



UNECE



**UNIVERSITÀ
DI SIENA**
1240

**ЕВРОПЕЙСКАЯ ЭКОНОМИЧЕСКАЯ КОМИССИЯ
ОРГАНИЗАЦИИ ОБЪЕДИНЕННЫХ НАЦИЙ
КОНФЕРЕНЦИЯ ЕВРОПЕЙСКИХ СТАТИСТИКОВ
Семинар по гармонизации статистики бедности для
измерения прогресса в достижении ЦУР 1 и 10**

Совершенствование методов обследования в Казахстане

Джанни Бетти

*Консультант ЕЭК ООН, Университет Сиены,
Италия, gianni.betti@unisi.it*

Женева, 7 декабря 2022 года



Оглавление

- **Введение**
- **Анализ методов формирования выборки, используемых в настоящее время Бюро национальной статистики Казахстана**
- **Решение вопроса качества данных: проблемы отсутствия ответа и ошибки охвата**
- **Разбивка данных о показателях бедности**
- **Опции для лонгитюдного обследования 2023-2030 гг.**
- **Основные выводы для других НСУ**
- **Литература**



UNIVERSITÀ
DI SIENA
1240

Введение

В презентации кратко излагается отчет, подготовленный в рамках консультаций, проводившихся по направлению «измерение бедности» проекта Счета развития Организации Объединенных Наций «Укрепление социальной защиты в ответ на пандемию». Его основной задачей является оказание технической помощи для выполнения рекомендаций, касающихся оценки и совершенствования методов обследования, изложенных в публикации ЕЭК ООН **«Измерение уровня бедности: Руководство по дезагрегации данных»** (ЕЭК ООН, 2020 г.), которое было основной темой семинара ЕЭК ООН в декабре 2019 года.



Анализ методов формирования выборки, используемых в настоящее время Бюро национальной статистики Казахстана

План выборки домохозяйств основан на многовероятностной (случайной) выборке. Этот принцип гарантирует независимость выбора единиц выборки (домохозяйств и отдельных лиц)

На **первом этапе** проводится стратификация генеральной совокупности по региональному распределению, в том числе на городские и сельские районы.

На **втором этапе** из каждой включенной в выборку ПЕВ отбирается определенное количество домохозяйств.



UNIVERSITÀ
DI SIENA
1240

Анализ методов формирования выборки, используемых в настоящее время Бюро национальной статистики Казахстана

Общий размер выборки рассчитывается на основе принципа оптимального сочетания затрат и заданных критериев точности результатов.

Такой общий (национальный) размер выборки определяется на основе стандартной ошибки (целевой) выборки и на основе дизайн-эффекта ($Deff$) для городских и сельских районов, которые были установлены, соответственно, на 1,0 и 2,0 .

Еще одним интересным аспектом плана выборки БНС для ОБДХ является лонгитюдное измерение.



Анализ методов формирования выборки, используемых в настоящее время Бюро национальной статистики Казахстана

Рекомендация 1. Рекомендуется принять новые значения $Deff = 1,24$ и $Deff = 1,89$ для городских и сельских районов, соответственно, для расчета оптимальных размеров выборки и их распределения по регионам в будущей стратегии выборки ОБДХ. Они были оценены на основе метода Верма и Бетти (2011).

Рекомендация 2. Предлагается более глубоко анализировать такой богатый лонгитюдный набор данных, изучая распределение лет, проведенных в бедности, т. е. процент / долю отдельных лиц и домохозяйств, которые находились в состоянии бедности на протяжении: 0 лет (никогда), 1 года или 2 лет (временная бедность) или всех 3 лет (постоянная бедность).

В качестве примера рекомендуется изучить онлайн-опрос Евростата об аналогичном «Распределении населения по количеству лет, на протяжении которых они находились в состоянии бедности в течение четырехлетнего периода» (https://ec.europa.eu/eurostat/web/products-datasets/-/ilc_li51).



Решение проблемы качества данных: проблемы отсутствия ответа и ошибки охвата

Отсутствие ответа

Первый эффективный способ для преодоления проблемы неполучения данных — тщательное документирование каждого случая. Так как в полевых условиях в Республике Казахстан всегда бывает сложно соблюдать принцип «без замещения», то обеспечивается список резервных домохозяйств в количестве одной трети от числа домохозяйств в каждом кластере (т.е. 10 домохозяйств в ПЕВ). По этим причинам в каждом кластере выбирается в общей сложности 40 жилищ: 30 в заданной выборке и 10 заменителей.



Решение проблемы качества данных: проблемы отсутствия ответа и ошибки охвата

Корректировка на случай отсутствия ответа проводится на уровне ПЕВ, где ПЕВ - это территории малого размера в Казахстане

Рекомендация 4: Использовать более крупные территории для корректировки на случай отсутствия ответа. Это уменьшит наличие экстремальных весов в некоторых конкретных ПЕВ и уменьшит текущий эффект Kish (1988) (эффект отклонения весов), который определяет более высокую изменчивость оценок (стандартные ошибки).

Рекомендация 6: Выполнить калибровку или постстратификацию для обнаружения ошибок охвата.



Решение проблемы качества данных: проблемы отсутствия ответа и ошибки охвата

Наконец, когда сбор данных осуществляется на региональном уровне, одной из серьезных проблем является качество данных, предоставляемых региональными отделениями.

Одним из хороших практических методов является введение обязательной подготовки отчетов о качестве данных на региональном уровне, а затем их включение в национальный отчет о качестве.

Рекомендация 8: Внедрить систему отчетности о качестве данных, которая будет готовиться на региональном уровне, а обобщаться на уровне центрального офиса.



Дезагрегация данных о показателях бедности

- С учетом обилия информации в руководстве ЕЭК ООН (2020) здесь мы просто предлагаем принять во внимание следующие переменные для дезагрегации для всех членов домохозяйства;

Пол (целевая группа: женщины и девочки);

Возраст (целевые группы: дети, молодежь и пожилые люди);

Инвалидность (целевая группа: люди с инвалидностью);

Миграционный статус (целевая группа: мигранты);

Этническая принадлежность (целевые группы: этнические меньшинства);



Дезагрегация данных о показателях бедности

Кроме того, можно рассматривать следующие переменные, относящиеся к социально-экономическому положению и географическим регионам:

Тип домохозяйства (характеристики состава домохозяйства);

Уровень образования (характеристики квалификации и социального положения);

Занятость (характеристики участия рабочей силы); **Статус владения домохозяйством** (характеристики организации проживания частного домохозяйства в жилищной единице);

Получение социальных пособий (характеристики состава дохода);

Степень урбанизации (характеристики, относящиеся к городскому/сельскому¹¹ составу).



Лонгитюдное обследование 2023-2030 гг - вариант 1

Вариант 1 - это разновидность СДУЖ-ЕС - ротационное панельное обследование продолжительностью 4 года, адаптированное к условиям Казахстана

	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



Лонгитюдное обследование 2023-2030 гг. - вариант 2

При следующем варианте 2 существует риск, что процент семей (выделено красным), которые могут наблюдаться более четырех лет, может достичь даже 50%.

Схематично выглядит следующим образом:

2023	2024	2025	2026	2027	2028	2029
1	1	1	1	1	1	1
2	2	2	2	2	2	2
3	3	3	3	3	3	3
4	4	4	4	4	4	4
5	5	5	5	5	5	5
6	6	6	6	6	6	6
7	7	7	7	7	7	7
8	8	8	8	8	8	8
9	9	9	9	9	9	9
10	10	10	10	10	10	10
11	11	11	11	11	11	11
12	12	12	12	12	12	12
percentage of updates:	25%	25%	25%	25%	50%	



Основные идеи для других НСУ- 1

- 1 - Документировать охват, точность и любые отклонения от международных стандартов в отчетах о качестве, которые должны составляться при каждой публикации данных.
- 2 - Строго придерживаться вероятностной выборки и воздерживаться от замены.
- 3 - Скорректировать план выборки и распределение для требуемой дезагрегации. Как минимум размер выборки для каждой группы не должен быть меньше 50 единиц. В идеале области оценки будут использоваться в качестве страт, чтобы можно было контролировать их размер.



Основные идеи для других НСУ- 2

4 - Обеспечить достаточное количество попыток контакта (например), использовать групповые переводы и адреса интервьюеров. Полевые инструкции подлежат дальнейшей доработке.

5 - Компенсировать отсутствие ответов и случайные отклонения путем корректировки распределений (калибровки). Убедиться, что эти корректировки (g-веса) остаются в разумных пределах. (например, от 1/3 до 3).

6 - Убедиться, что уязвимые лица участвуют в разработке исследования, защищены от любого вреда и могут извлечь выгоду из его результатов.



Основные идеи для других НСУ- 3

7 - Минимизировать количество недостающей информации. Поскольку статус бедности обычно выводится из нескольких переменных, которые должны наблюдаться одновременно, крайне важно, чтобы репрезентативность групп сохранялась с помощью адекватных и прозрачных правил импутации.

8 - Тщательно оценить дисперсию всех ключевых оценок. Публиковать коэффициенты вариаций в отчетах о качестве, чтобы точность можно было сравнить с другими опросами и на протяжении определенного периода времени. Как минимум, представить оценки дизайн-эффекта для уровня бедности.

9 - Установить правила фильтрации публикаций, исключая недостоверные оценки. Никогда не публиковать результаты, основанные на менее чем 20 наблюдениях или с соответственно высоким коэффициентом вариаций.



UNIVERSITÀ
DI SIENA
1240

Литература

Eurostat (2020), METHODOLOGICAL GUIDELINES AND DESCRIPTION OF EU-SILC TARGET VARIABLES, Luxembourg: Publications Office of the European Union.

Kish, L. (1988), Multipurpose sample designs. Survey Methodology, 14, 19-32.

UNECE (2020), Poverty Measurement: Guide to Data Disaggregation, Geneva.

Verma V., Betti G. (2011), Taylor linearization sampling errors and design effects for poverty measures and other complex statistics, Journal of Applied Statistics, 38(8), pp. 1549-1576.



UNIVERSITÀ
DI SIENA
1240

**СПАСИБО
ЗА ВНИМАНИЕ!!!**