project phases. This year's timetable was very challenging, not only because of not enough time for fieldwork preparation, but time foreseen for fieldwork (for the reason explained above). All future waves should consider project start around March,

so preparatory phases could be properly completed by summer time when fieldwork could be completed more efficiently. It is recommended to maximize the time between the pilot and the mainstage fieldwork so that both the contractor and UNDP/World Bank have sufficient time to consider findings and recommendations from the pilot and make amendments for the main stage.

