

# Компьютерный практикум

- **Заработная плата и факторы ее формирования**
- **Измерение и регулирование неравенства**

Арбузова Дарья  
Дробот Мария  
Топорова Анастасия

# Цель и актуальность

## Цель

разработка алгоритма,  
позволяющего провести  
измерение неравенства и  
симуляцию эффектов  
перераспределительной  
политики

## 🔍 Проблема

рост напряженности в обществе  
вследствие социального неравенства

## 🔍 Актуальность

1

практическая значимость для  
государственного регулирования  
социального неравенства в обществе

2

научная новизна:  
симбиоз микро- и макро- подходов

# Задачи

## Обработка данных

- 1 Загрузка и предварительная обработка данных ОДН и ОРС
- 2 Выделение и гармонизация общих переменных
- 3 Импутация недостающих значений зарплаты в массиве ОРС

## Анализ

- 4 Расчёт показателей неравенства: коэффициент Джини и кривая Лоренца
- 5 Симуляций двух видов перераспределительных политик

## Литературный обзор

### **Chetty et al. (2016, Journal of Political Economy)**

Точечные безвозмездные трансферты для малообеспеченных домохозяйств в США.

Вывод: адресные выплаты приводят к сокращению коэффициента Джини на 0.5–1%, однако масштабное смягчение неравенства ограничено рамками программы. Требуется сочетать адресные меры с системными.

### **Kanbur et al. (2021, Journal of Development Economics)**

Исследование межрегиональных программ перераспределения в развивающихся странах: моделирование перевода средств из богатых регионов в бедные.

Вывод: значимое снижение неравенства достигается при таргетировании наиболее уязвимых групп внутри регионов.

### **Piketty & Saez (2014, American Economic Review)**

Исследование исторических данных Франции и США, показавшее, что прогрессивное налогообложение и трансферты имеют нелинейные эффекты: после определённого уровня ставки налоговая база инертна.

Вывод: оптимальный уровень налоговой нагрузки должен находиться в диапазоне, где сокращение неравенства максимизируется без значительного падения стимулов.

# Гипотезы

- 1** Точечные денежные трансферты снижают неравенство на макроуровне  
на основе работ Chetty et al., 2016 в Q1 Journal of Political Economy
- 2** Межрегиональное перераспределение эффективнее адресных выплат  
согласно исследованиям Kanbur et al., 2021 в Journal of Development Economics
- 3** Эффект перераспределения нелинеен и требует критической массы средств  
по данным Piketty & Saez, 2014 в American Economic Review

# Практическая реализация

Выгрузка архива из Яндекс Диска с помощью API

Наборы ОДН и ОРС содержат разные колонки:  
в ОДН — доход, в ОРС — демография и признаки трудовой активности

Общие переменные:

- пол
- регион проживания
- возраст
- уровень образования

## 🔍 Код

```
import requests, subprocess, pathlib, json, os, textwrap, sys

PUBLIC_LINK = "https://disk.yandex.ru/d/W2tG4krbjk4mw"
API = "https://cloud-api.yandex.net/v1/disk/public/resources/download"
resp = requests.get(API, params={"public_key": PUBLIC_LINK})
href = resp.json()["href"] # одноразовая прямая ссылка на ZIP

TARGET = "/content/rosstat_data.zip" # куда сохранить
subprocess.run(["wget", "-q", "-O", TARGET, href], check=True)

!unzip -q /content/rosstat_data.zip -d /content/rosstat_data

frames = []
for year in YEARS:
    for fp in (DATA_DIR / "ODN").rglob(f"*{year}*.sav"):
        frames.append(read_odn_one(fp, colmap))

COMMON_COLS = {
    "gender" : ["nas_pol", "NAS_POL"],
    "subject_code" : ["territ", "TERRIT", "SETKA", "setka"],
    "age_group_5y" : ["nas_voz4", "NAS_VOZ4"],
    "education" : ["nasobraz", "NASOBRAZ", "nas_obr1", "NAS_OBR1"],
}
```

# Практическая реализация

## Импутация зарплат в ОРС

Расчет расстояния до всех индивидов из ОДН по общим параметрам, используя расстояние Говера

→ N = 5 ближайших соседей

→ случайный выбор одного донора для копирования величины предналоговой зарплаты

Устранение проблемы отсутствующих значений и минимизация смещения

## 🔍 Формула и код

$$S_{ij} = \frac{\sum_{k=1}^p w_{ijk} s_{ijk}}{\sum_{k=1}^p w_{ijk}}$$

```
ors_df = add_aux_strata_cols(ors_df)

COMMON_VARS = ["gender", "subject_code", "age_group_5y", "education"]

ors_df = impute_wages_rowwise(
    donor_df = donor_small,
    recipient_df = ors_df,
    common_vars = COMMON_VARS,
    strata = STRATA,           # те же три страты
    N = 5,
    wage_column = "gross_wage",
    random_state = 42
)
```

Доля пропусков: 0.0

	gender	subject_code	age_group_5y	education	imputed_wage
0	male	26	15-19	basic_general	17241.378906
1	female	40	55-59	среднее профессиональное образование по програ...	482758.625000
2	female	8	50-54	среднее профессиональное образование по програ...	606896.562500
3	female	26	15-19	среднее общее образование (ранее – среднее (по...	62068.964844
4	female	78	40-44	среднее профессиональное образование по програ...	300000.000000

# Расчёты

Коэффициент Джини: классическая формула для взвешенных данных

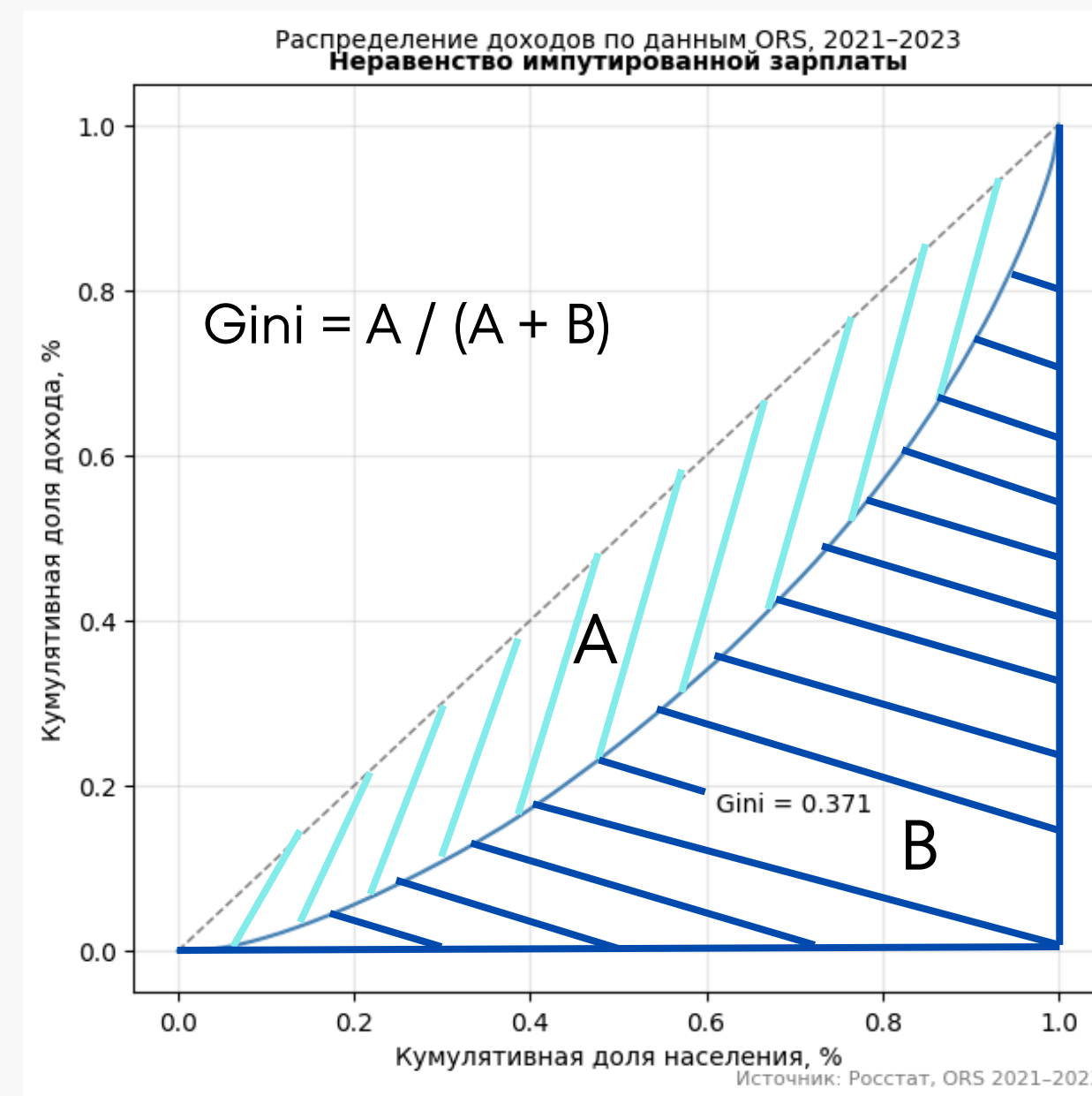
$$G = \frac{\sum_{i=1}^n w_i \sum_{j=1}^n w_j |x_i - x_j|}{2 \sum_{i=1}^n w_i \sum_{i=1}^n w_i x_i}$$

$$G = \frac{A}{A + B} = 2A = 1 - 2B$$

Кривая Лоренца

Зависимость кумулятивной доли дохода от кумулятивной доли населения

## График







# Симуляция сценариев

## Безвозмездный трансферт

Фиксированная сумма  $T = 30\,000$  руб выделяется  $N = 50\,000$  единицам внутри однородных групп (по региону)

Перебираем каждую группу, рассчитываем новый коэффициент Джини, выбираем группу с максимальным снижением неравенства

## Перераспределительная политика

Налог  $T$  взимается с  $N$  богатейших жителей одного региона и поровну распределяется между жителями другого региона

Выбираем наиболее населенные регионы, перебираем значения параметров размера налога и региона-донора, а также региона-получателя, оцениваем эффект

Метрика	Базовое значение	Адресный сценарий	Межрегиональный сценарий
Коэффициент Джини	$0,371 \pm 0,004$	$0,3706 (-0,0004)$	$0,3698 (-0,0012)$
Изменение, %	—	-0,11 %	-0,32 %
p-value	—	0,34	0,13

# Результаты

## Гипотеза 1

Прогноз: падение  
Джини на  $\approx 0,02$   
Фактически:  $-0,0004$   
пункта  
Статистически  
случайно ( $p > 0,3$ )

Гипотеза отвергнута

Точечные меры не  
эффективны

## Гипотеза 2

Прогноз: 2-ой сценарий  
эффективнее  
Фактически: улучшение  
в 3 раза, но низкая  
статистическая  
значимость ( $p > 0,1$ )

Подтверждена частично

Комбинация для  
устойчивых  
изменений

## Гипотеза 3

В ходе перебора  
значений во 2-м  
сценарии наблюдалось  
скачкообразное  
поведение  
коэффициента

Гипотеза подтверждена

Системные реформы  
Фискальные инструменты  
Институциональные изменения

## Библиография

- Chetty et al., 2016 в Q1 Journal of Political Economy
- Chetty, 2016; World Inequality Report, 2022
- Kanbur et al., 2021 в Journal of Development Economics
- Piketty & Saez, 2014 в American Economic Review
- Banerjee et al., 2017
- Milanovic, 2016
- Saez & Zucman, 2020
- Kanbur, 2021
- Alvaredo et al., 2018, Econometrica
- Налог на доходы физических лиц как инструмент регулирования социального неравенства - О. В. Староверова, Р. Г. Ахмадеев, А. А. Агапова

**Спасибо за  
ВНИМАНИЕ**

