



## Improving survey methods in Kazakhstan

### Improving the sampling methodology of household living standard survey

**Guljan Daurenbekova**

Group of Experts on Measuring Poverty and Inequality  
Geneva (Switzerland), 7-9 December 2022

## Sample Design

---

**Sample frame** - housing stock register

**Sampling method** - two-stage probabilistic (random) sampling using stratification and random selection procedures at each stage

**Assessment of the accuracy of indicators** - the sample size (12,000 households) provides results with a sampling error of no more than 1% at the national level and no more than 4% at the regional level

**Selection principle** - disproportionate stratification selection

**Sampling unit** - household

## Sampling stages

---

### Phase 1

- Stratification by geographic region (31 strata)
- Selection of primary sampling units (400 PSUs)

### Phase 2

- Segmentation (grouping settlements by geographic location)
- Selection of households in each primary sampling unit

The sample design follows the principle of the optimal combination of costs and specified criteria for the accuracy of the results

## Annual rotation

---

**The purpose of rotation** - avoiding the effect of household fatigue from participating in the survey

**Rotation period** - every year

**Amount of rotation** - 1/3 part

## Rotation signs

---

**All households in the sample have a rotation sign and are removed from the sample annually – substituted with new households**

**1 year** - removal of households with the rotation sign "1" from the list

**2 year** - removal of households with the rotation sign "2" from the list

**3 year** - removal of households with the rotation sign "3" from the list

The annual rotation is the most difficult period in the work of the household survey departments of the Oblast Departments of Statistics and interviewers.

### **Main reasons:**

- It is difficult to persuade people to participate in the survey,
- documents from households are required,
- first people agree and then refuse,
- no substitutions for the sampled household are provided for
- the need to use only backup lists when recruiting the network
- the pandemic and social tension, distrust of the population

## Current situation evaluation

---

- **The current version of the sample design** was developed by expert Juan Munoz as part of the joint economic study for 2010 "Development of a new sample for a household survey on living standards" of the World Bank's Living Standards Measurement Study project
- **Quality reports:** since 2019, tables with sampling errors and quality reports with detailed calculations have been published
- Sampling **error calculations** based on the results of the household budget survey are carried out by the division of statistical registers and classifications

# Current situation evaluation

## Accuracy of 2021 Living Standards Sample Survey Results

	Survey indicators (average per month per person, tenge)	Standard error	Relative standard error	Limit of error
				per cent
Income used for consumption	69,038	192.4	0.29	377.1
Monetary expenditure	67,440	190.6	0.30	373.5
Monetary income	69,111	190.4	0.28	473.1
Consumption expenditure	62,711	170.8	0.29	334.8

	Income used for consumption	Monetary expenditure	Monetary income	Consumption expenditure
<b>Қазақстан Республикасы</b>	0.29	0.30	0.28	0.29
Ақмола	1.08	1.13	1.12	1.09
Ақтөбе	1.01	1.08	0.89	1.08
Алматы	1.08	1.09	1.13	1.03
Атырау	0.94	0.93	1.09	0.84
Батыс Қазақстан	0.97	1.07	1.13	1.04
Жамбыл	0.88	0.89	0.94	0.90
Қарағанды	1.13	1.16	1.06	1.12
Қостанай	1.06	1.12	1.02	1.05
Қызылорда	0.86	0.87	0.83	0.86
Маңғыстау	0.75	0.75	1.17	0.73
Павлодар	1.12	1.15	1.03	1.13
Солтүстік Қазақстан	1.34	1.39	1.24	1.35
Түркістан	0.71	0.72	0.79	0.70
Шығыс Қазақстан	1.25	1.30	1.12	1.26
Астана қаласы	1.22	1.22	1.14	1.16
Алматы қаласы	1.37	1.37	1.25	1.35
Шымкент қаласы	0.81	1.81	0.96	0.80

## Problem issues:

---

- Not agreeing to participate in the survey when recruiting the network, households withdrawal after the first year of participation in the survey
- Sample bias as high-income households do not agree to participate
- Low pay for interviewers and turnover
- Non-observance of rotation rules - participation of households is longer than the specified period (in practice, the network is updated by more than 30%)
- Questions on the participation in the survey of recipients of social benefits
- Increased negative attitudes and social distrust of the population, especially after the COVID-19 pandemic
- Substitution of households in the sample with others
- Formulation and approval of the sampling methodology for the standard of living survey

## Improving survey methods

---

The **Department** adhered to international guidelines on conducting a survey during the COVID-19 pandemic

- Crowdsourcing (use of volunteers as part of the household survey for cities of republican significance)
- Telephone surveys

The Department also carries out the following activities:

- Comparing the data with data from register systems and demographic statistics to ensure that the sample represents the general population (education attainment, employment status, gender, age, etc.)
- Offsite events are conducted by visiting households in person
- Methodological work with the UNECE experts to improve the sample is performed

## Improving survey methods

---

**As part of cooperation with UNECE** on poverty statistics, work is underway to update the methodology for sampling households by standard of living with the participation of UNECE expert on sampling survey methods Gianni Betty

The **goal** is to improve the quality of survey data in the face of increasing non-response, progress in data disaggregation, design optimization and accuracy improvement

### **Expert assessment:**

“The HBS sampling methodology currently used by the BNS is generally consistent with the universally accepted sampling methods and ensures a high quality of statistical data”

The sampling methodology introduced in 2010 was reviewed by World Bank staff in 2017 and rated “still good”

Implementing the recommendations will have a positive impact both on reducing the so-called dropout rate, and on non-response rate due to respondent burden

Transforming theoretical and technical aspects into concrete actions



### Recommendations by expert Gianni Betti (implemented):

- **Recommendation no. 1:** It is recommended to adopt the new  $Deff = 1.24$  and  $Deff = 1.89$  for urban and rural areas respectively, for the calculation of the optimal sample sizes, and their allocation over regions in the future HBS sampling strategy.
- **Recommendation no. 3:** It is recommended to transform the current sample design for HBS into a 4-year rotating survey design similar to EU-SILC.
- **Recommendation no. 6:** Perform calibration or post-stratification.
- **Recommendation no. 8:** Introduce a system of data quality reporting to be prepared at regional level, and the aggregated at central office level.

# Improving survey methods

## Network redistribution

Region	For 2022			For 2023			increase/decrease		
	number of households						total	urban	rural
	total	urban	rural	total	urban	rural			
Abai	360	180	180	510	240	270	150	60	90
Akmola	840	360	480	750	360	390	-90	0	-90
Aktobe	780	360	420	720	360	360	-60	0	-60
Almaty	450	120	330	480	180	300	30	60	-30
Atyrau	510	300	210	510	300	210	0	0	0
West Kazakhstan	660	240	420	660	240	420	0	0	0
Zhambyl	630	270	360	630	270	360	0	0	0
Zhetysu	270	120	150	480	240	240	210	120	90
Karaganda	810	480	330	630	390	240	-180	-90	-90
Kostanay	810	360	450	780	390	390	-30	30	-60
Kyzylorda	600	240	360	600	240	360	0	0	0
Mangistau	540	330	210	570	330	240	30	0	30
Pavlodar	840	360	480	780	360	420	-60	0	-60
North Kazakhstan	660	270	390	660	270	390	0	0	0
Turkestan	600	210	390	600	210	390	0	0	0
Ulytau	150	120	30	390	210	180	240	90	150
East Kazakhstan	540	270	270	480	240	240	-60	-30	-30
City of Nur-Sultan	600	600	-	570	570		-30	-30	
City of Almaty	810	810	-	720	720		-90	-90	
City of Shymkent	540	540	-	480	480		-60	-60	
<b>Total</b>	<b>12,000</b>	<b>6,540</b>	<b>5,460</b>	<b>12,000</b>	<b>6,600</b>	<b>5,400</b>	<b>0</b>	<b>60</b>	<b>-60</b>

## Rotation options

### First option

KAZ variant					100% new sample			
	2023	2024	2025	2026	2027	2028	2029	2030
10%	I	new 1 group			I			
20%	II	II			II	II		
20%	III	III	III		III	III	III	
50%	IV	IV	IV	IV	IV	IV	IV	IV

### Second option - Rotation without rotation sign.

Households are observed without the rotation sign and annual withdrawal occurs voluntarily according to the conclusions that 25% households on average do not complete the 3-year cycle. I.e. Annual “voluntary” renewal will occur at 25% level. An additional condition is that households cannot be observed for more than 4 years.

### Third option - rotation similar to EU-SILC.

The rotation pattern is replaced with a 4-year cycle with annual renewal by 25%

	2023	2024	2025	2026	2027	2028	2029	2030
25%	1	2	3	4	1	2	...	...
25%	4	1	2	3	4	1	2	...
25%	3	4	1	2	3	4	1	2
25%	2	3	4	1	2	3	4	1

### Recommendations by expert Gianni Betti (being implemented):

- **Recommendation no. 2:** As an example, it is recommended to explore the Eurostat web survey on the similar “Distribution of population by number of years spent in poverty within a four-year period”
- **Recommendation no. 4:** Use larger areas for non-response adjustment. This will reduce the presence of extreme weights in some specific PSU and will reduce the current Kish effect (effect of weights), which determine larger variability in the estimates (standard errors).
- **Recommendation no. 5:** Use hot-deck data imputation methods to adjust for item non-response.
- **Recommendation no. 7:** Adjust calibrated weights to avoid significant increases in variance or standard errors in the final estimates.
- **Recommendation no. 9:** When disaggregating data to measure poverty, it is strongly recommended, where possible, to follow the target variables proposed in the UNECE Guidance (2020) .

---

**Thank you for your  
attention!**