

Коронавирус поражает и опросы: систематическая ошибка в отношении неполучения ответов и пандемия коронавируса

Джонатан Ротбаум

- US Census bureau

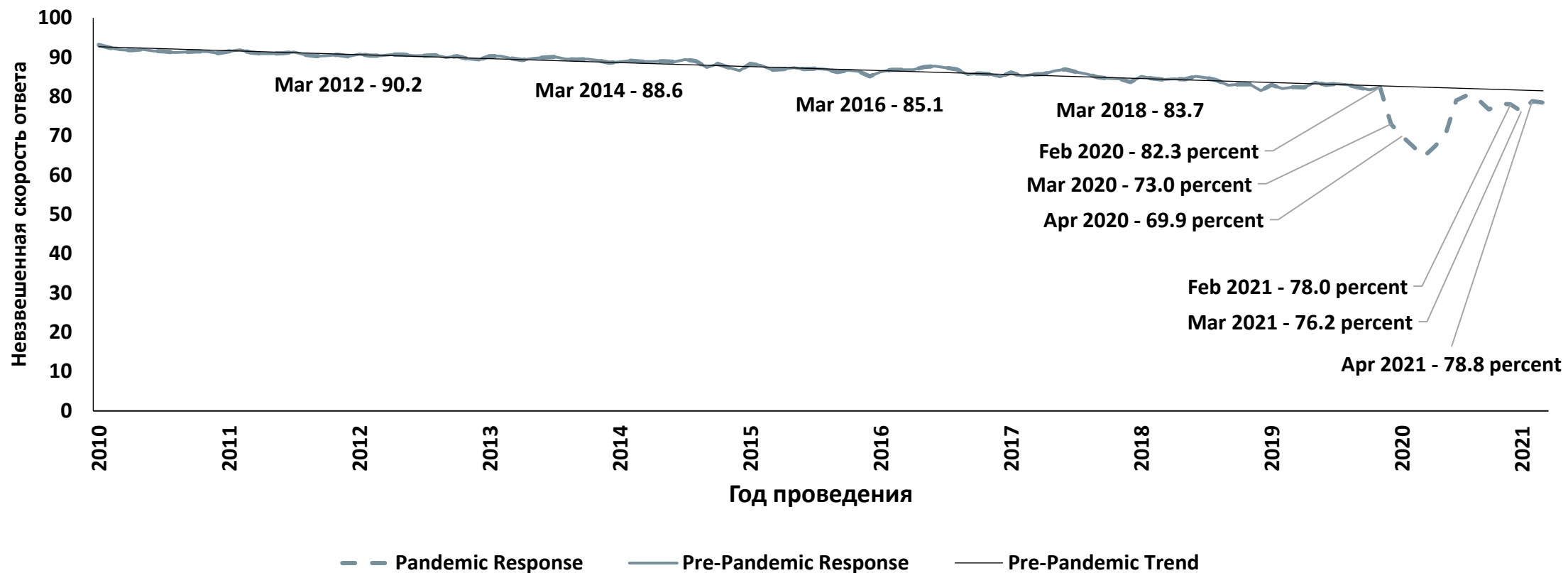
Адам Би

- US Census bureau

Семинар по гармонизации статистики бедности для измерения ЦУР 1 и 10
1 декабря 2021 г.

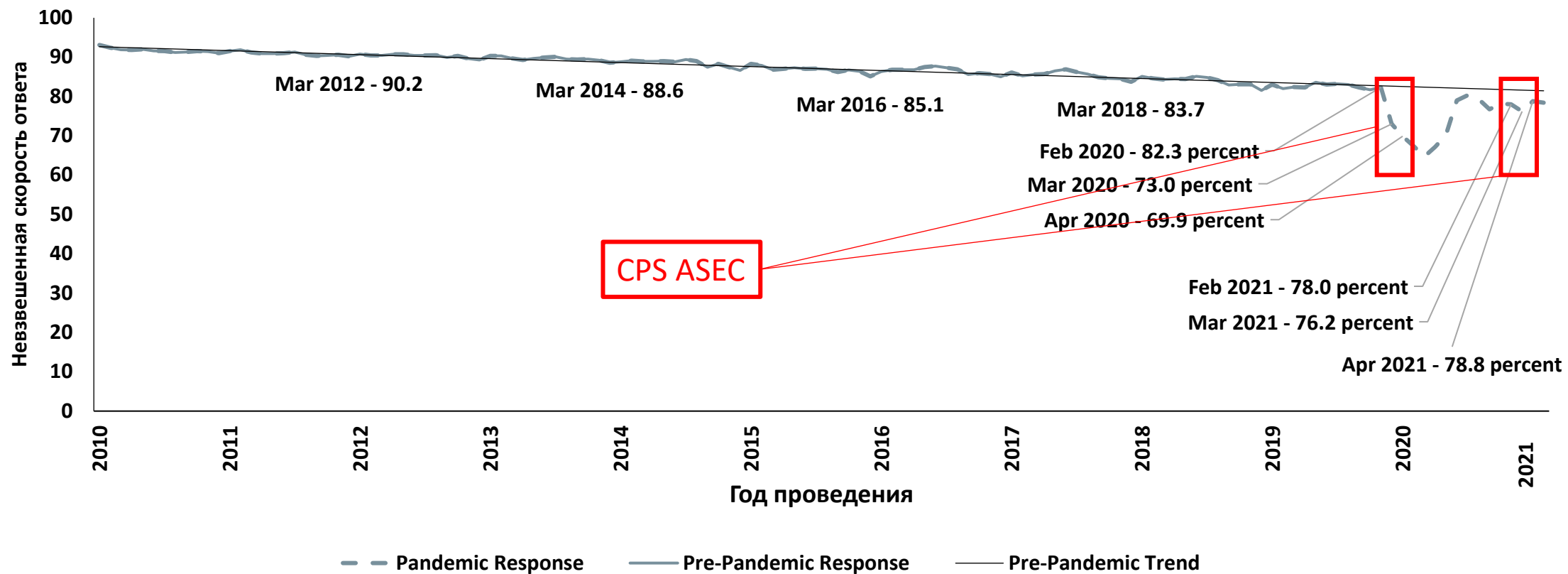
Эта презентация выпущена для информирования заинтересованных сторон о текущих исследованиях и поощрения обсуждения. Любые мнения, выраженные по статистическим, методологическим, техническим или эксплуатационным вопросам, принадлежат автору, а не обязательно Бюро переписи населения США. Бюро переписи населения США проверило этот информационный продукт на предмет несанкционированного раскрытия конфиденциальной информации и одобрило методы предотвращения раскрытия, применяемые к этому выпуску. CBDRB-FY20-380, CBDRB-FY20-414, CBDRB-FY21-POP001-0060 и CBDRB-FY21-POP001-0199. Гири для общего пользования одобрены для выпуска согласно CBDRB-FY21-126.

Ответ в текущем обследовании населения



Источник: Бюро статистики труда - <https://www.bls.gov/osmr/response-rates/household-survey-response-rates.htm>

Ответ в текущем обследовании населения



Источник: Бюро статистики труда - <https://www.bls.gov/osmr/response-rates/household-survey-response-rates.htm>

Результаты опроса, вызывающего озабоченность

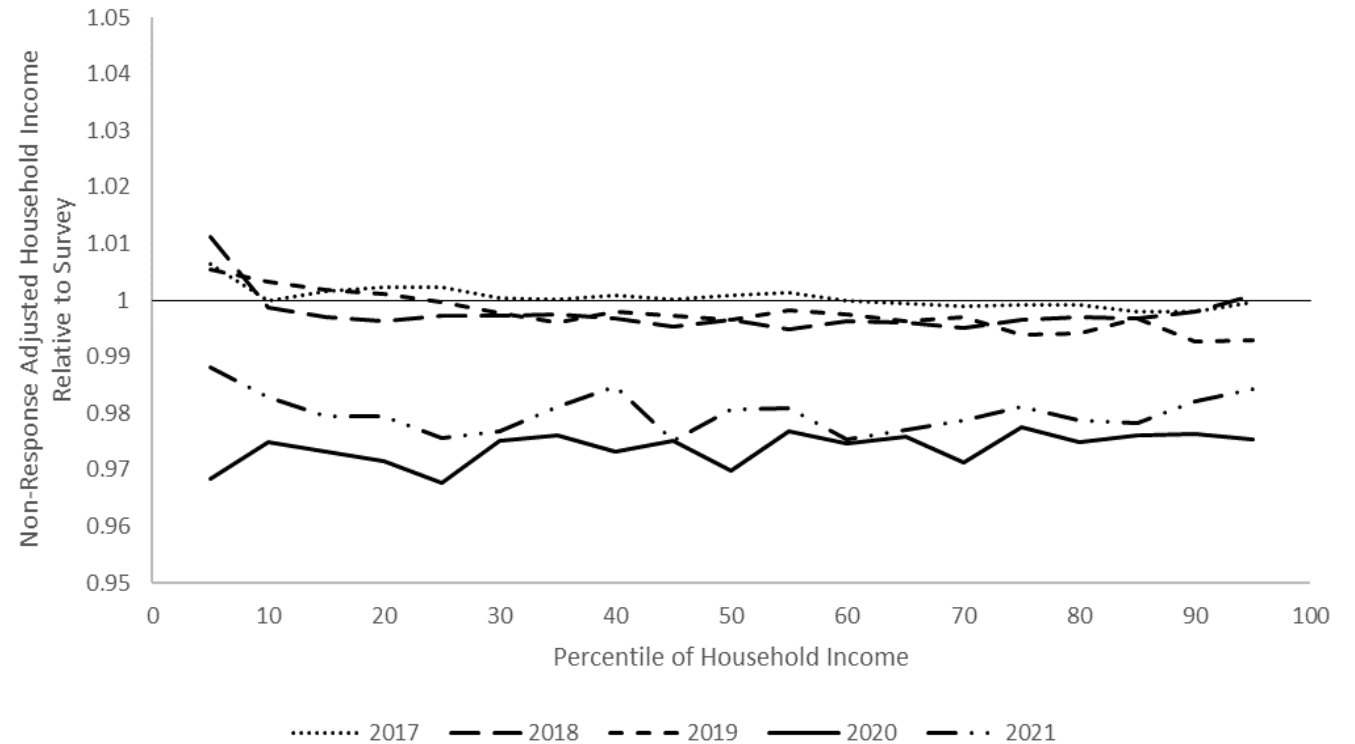
- Удивительно высокий рост медианного дохода домохозяйств в 2019 году
 - 6,8 процента - самая высокая оценка за всю историю
- Изменения характеристик, не учитываемых при взвешивании
 - Гражданство - меньше неграждан
 - Образование - больший, чем ожидалось, сдвиг в сторону более высокого образования

Что мы делаем

- Используем связанные административные данные, данные переписи и предыдущих обследований для оценки систематической ошибки, связанной с неполучением ответов.
- Применяем балансировку энтропии, чтобы повторно взвесить выборку, чтобы скорректировать систематическую ошибку неполучения ответов
 - Условие для связанных данных и внешнего контрольного населения
- Создаем веса для общего пользования, которые не зависят от административных данных

Предварительный просмотр результатов

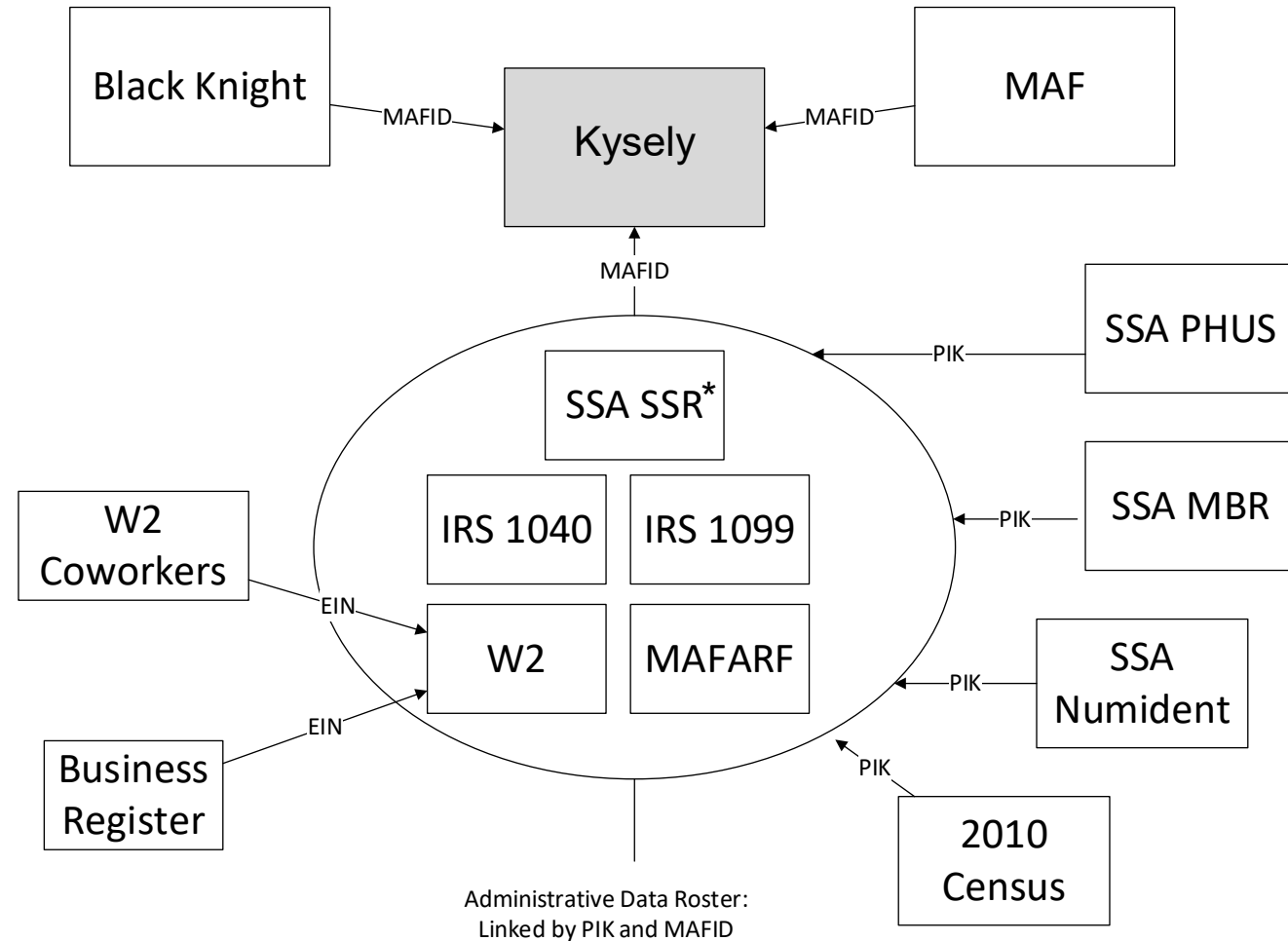
- Нет доказательств систематической ошибки неполучения ответов в отношении дохода или бедности в 2017-2019 гг.
- Согласно данным опроса 2020 и 2021 годов, доход завышен на 2-3 процента по всему распределению.



Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

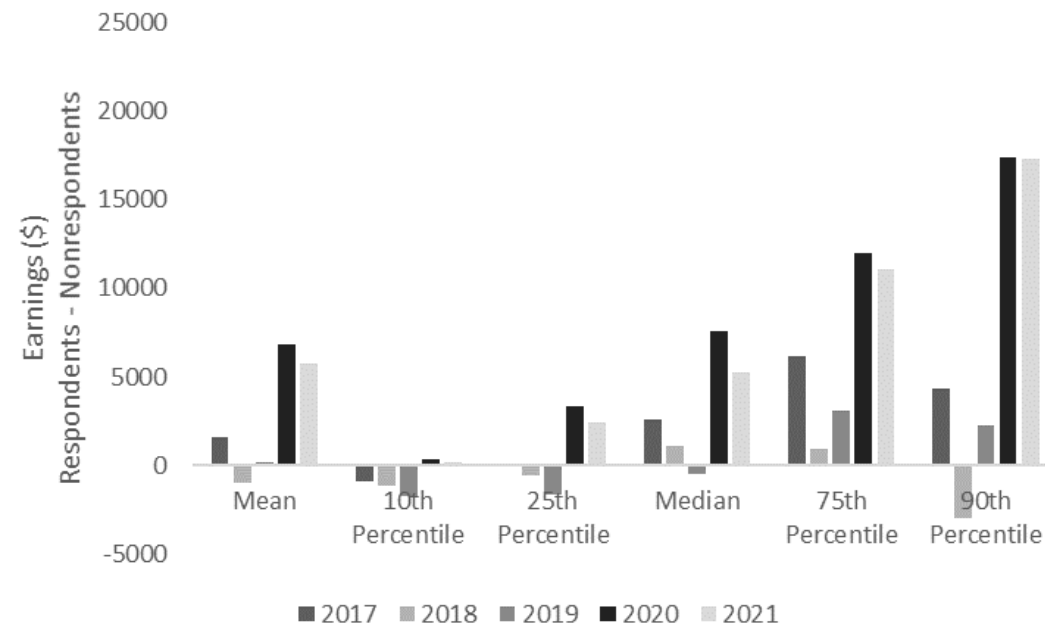
Дополнительная информация

- Увязка данных
 - Свяжите административные данные с адресами респондентов, *а также не ответивших*
- Что мы имеем
 - Жилье - MAF, Black Knight
 - Доход / занятость - W2, 1040, 1099s, PHUS, SSR, MBR, Business Register
 - Демография - Numident, перепись 2010 г. Включает расу, испаноязычное происхождение, гражданство, возраст, пол.
 - Состав домохозяйства / семейное положение - 1040
- Чего у нас нет
 - Прямые показатели образования, статуса медицинского страхования и прочей статистики
 - Любой информации в связанных данных для домохозяйств, которые не могут быть связаны MAFID



Неслучайное отсутствие ответа

Разница в доходах респондентов и неответивших



Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This figure shows the difference in income by address between respondents and nonrespondents.

Это shows total W-2 earnings at that address in the reference year of the survey.

Неслучайное отсутствие ответа

Разница в доходах респондентов и неответивших



Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Notes: This figure shows the difference in income by address between respondents and nonrespondents. Это shows total W-2 earnings at that address in the reference year of the survey.

Взвешивание в CPS ASEC

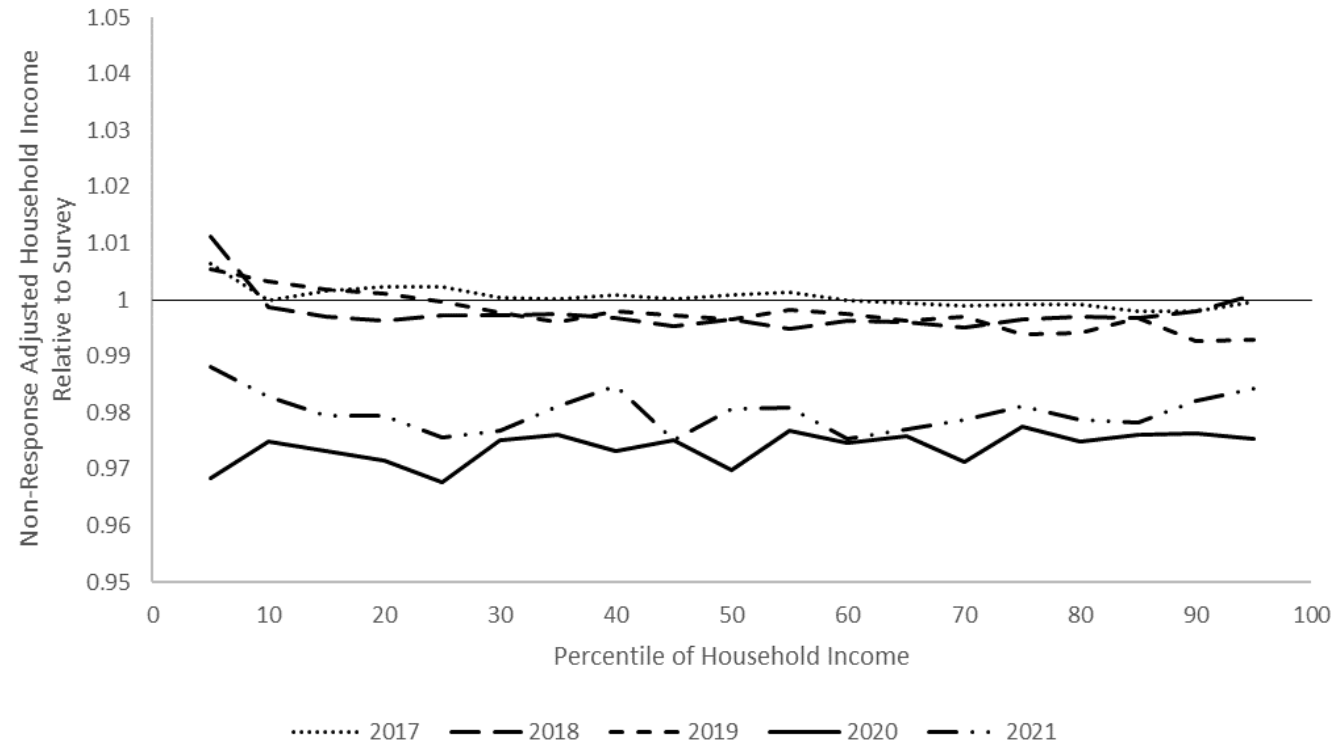
- Вес для контрольных групп населения (раса x возраст x пол x штат)
- Отношение дохода к ответу устойчиво для учета демографических данных в контрольных группах населения.

Балансировка энтропии

- Хайнмюллер, 2012. «Балансировка энтропии для причинных эффектов: метод многомерного повторного взвешивания для получения сбалансированных выборок в наблюдательных исследованиях»
 - Повторно взвесьте контрольную группу, чтобы она соответствовала характеристикам экспериментальной группы.
 - Применение экспоненциальной эмпирической калибровки (Девиль и Сэрндал, 1992)
 - Рассматривайте изменение веса как проблему оптимизации, а не начинайте с регрессии.

Результаты - доход

- Нет доказательств систематической ошибки неполучению ответов в отношении дохода или бедности в 2017-2019 гг.
- Согласно данным опроса 2020 и 2021 годов, доход завышен на 2-3 процента по всему распределению.



Source: U.S. Census Bureau 2017-2021 Current Population Annual Social and Economic Supplement linked to administrative, census, and survey data as indicated in Table 1. The 2017 and 2018 files are the CPS ASEC Research and Bridge Files, respectively.

Результаты - Бедность

- Бедность выше в оба года
 - На 0,46 процентных пункта выше в 2019 году (10,92 против 10,46)
 - На 0,44 выше в 2020 г. (11,88 против 11,44)

Веса для общественного пользования

- Проблема - веса были оценены путем кондиционирования на многомерном наборе административных характеристик.
 - Как показать, что они не раскрывали информацию о респондентах?
- Решение
 - Не делать
 - Создать новые веса, которые зависят только от информации в опросе
- Веса доступны на <https://www.census.gov/data/datasets/time-series/demo/income-poverty/data-extracts.html>

Планы на будущее

- Распространите эту работу на другие опросы, на которых работаете в настоящее время или планируете работать в будущем:
 - ACS
 - Обзор потребительских расходов (совместно с BLS)
 - Экспресс обследование домашних хозяйств
 - Десятилетняя перепись 1940 года (для использования со связанными данными для оценки долгосрочных результатов у детей)
- Использование балансировки энтропии для взвешивания связанных опросов и административных данных, чтобы быть репрезентативными для населения (проект NEWS)
- Пакет Stata, который охватывает файл ebalance.ado, для облегчения повторного взвешивания опроса с использованием балансировки энтропии с несколькими наборами моментных условий (в основном, завершенных)

Контактная информация

Джонатан Ротбаум

Экономист-исследователь

Отдел социальной, экономической и жилищной статистики

jonathan.l.rothbaum@census.gov

(301) 763-9681