

Показатели репрезентативности ответов на обследование

Семинар ЕЭК ООН по измерению бедности во время пандемии

Barry Schouten (Статистическое Управление Нидерландов и
Университет Утрехта)

26 марта 2021 года



Academy

Отказы от ответа

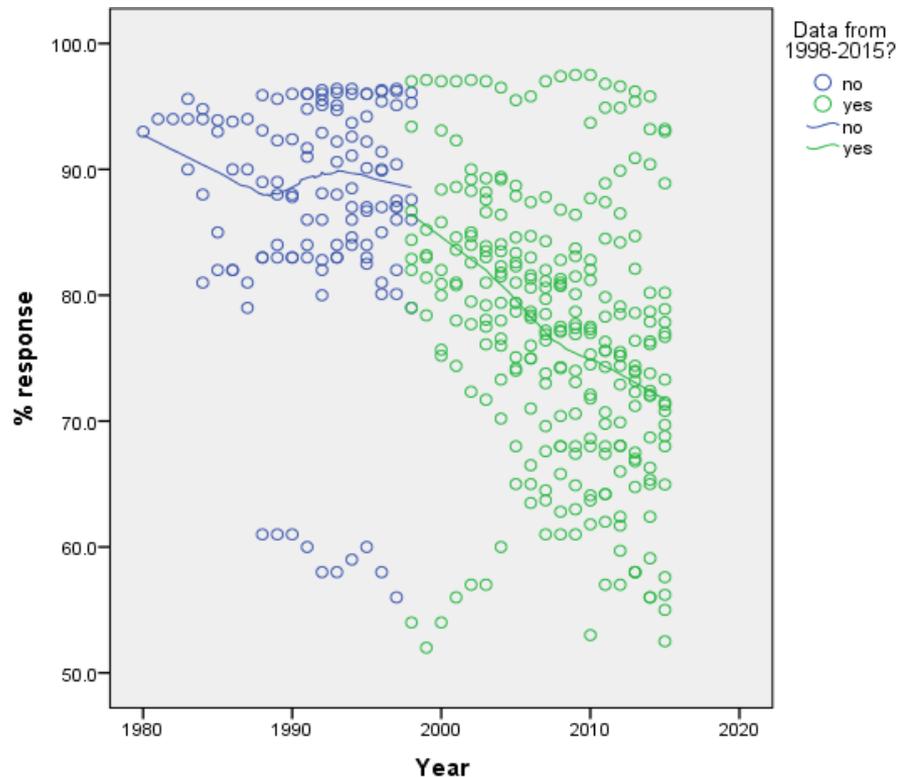
- Отказы от ответа случаются во всех обследованиях;
- Отказ от ответа является результатом недостаточно установленного контакта, отсутствия физических возможностей или языковой барьер, недостаток времени и нежелание;
- Отказ от ответов приводит к смещениям в расчетах;
- Отказ от ответа приводит к уменьшению выборки;
- Последний опыт показывает, что проблема отказа от ответа увеличивается, т.е. требуется все больше усилий/бюджетных средств, чтобы получить такой же уровень участия;

- Варианты:
- Предотвратить его появление (снижение);
- Скорректировать после с помощью статистических моделей (досчеты);
- Гибридный подход: скорректировать в процессе разработки дизайна (адаптивный дизайн обследования)



Тренды в уровне участия в обследованиях

Уровень участия в обследованиях рабочей силы 1980 – 2015гг в разных странах



Анализ отказов от ответа

Ключевой компонент это вспомогательная информация:

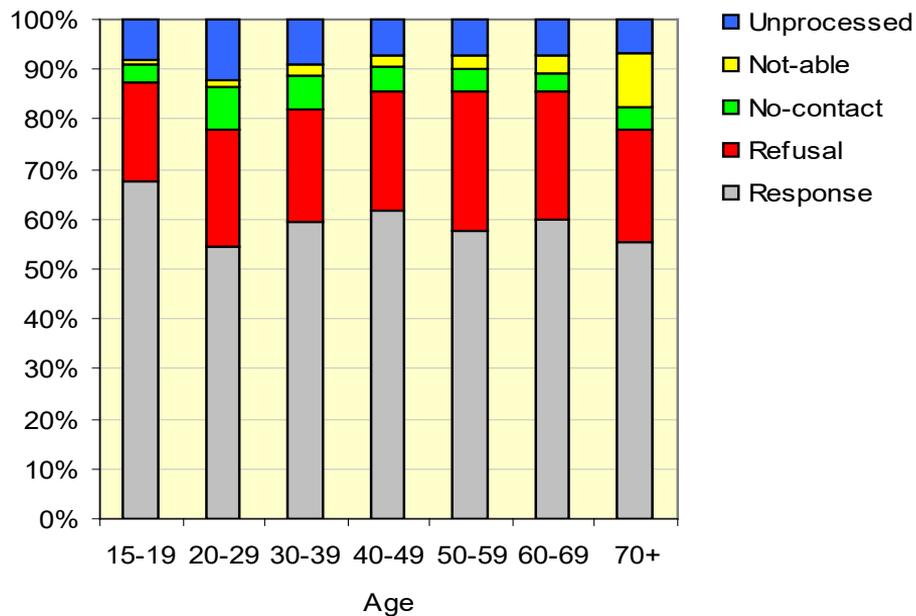
- Информация доступная в рамках выборки
- Информация из административных источников, которая может быть увязана
- Информация, которая доступна по общей численности населения
- Информация, которая собирается интервьюерами по всей выборке



Глобальный анализ отказа от ответа

Возраст

- Age

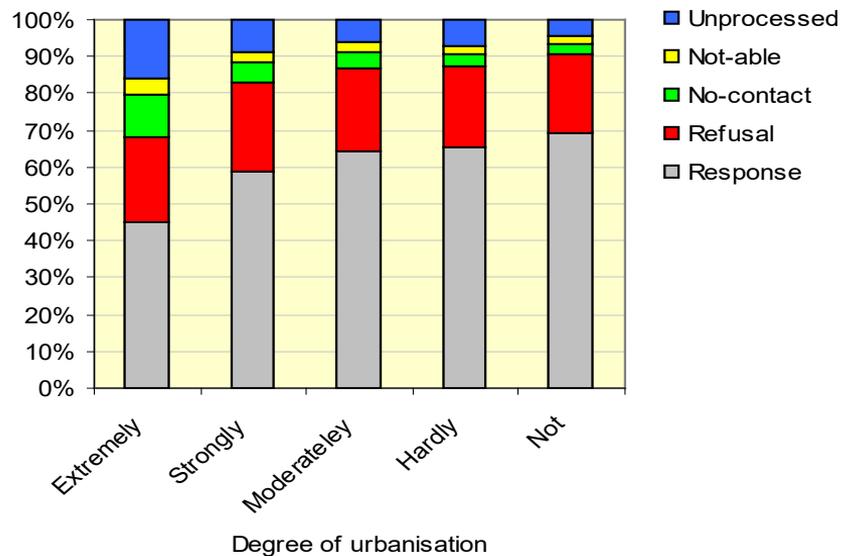


Academy

Глобальный анализ отказа от ответа

Уровень урбанизации

- Degree of urbanisation



R-индикаторы

- R-индикатор основывается на варьировании склонностей индивидуальных ответов

$$R(\rho) = 1 - 2S_\rho$$

- Два типа:
 - На основе выборки: ответ сравнивается с показателями по всей выборке
 - На основе населения: ответ сравнивается с показателями по всему населению
- На www.risq-project.eu вы найдете кодировку в SAS и R плюс учебные и тестовые наборы данных



R-индикаторы

- Смещения среднего показател участия ввиду отказа от ответа

$$|B(\bar{y}_R)| = \frac{|R_{\rho Y}| S_{\rho} S_Y}{\bar{\rho}} \leq \frac{S_{\rho} S_Y}{\bar{\rho}} = \frac{(1 - R(\rho)) S_Y}{2\bar{\rho}}$$

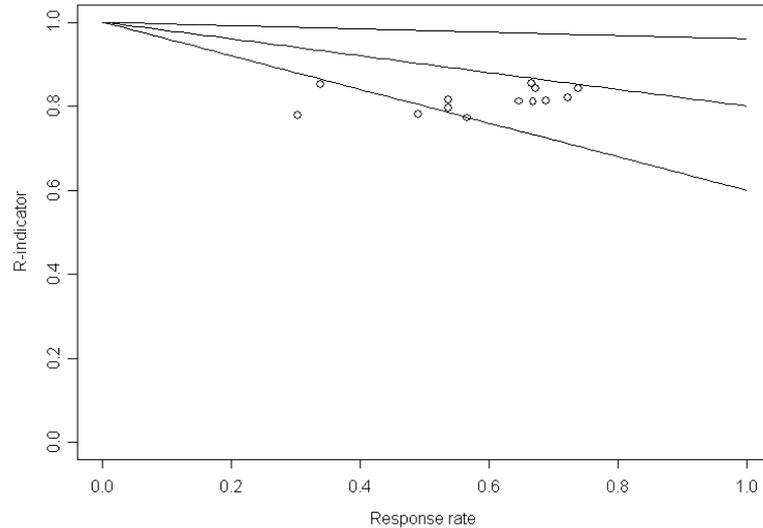
- Ограничивающий R-индикатор: диаграммы ответов-репрезентативности

$$\frac{|B(\bar{y}_R)|}{S_Y} < \gamma \quad \longrightarrow \quad R(\rho) > 1 - 2\bar{\rho}\gamma$$



R-индикаторы

- Примеры уровней участия и R-индикаторов (включая три кривые $\gamma = 2\%, 10\%, 20\%$)



Пример 1 – разные обследования ЕСС

X = пол, возраст, урбанизация

	Размер выборки	Уровень участия	R-индикатор
Обследование состояния здоровья 2005 (Голландия)	15,411	67.3%	0.832
ЕСС 2006 (Бельгия)	2,927	61.4%	0.807
ЕСС 2006 (Норвегия)	2,673	65.6%	0.762
Уровень жизни 2004 (Норвегия)	4,837	69.1%	0.872
ОПС, 3 квр – 2007 (Словения)	2,219	70.1%	0.854
ОПС, 4 квр – 2007(Словения)	2,215	69.3%	0.807



Пример 2 – Обследование серой экономики

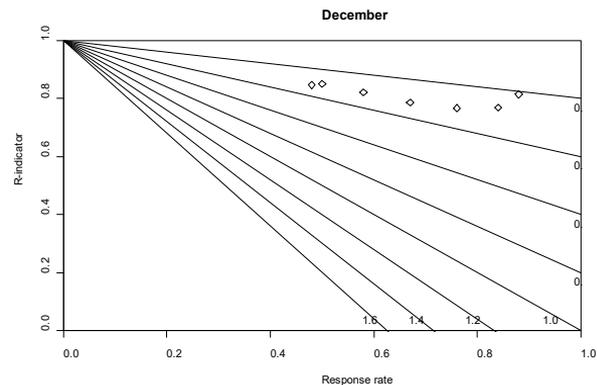
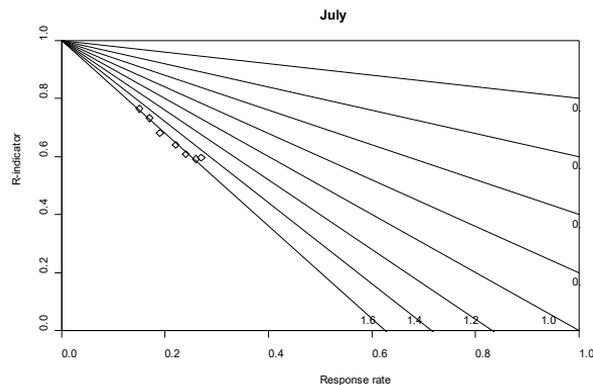
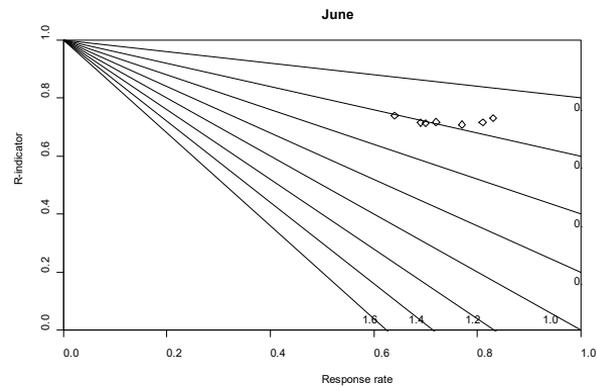
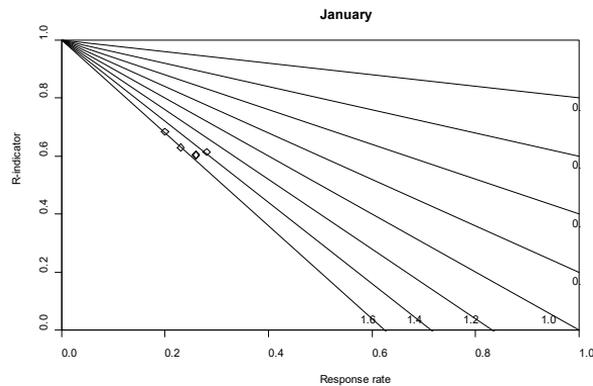
X = возраст, стоимость жилья, этническая принадлежность, тип домохозяйства, трудоустройство, городское население

Группа респондентов	Уровень участия	Показатель репрезентативности R	Доверительный интервал	CV
Встреча лицом к лицу	56.7%	77.8%	74.4% - 81.3%	0.102
В сети/на бумаге	33.9%	86.3%	83.1% - 89.4%	0.112
В сети/на бумаге + телефон	49.0%	79.3%	75.6% - 83.0%	0.11.3



Пример 3 – Обследование хозяйственной деятельности

$X = \text{зарплаты}(t), \text{КДЭС}, \text{НДС}(t-12) \times \text{размер}$

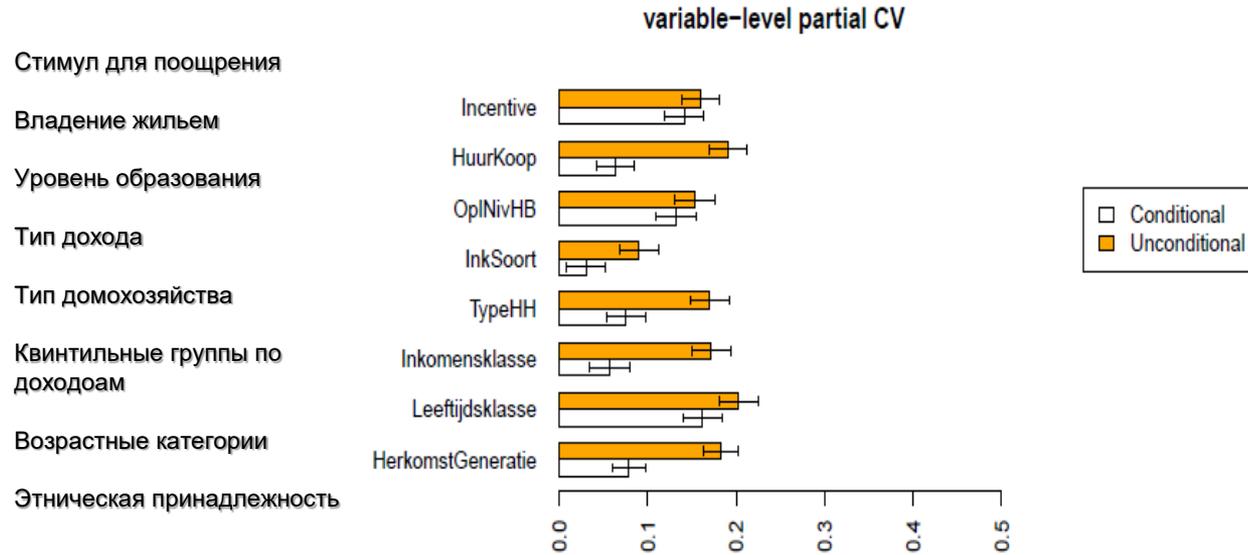


Частичные R-индикатор

- Частичные R-индикаторы раскладывают R-индикатор в зависимости от влияния единственных переменных
общая дисперсия = промежуточная дисперсия + внутренняя дисперсия
- Безусловный частичный R-индикатор для единственной переменной Z:
промежуточная дисперсия склонностей к ответу
- Условный частичный R-показатель для единственной переменной Z обусловленный X: *внутренняя дисперсия склонностей к ответу обусловленная стратификацией X*
- Оба типа индикаторов в идеале должны стремиться к 0 и позволять мониторинг сбора данных и распределение ресурсов

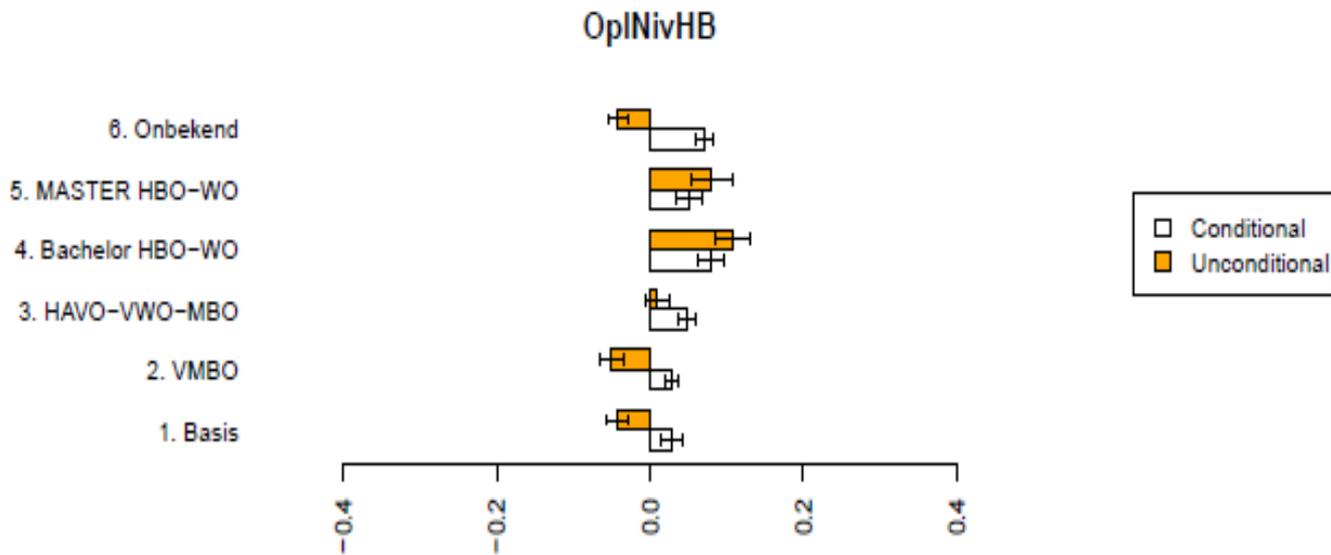


Пример EU-SILC



Пример – EU-SILC

Частичные R-индикаторы по уровню образования



Краткий итог

- Эти индикаторы были разработаны для своевременного мониторинга обследований, а также мониторинга в процессе сбора данных.
- Индикаторы могут использоваться для адаптации и разработки специализированных стратегий работы на местах.
- Ключевой компонент это вспомогательные переменные.

