SIAR
research and consulting

The Household Survey on Corruption and Social Assistance

TECHNICAL REPORT

Prepared for:

World Bank

Done by:

SIAR research&consulting

Kyrgyzstan, Bishkek

147 Moskovskaya str., Office 1

Tel: +996 312 976331

E-mail: office@siar-consult.com

December,2014

The goal of the overall project is to evaluate household's views and experiences with corruption practices, as well as attitudes towards social assistance and actual practices of heating.

The main objective of this consultancy service is to carry out a household survey to provide primary data.

METHOD: CAPI

For the successful implementation of the project and achievement of goals of the research, quantitative method was used.

- ✓ *Individual (face-to-face) interview*

Quantitative survey conducted in 7 oblasts including Bishkek and Osh cities of Kyrgyzstan. The total number of respondents was 1,080.

Object: respondents aged 16 and senior.

 Subject: opinions, judgments and assessments.

Pilot: 17 interviews (with providing results to the Client)

FIELD PERIOD:

Milestones	Description	Timeline
1.	Training for supervisors in Bishkek	05.11.2014
2.	Training for interviewers in regions	06.11.2014
3.	Data collection	11.11.2014 - 7.12.2014
4.	Quality control	14.11.2014-19.12.2014
5.	Coding and data entry	15.12.2014-18.12.2014
6.	Database cleaning	18.12.2014-22.12.2014
7.	Technical report	22.12.2014
8.	Final database	22.12.2014

SAMPLE DESIGN DESCRIPTION

1 Target Population and Survey Population

The primary target population is all regular (non-institutional) households in the Kyrgyz Republic. Other target populations are all population oldest 16 yearsliving in non-institutional households. The survey population is identical to the target population; the survey covers all areas within the national borders.Table

1 shows the number of households by region and urban/rural area while Table 2 shows distribution of population.

Table 1: Distribution of Total Number of Households by Region and Urban/Rural Area.

Region	Total		Urban		Rural	
	Number of H'holds	% of Total Number of H'holds	Number of H'holds	% Urban Households in Region	Number of H'holds	% Rural Households in Region
Batken	80 160	6.9	20 769	25.9	59 391	74.1
Djalal-Abad	187 782	16.1	51 622	27.5	136 160	72.5
Issyk-Kul	103 323	8.9	33 427	32.4	69 896	67.6
Naryn	52 022	4.5	8 401	16.1	43 621	83.9
Osh Province	188 718	16.2	15 152	8.0	173 566	92.0
Talas	44 693	3.8	8 375	18.7	36 318	81.3
Chui	206 892	17.8	48 370	23.4	158 522	76.6
Bishkek	240 828	20.7	240 828	100.0	-	-
Osh City	58 393	5.0	58 393	100.0	-	-
Kyrgyz Republic	**1 162 811**	**100,0**	**485 337**	**41,7**	**677 474**	**58.3**

Table 2: Distribution of Population by Region and Urban/Rural Area.

Region	Total population	Urban population	Rural population
Batken	423 288	91 586	331 702
Djalal-Abad	1 005 743	217 420	788 323
Issyk-Kul	436 144	117 691	318 453
Naryn	258 510	34 949	223 561
Osh Province	1 103 278	86 515	1 016 763
Talas	225 939	32 607	193 332
Chui	791 744	142 399	649 345
Bishkek	862 408	862 408	-
Osh City	255 858	255 858	-
Kyrgyz Republic	**5 362 912**	**1 841 433**	**3 521 479**

2 First stage Sampling Frame and Primary Sampling Units (PSUs)

The primary sampling frame will be the list of census enumeration areas (EA) from the Census 2009. There are in all 13 297 enumeration areas in the Republic. Many of the rural EAs are formed around settlements (villages) so the EA coincides with the settlement. Larger settlements contain two or more EAs.

For each EA there is information on total number of people (by sex) and total number of households .There is also administrative information on urban/rural classification, municipality, district and oblast as well as a sketch map of the area. The census maps are kept at the municipality offices. There have been no changes in boundaries between administrative units (municipality, district, and region) since the census, but 50 EAs have been reclassified from urban to rural. The frame has been updated accordingly.The Kyrgyz projectemploys a

stratified two stage design. The first stage of sampling entails sampling of areas. In the second stage a sample of households are selected in each selected area. The census enumeration areas serve as first stage sampling units – Primary Sampling Units (PSU). Within selected PSUs a sample of households is drawn..

The distribution of enumeration areas by region and urban/rural area is shown in table 3.

Table 3: Distribution of EAs by Region and Urban/Rural Area in the Sampling Frame

	Total	Urban	Rural
Region	Number of EAs	Number of EAs	Number of EAs
Batken	1 041	194	847
Djalal-Abad	2 539	444	2 095
Issyk-Kul	1 049	244	805
Naryn	668	85	583
Osh Province	2 854	174	2 680
Talas	554	72	482
Chui	2 014	301	1 713
Bishkek	1 972	1 972	-
Osh City	606	606	-
Kyrgyz Republic	13 297	4 092	9 205

There are on average 86 households per EA, see table 4. There is a clear urban/rural difference in the size of enumeration areas; the urban EAs are on average 60 % larger than the rural EAs (in terms of number of households).

Table 4: Average Number of Households per EA by Region and Urban/Rural Stratum

	Total	Urban	Rural
Region	Average no of HHs per EA	Average no of HHs per EA	Average no of HHs per EA
Batken	77	118	67
Djalal-Abad	73	118	64
Issyk-Kul	98	144	84
Naryn	77	106	72
Osh Province	66	88	64
Talas	80	117	74
Chui	103	161	92
Bishkek	116	116	-
Osh City	95	95	-
Kyrgyz Republic	86	117	73

The sizes of the EAs (in terms of number of households) vary. Table 5 shows the distribution of EAs by size. There are 420 small EAs with less than 30 households. Very small EAs, with less than 20 households, are too small to serve as PSUs. These small EAs can be combined with neighboring EAs prior to the sampling. Alternatively, it is possible to combine the small sample EAs with neighboring EAs to form PSUs after the first

stage selection of EAs, so that it would only be necessary to combine the small EAs that are selected. The number of small EAs in the sample should be small.

Table 5: Distribution of EAs by size (number of households)

Number of households in EA	Number of EAs	Percent
0-29	420	3.2
30-59	1 988	15.0
60-89	5 830	43.8
90-119	2 835	21.3
120-149	1 319	9.9
150-179	563	4.2
180-209	210	1.6
210-239	81	0.6
240-269	26	0.2
270-299	11	0.1
300-349	6	0.0
350 +	8	0.1
Total	**13 297**	**100.0**

3 Survey Domains; Stratification

The survey domains are the eight administrative regions plus BishkekCity. It was decided to stratify the sample on survey domain by urban/rural area and Bishkek City. There are therefore 3strata in all. It was discussed if a deeper stratification- further stratification within region - would give further gains in precision of the estimates. It was concluded that the gains would be small so no further stratification was done. There is, however, an *implicit* geographical stratification within each province. This is achieved by a geographical ordering of the PSUs and systematic sampling of PSUs in the first stage of sampling.

4 Sample Size and Allocation

4.1 Sample size

Thesamplesizefor«Изучения общественного мнения относительно вопроса коррупции и социальной помощи»was1080households. It was further decided to select 12households per PSU. With 12 households per PSU, there will be a need for 90PSUs in the sample.

Table 6: Proposed allocation of the sample of 90PSUs to strata

Region	Total	Urban	Rural
Batken	6	3	3
Djalal-Abad	14	7	7
Issyk-Kul	7	4	3
Naryn	3	1	2
Osh Region	9	2	7
Talas	3	1	2
Chui	10	4	6
Osh city	8	8	
Bishkek	30	30	
Kyrgyz Republic	90	60	30

Assuming that 12 households are selected in each PSU , the sample of households by stratum will be as shown in table 7.

Table 7: Proposed allocation of the sample of 1080 households to strata

Region	Total	Urban	Rural
Batken	72	36	36
Djalal-Abad	168	84	84
Issyk-Kul	84	48	36
Naryn	36	12	24
Osh Region	108	24	84
Talas	36	12	24
Chui	120	48	72
Osh city	96	96	
Bishkek	360	360	
Kyrgyz Republic	1080	720	360

5 Number of Sample Households per Cluster

The project had a cluster size of 12households. This is a rather small cluster size compared to what is used in many other surveys focusing on demographic indicators. With a cluster size of 10 households altogether 108 PSUs would be needed to achieve a sample of 1080 households. A larger cluster size than 10 would mean that fewer PSUs than 108are needed. Consequently, the field work costs per household would be lower. On the other hand, the standard errors would be larger. To find out the theoretically "optimum" cluster size – the cluster size that gives the best precision per unit of cost – detailed data on field work costs would be needed. These data are not at hand but some crude calculations could still be done using "guesstimates" on average time for transports between PSUs, time for listing of households in sample PSUs and time for interview per household.

Based on the calculations described above it was decided to increase the cluster size from 10to 12 households. This size may well differ from the theoretical optimum value but it is suitable from a practical field work point of view. The total sample size is1080 households. So, the number of PSUs will be 1080/12= 90.

6 Sampling of PSUs

The PSUs are selected by systematic PPS sampling (PPS= sampling with probabilities proportionate to size). The PSUs are ordered in geographical sequence within the stratum before the selection of PSUs is done. The size measures are the number of households in the PSU from the Census 2009.

It was discussed within the sampling group whether there was a need for adjusting the size measures of some PSUs. This should be done in areas where it is known that substantial changes have taken place since the census (e.g. new large scale housing projects or clearance of squatter areas). The conclusion of the discussion was that updating of size measures should not be done. The opinion was that there should be rather few cases of radical population changes and also, that it will be difficult/costly to obtain updated information at the PSU level.

The sampling of PSUs will be done in Excel by a standard procedure which is used for all household surveys.

7 Sampling of Households

For each selecsd PSU, the starting points will be defined by regional supervisors. Big cities are divided into several **territorial units**, these units placed in the program Excel, and then the program will randomly select starting points.

In the rural area starting point is also determined by supervisors. The interviewer reaches selected from the sample village and then interviewer describes the layout of the streets supervisor and administrative buildings (the number of streets, crossing the street, the number of administrative buildings). This all is entered in the program Excel and then the program will randomly select starting point for this village.

The interviewer has no right to select the starting point, or change it.

Starting from the given address/point, an interviewer will follow strict rules to select a household and a respondent within selected household. The random route method using the right-hand rule is used with the predefined interval of three to select the household (counting each third household, excluding the starting point). Never move on the left side! In the deadlock - Interviewers cover only the right side of the street. In apartment building interviewers begin to move from the top floor downin a clockwise direction, also adhere to step 3.

Each third household is considered as main household, where up to three contacts must be attemptedat different times of the day, days of the week, and the weekend within the fieldwork period to conduct a successful interview. In areas where the interviewer will not be able to return on a different day, the interviewer will make attempts with at least a two-hour gap between each attempt before substituting the household.

If the interviewer cannot obtain an interview at the main sample household, the interviewer selects the household to the immediate right of the main household as the first substitute. In the event that the attempt at the substitute household also fails, then the interviewer selects the house immediately to the left of the initial/main household as the second substitute. In the event that an interviewer fails to obtain an interview at all three households, the interviewer selects another main household continuing with the same interval and numbering sequence of questionnaires can be saved.

All the routes and visits are to be recorded in Contact sheets completed by interviewers for each starting point.

8 Estimation Procedures

Weighting Procedures

The weights are equal to the inverse of the probability of selection of the units (households). The weights should be included in each sample unit record in the data sets. **Types of Survey estimates**

All of the survey estimates are in the form of proportions (percentages) or other ratios.

Due to disproportionate distribution of the sample to the urban and rural areas, as well as because of the possible differences in responses, data analysis will require weighting for actual representative research at the national level and at the domain level. Since the sample is a two-step process, the sample weighting was based on the probability of the each sample step for each cluster. We use the following notation:

P_{1hi}:   probability sample for the first phase of ith cluster in stratum h

P_{2hi}:   probability sample of the second phase inside i^{th} cluster (households)

Let us assume that a_h– it's the number of PU in the selected stratum h, M_{hi}– it's the number of households according to the total sample in i^{th} PU, и $\sum M_{hi}$ - the total number of households in the stratum. The probability of choosing ith-PU is calculated as follows:

$$\frac{a_h\, M_{hi}}{\sum M_{hi}}$$

Let us assume that b_{hi} - is the proportion of households in the selected clusters relative to the total number of households in PUi in stratum h, in case if the PU is segmented, or in other words $b_{hi}=1$.

So the probability of selecting a cluster i in the sample will be:

$$P_{1hi} = \frac{a_h\, M_{hi}}{\sum M_{hi}} \times b_{hi}$$

Let us assume that L_{hi} - - it's the number of households listed in operation list of households in the cluster istratuma h and g_{hi} - it's the number of households selected in the cluster. The second step of calculating the probability of selection for each household will be done in the following way:

$$P_{2hi} = \frac{g_{hi}}{L_{hi}}$$

The overall probability of selection for each household in the cluster istratum h therefore will be a derivative of the probability of selection in two stages:

$$P_{hi} = P_{1hi} \times P_{2hi}$$

Estimated weighting for each household within the cluster istratum h will be the opposite to the general probability of selection:

$$W_{hi} = 1/P_{hi}$$

RESPONDENT SELECTION

Interviewers surveyed the heads of households.

Head of household - is a household member who manages the household finances.

There are cases when for example the head of household is considered the father, but earn and manage finances son, in this case, the interviewer questioned his son.

CONTACT SHEET:

Region_____ Settlement_____ SSU _____

Interviewer _____

Adress/ Phone	Name		Date	Result	QN	Notes
		Visit I				
		Visit II				
		Visit III				
		Visit I				
		Visit II				
		Visit III				

		Visit I				
		Visit II				
		Visit III				
		Visit I				
		Visit II				
		Visit III				
		Visit I				
		Visit II				
		Visit III				
		Visit I				
		Visit II				
		Visit III				

TRAINING

During the survey period the interviewer is the main executor of works, and provides qualitative level of research. Supervisor's role is also very important. He coordinates the work of interviewers, provides control over conducting the survey by checking filled questionnaires and route sheets.

In total **59** interviewers were involved.

The training was conducted in three phases, the first was devoted to the theoretical basis of the survey, the second phase was given to training on the use of tablets and the third stage was aimed to the practical exercises.

1 phase:

Explanation of the basic techniques for the research and discussion of organizational questions

Explanation of common standards and methodologies of "F2F" interviews

Explanation of rules for choosing required respondent (head of H / H)

Explanation of the rules for filling the route sheet

Explanation of step along the route

2 phase:

Explanation of program on tablets

Examining the questionnaire on tablets

3 phase:

Survey (interviewers were divided into pairs and had the opportunity to interview each other).

FOLLOWING MATERIALS HAVE BEEN USED ON THE TRAINING:

- Tablets with questionnaires (russian, kyrgyz langusges)
- Cards
- Instruction for interviewers
- Routeing sheets
- Code pages for route lists
- Cover letter
- ID for each interviewer
- ID for each PSU

DIFFICULTIES DURING THE FIELD WORK:

In general, field work was carried out successfully and in time.

The most common problem in all oblasts was annoyance of the respondents regarding the presence of cattle, enumeration of property and enumeration of family members, even babies.

For example, in Batken oblast there were cases when respondents were trying to hide the existing property and cattle, even though interviewers saw standing in the courtyard a car or cattle. In this case, the interviewer had to ask again and it was very inconvenient. In Djalal-Abad interviewers said that many respondents were afraid to call the number of cattle due to the fact that they are obliged to pay tax for the cattle.

In Chui and Issyk-Kul oblasts and in Bishkek, many respondents did not want to give their contact details, fearing for their safety.

Many respondents did not understand why this study requires such many details, does it reflect positively on the respondents and will it work at all. Overall, there was distrust among the population.

Interviewers from Chui and Issyk-Kul Oblasts, as well as the Bishkek city faced with a large number of refusals. Moreover, many respondents were very upset with the amount of questions.

In view of the fact that the field work was held at the end of autumn, the weather conditions left much to be desired. There were complaints from interviewers from different oblasts that it is very cold and it was very difficult to work in a short time due to the severe weather conditions. For example, interviewers from Talas oblast said that it is very difficult to work with tablets when it rains. Furthermore, many interviewers from Osh, Batken and Naryn noted that during heavy snowfall it is impossible to go to the remote areas.

Such oblasts as Batken and Naryn noted a poor internet connection. Many were unable to send the completed questionnaire to the server. The tablets were sent to Bishkek (in SIAR company) and employees uploaded them to the server.

Desposition codes

A	B	C	D	E	F	G
1080	17	284	139	7	15	1

	rate	formula used
RR1	0,700	A/(A+B+C+D+E+F+G)

A	Completed Interview
B	Broke off interview
C	Refusal
D	NON-CONTACT*
E	OTHER
F	UNKNOWN
G	NOT ELIGIBLE

SCRIPT/DATA TRANSFER:

• The program for the data collection was executed by high- qualified specialist

• There are 60 available tablets in SIAR research&consulting company.

• Daily, every interviewer sent completed questionnaires to the SIAR server, then data was verified in the database and open-ended responces were coded

• Database was provided to customers in 3 phasese: 10%; 50%; and 100%.

• Geolocation determined at the level of one village (portfolio), because in regions we have very poor internet connection.

QUALITY CONTROL:

Quality control was done on three phases:

1-phase. Control by accompanied the interviewers (5%). This phase of control was performed by supervisor, in order to see :

- Right step upon the route

- Understanding of program on tablet

- Understanding of questionnaire

2-phase. Control by phone (5-10%). This type of control was carried out by independent controllers which were hired by SIAR. Supervisors were refined following questions:

• Have the interviewer came and interviewed you?

• Check the duration of the interview.

• In order to check answers in the questionnaire, respondents were asked several questions that we believe the respondent did not have to change the point of view

• Controllers asked questions about the number of residing in the selected household

• Controllers asked respondents to evaluate the performance of the interviewer on the 5-point scale, how he was professional, polite, etc.

• If the controller does not reach the respondent, the 3 phase of control is used

3-phase. Control by attending households (10.5%). This type of control was also carried out by independent controllers which were hired by SIAR.

This type of control includes all the above steps for 2-phase control.

The findings of the three-stage quality control allow us to conclude a good faith and quite high quality of primary data collection. With the difficulties during the interview (the correct route and respondent selection) faced a few. The presence of clear control mechanisms has allowed to identify and to eliminate the following violations: 31 cases of wrong route and 2 cases of wrong choice of the respondent. In both cases, SIAR research & consulting sent other interviewers, and the work was carried out again.

EVENTS:

1. Basic socio-economic indicators of living standards
   http://stat.kg/images/stories/docs/Yearbook/Living/LIVING%20STANDARDS%2017.1.pdf

2. Kyrgyzstan joined the list of countries in the World Bank with an average level of per capita income
   http://kabar.kg/rus/economics/full/79506

3. Education in Kyrgyzstan has become a zone of payments
   http://www.vesti.kg/index.php?option=com_k2&view=item&id=24687:sfera-obrazovaniya-v-kyirgyizstane-prevratilas-v-zonu-platezhey&Itemid=80#ixzz3MDtcObXH
   http://www.vesti.kg/index.php?option=com_k2&view=item&id=24687:sfera-obrazovaniya-v-kyirgyizstane-prevratilas-v-zonu-platezhey&Itemid=80

4. The employment rate of the population of Kyrgyzstan is 60%

   http://www.knews.kg/society/44030_uroven_zanyatosti_naseleniya_kyirgyizstana_sostavlyaet_60_/

5. The role of informal employment in Kyrgyzstan
   http://nisi.kg/ru-analytics-478

6. The unemployment rate in Kyrgyzstan on January 1, 2014 was 8.4%
   http://catoday.org/centrasia/kgru/12462-uroven-bezraboticy-v-kyrgyzstane-na-1-yanvarya-2014-goda-sostavil-84.html

7. "Bishkekteploset": The new tariffs are designed to protect vulnerable citizens
   http://www.vb.kg/285980

8. Coal prices in Kyrgyzstan in the past month increased by 20%
   http://rus.azattyk.org/content/news/26622861.html

9. Deputies discuss increased penalties for corruption
   http://www.vb.kg/297095

10. Kyrgyzstan will count the number of livestock and poultry
    http://www.vb.kg/295929

11. Kyrgyzstan will implement the project for the identification of domestic animals
    http://rus.azattyk.org/content/news/25307068.html

12. Corruption in Kyrgyzstan today
    http://www.knews.kg/tags/korruptsiya/

13. K.Baybolov (deputy) - "Corruption in Kyrgyzstan. The law does not change anything but we should fight"
    http://www.centrasia.ru/newsA.php?st=1049958300