



*Mobile Network Operator data –  
Methods for Integrating New Data Sources*

2022-IT-TSS-METH-TOO  
Project n. 101132744

***ISI WSC 2025***, 9 October 2025, The Hague

Integrating MNO Data into Official Statistics: Challenges and Innovation

# Can a traditional survey reduce the bias of mobile phone data based statistics?

Alexander Kowarik – Statistics Austria



[MNO-MINDS | Eurostat CROS \(europa.eu\)](#)

Project e-mail: [tss-meth-too@istat.it](mailto:tss-meth-too@istat.it)



This presentation is mainly based on the deliverable  
“***Deliverable 4.2 Report on a possible ad-hoc survey to improve the usage of MNO data in official statistics***” by

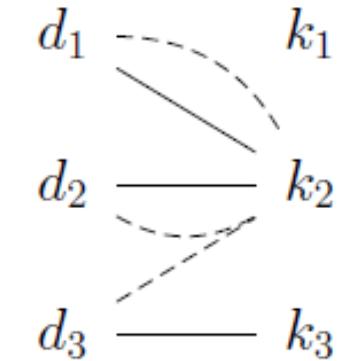
- Alexander Kowarik, Tiziana Tuoto, Loredana Di Consiglio, Gloria Deetjen, Maurice Brandt, Remy Kamali, Kira Gylling, Cristina Faricelli, Sara Piombo and Tiziana Pichiorri
- **FAKE:** Numeric results are just for illustrative purposes. They are extrapolated from the non-representative friendly user test in a very simple manner.



# Current Challenges with MNO Data



- **User ambiguity:** contract holder  $\neq$  primary user (family plans, business contracts)
- **Device duplication:** multiple devices/SIM cards  $\rightarrow$  duplicate records.
- **Behavioural variation:** differences by age, employment, lifestyle.



$$Z = \sum_{d \in D} y_d \left( \sum_{j, k \in U} c_{dk} a_{dj} \right) = \sum_{j, k \in U} \nu_{kj} y_j = \sum_{\substack{k \in U \\ \zeta_k > 0}} y_k \nu_{kk} + \sum_{\substack{k \in U \\ \zeta_k > 0}} \sum_{\substack{j \in U \\ j \neq k}} y_j \nu_{kj}$$
$$\nu_{kj} \stackrel{\uparrow}{=} \sum_{d \in D} c_{dk} a_{dj} \quad \zeta_k \stackrel{\uparrow}{=} \sum_{j \in U} \nu_{kj}$$

# Segmentation of MNO User Group



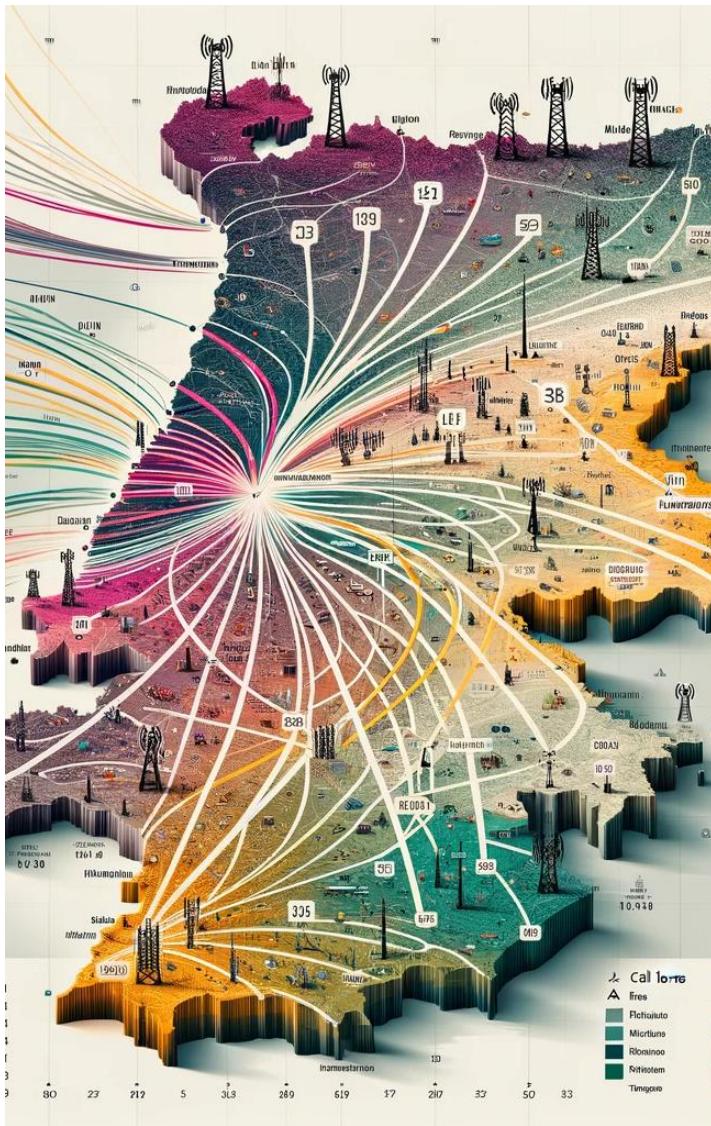
- Looking at the actual main user of a certain mobile phone
- Covered by standard **NSI sampling frame**:
  - Registered resident
  - Children via proxy information.
- Not covered in traditional surveys – **no frame readily available**:
  - Tourists, persons in transit
  - Non-registered residents, irregular migrants

# Need for (a) Dedicated Survey(s)



- **Purpose:** Collect systematic data on device usage and user profiles.
- **Objectives:**
  - Determine devices/SIM cards per individual.
  - Quantify mismatches between contract holders and users.
  - Analyse demographic usage patterns.
  - Gather socio-demographic background.

# Survey Questionnaire

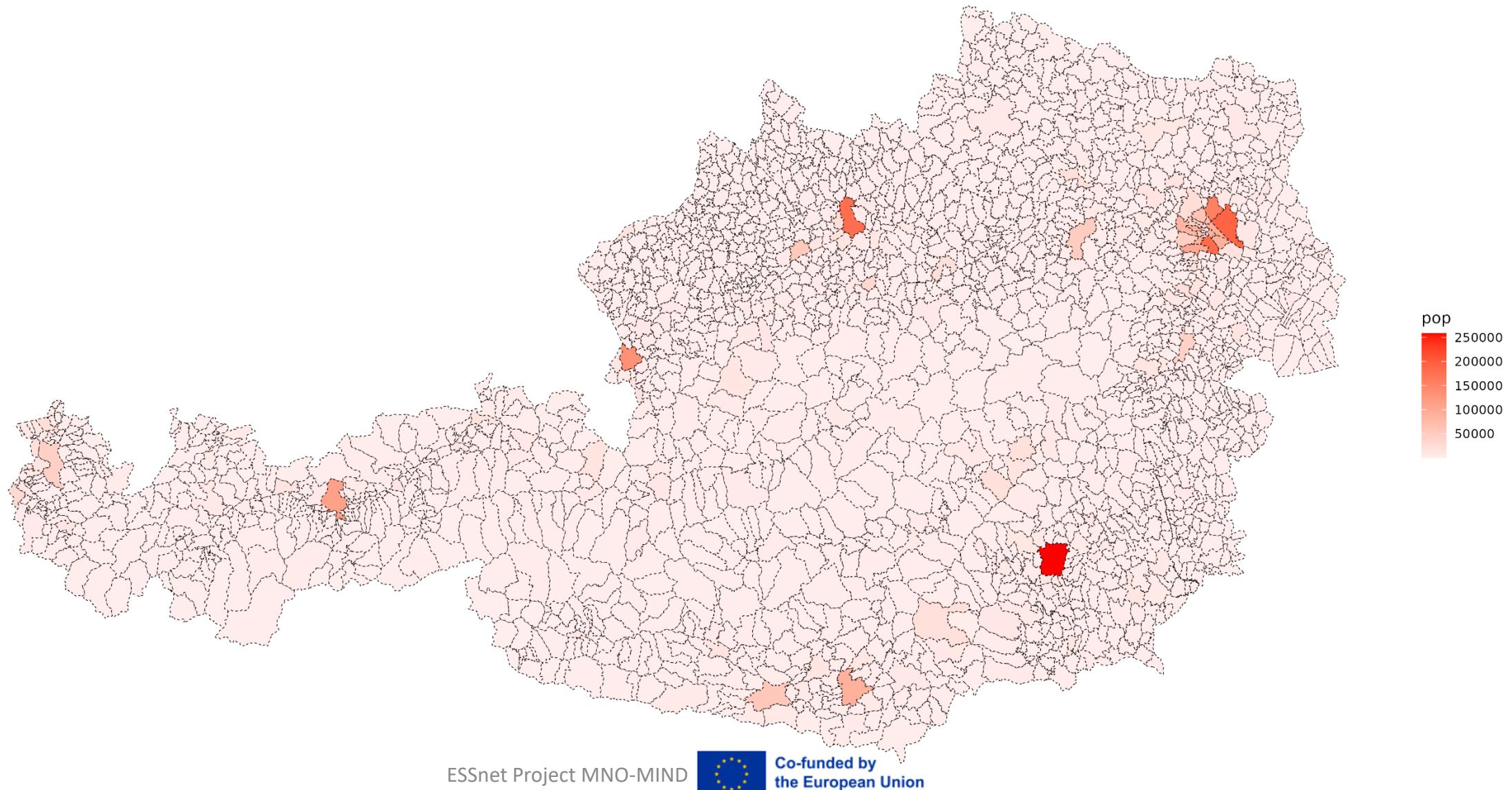


- Questionnaire Design
    - Core module with potential extensions for specific domains.
    - Sections:
      - Device and SIM Card Usage
      - Contract Details
      - Demographics & Lifestyle



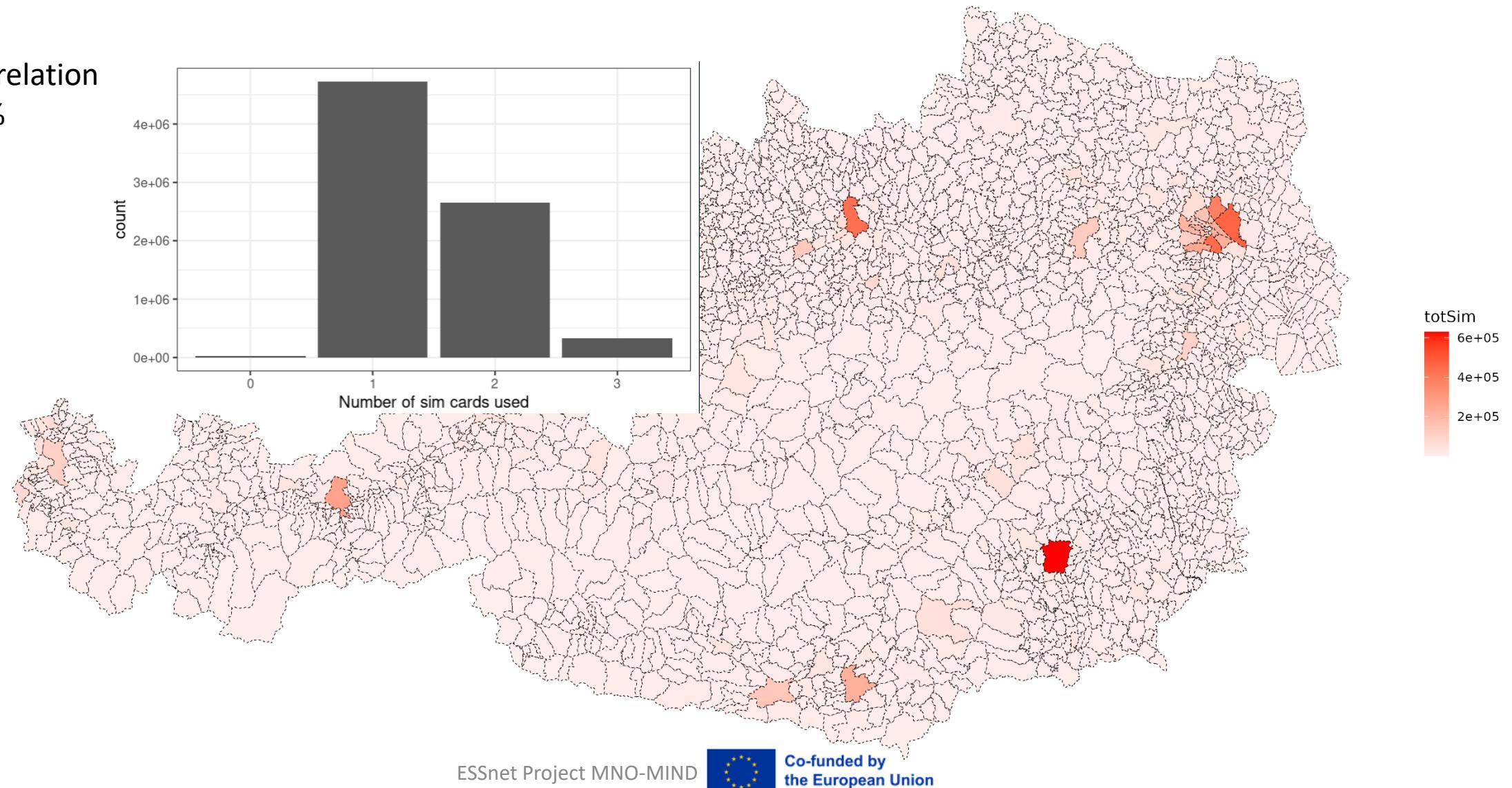
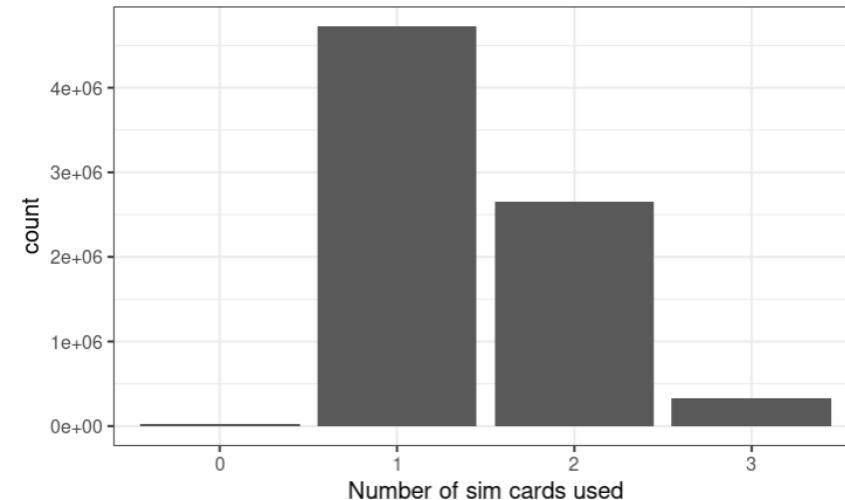
[https://ec.europa.eu/eusurvey/runner/MNOMINDS\\_WP4](https://ec.europa.eu/eusurvey/runner/MNOMINDS_WP4)

# Population in Austria >15years, about 8 million



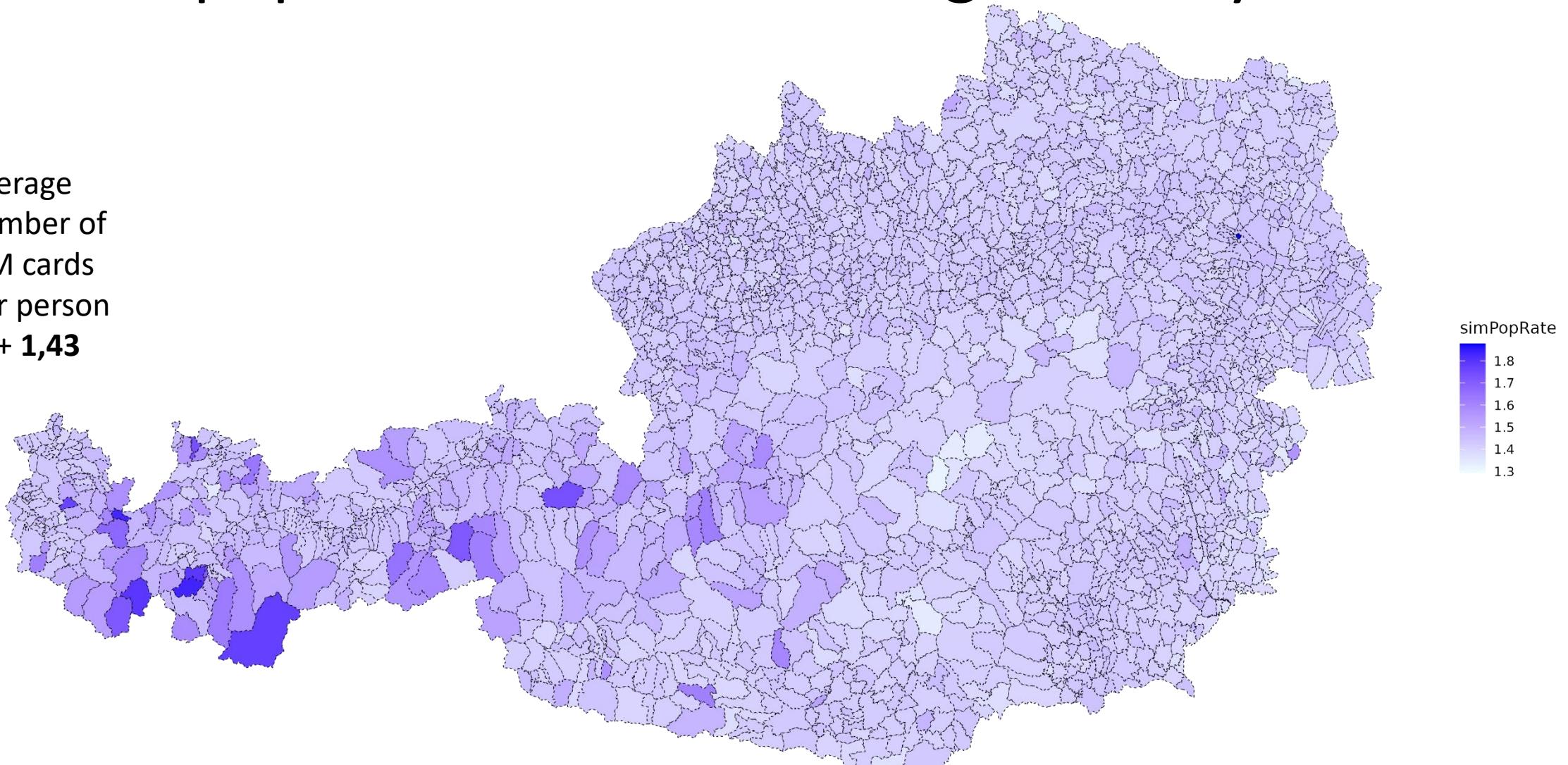
# Approx. 11 million local SIM cards in use

Correlation  
97%



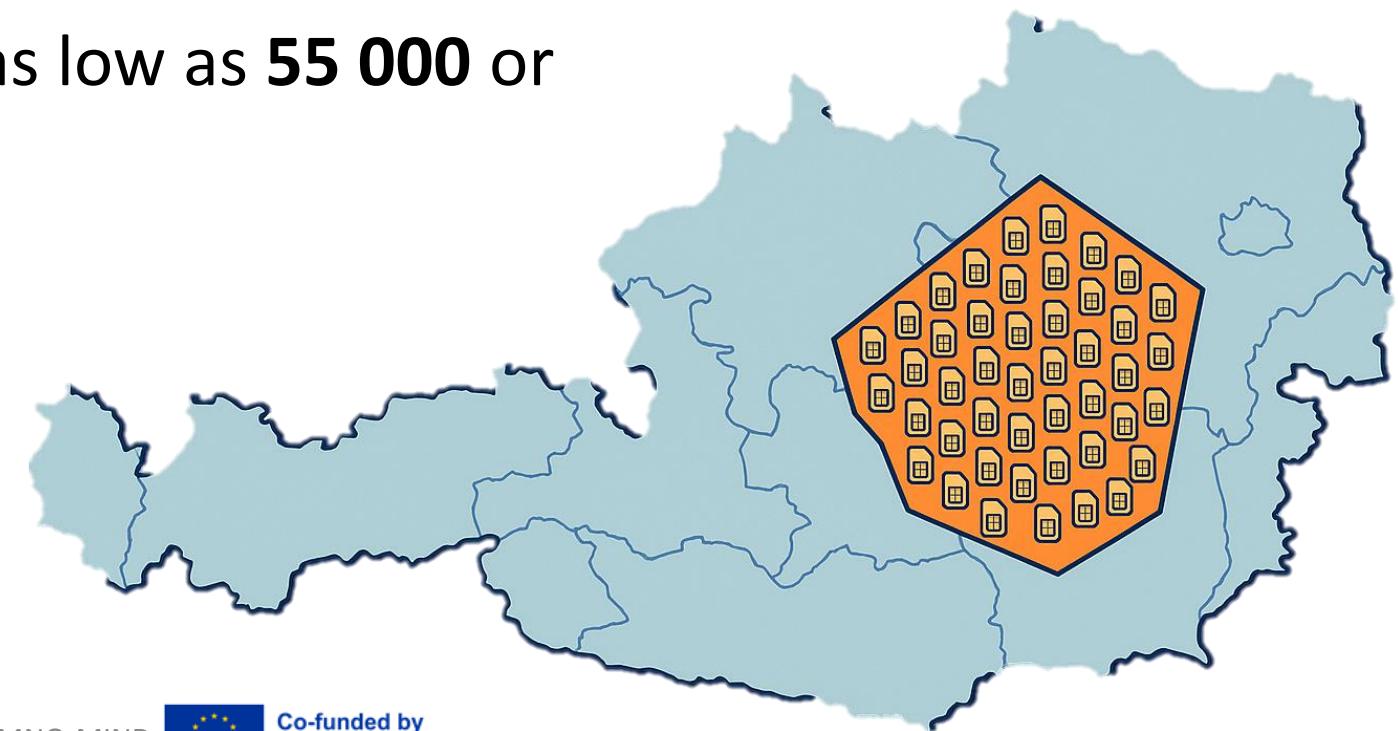
# SIM-to-population ratio varies significantly.

Average  
number of  
SIM cards  
per person  
15+ **1,43**



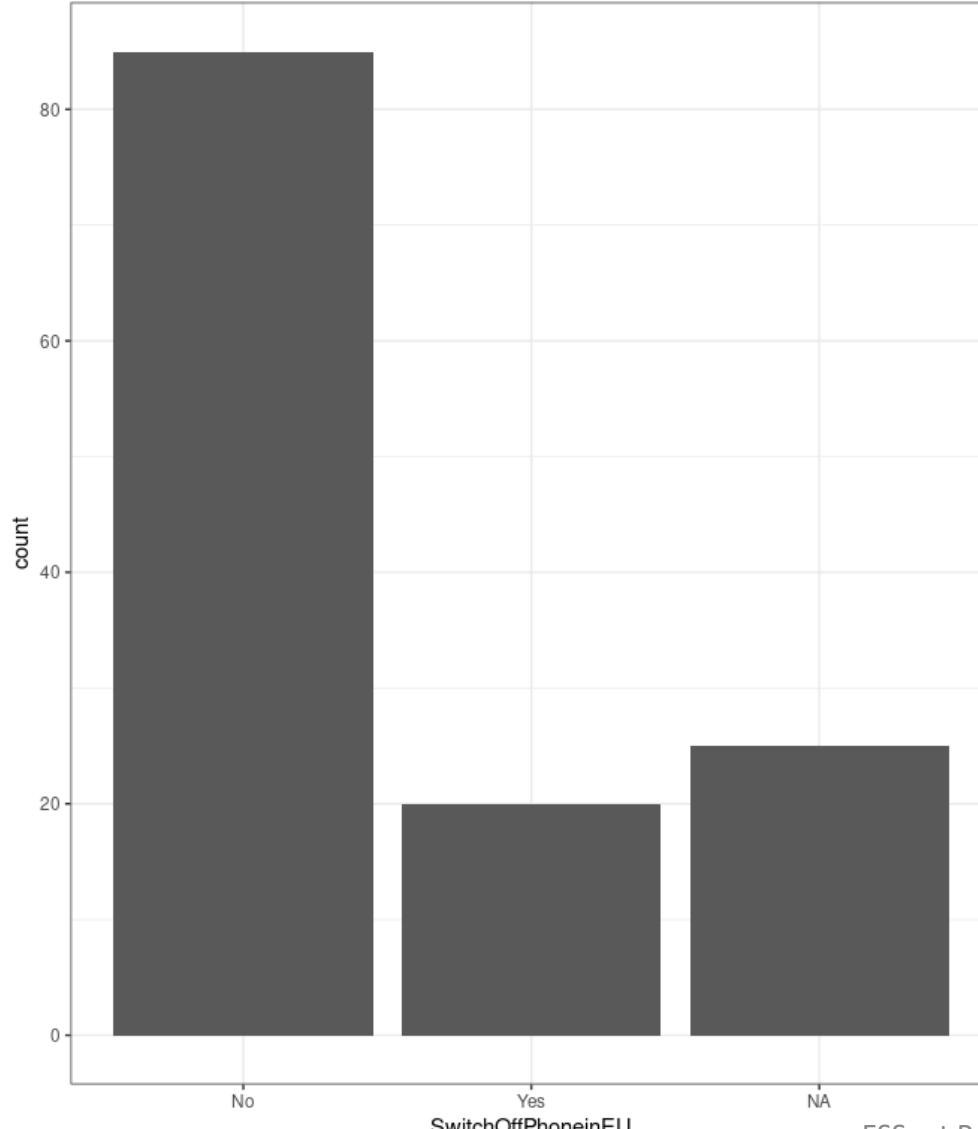
# Illustrative Example (not real data).

- Example: **100 000 SIM cards** observed in an area within a time window.
- With the mean ratio SIM cards/person this leads to an estimate of **70 000 persons**
- But in specific areas it could be as low as **55 000** or as high as **77 000**

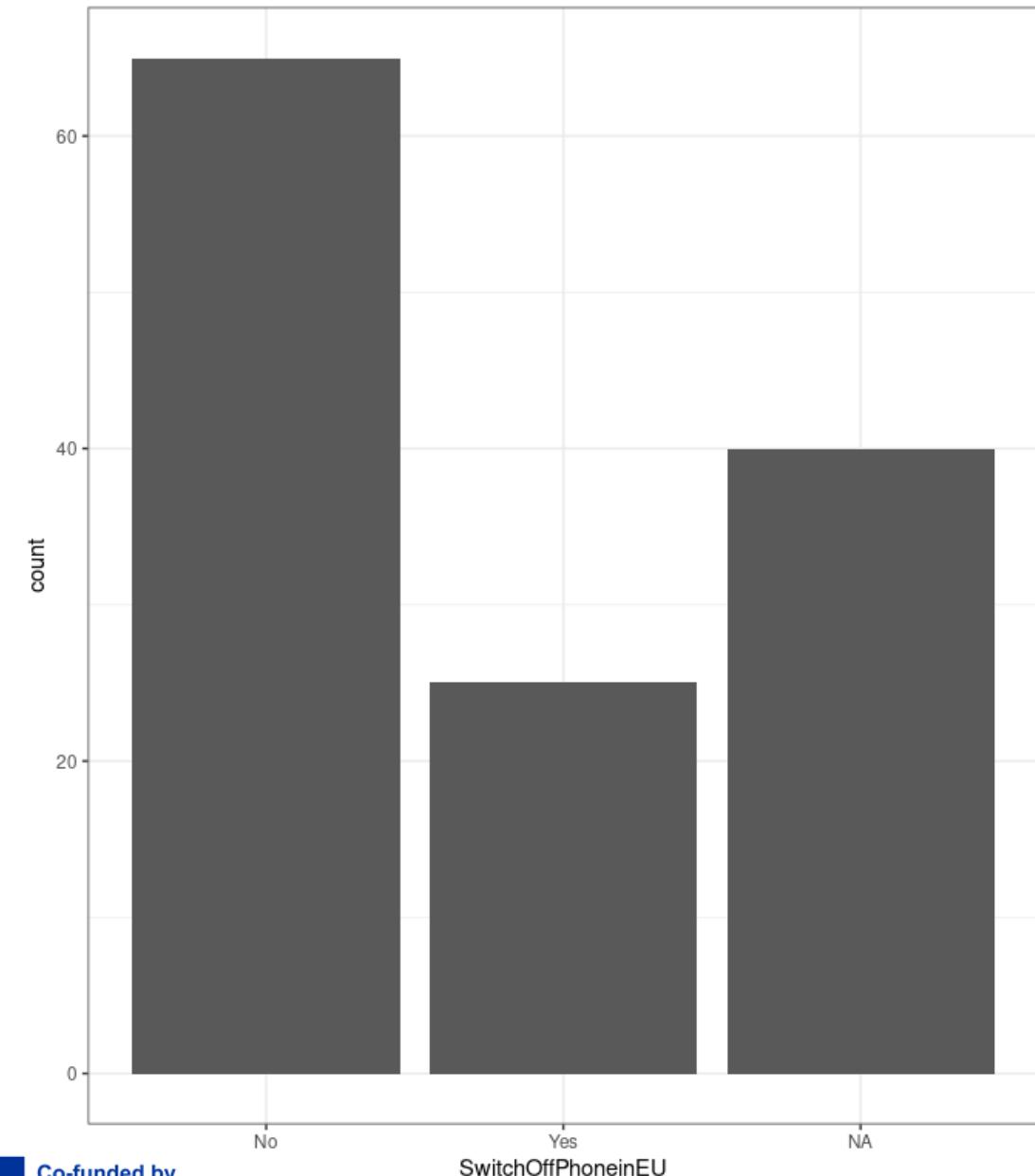


# Phone Usage During Travel

Does switch off at least one phone during EU travels?



Does switch off at least one phone during outside EU travels?



# Alternative Approaches



- Area Sample Surveys

- Target population : National SIM cards
- Sampling based on geographical areas (antennae).
- Consideration of technical and legal challenges.
- Should cover all MNO users
- Segmentation can be estimated
- # SIM cards per user

- Border Surveys

- Airports, border crossings etc.
- Covers international tourists.

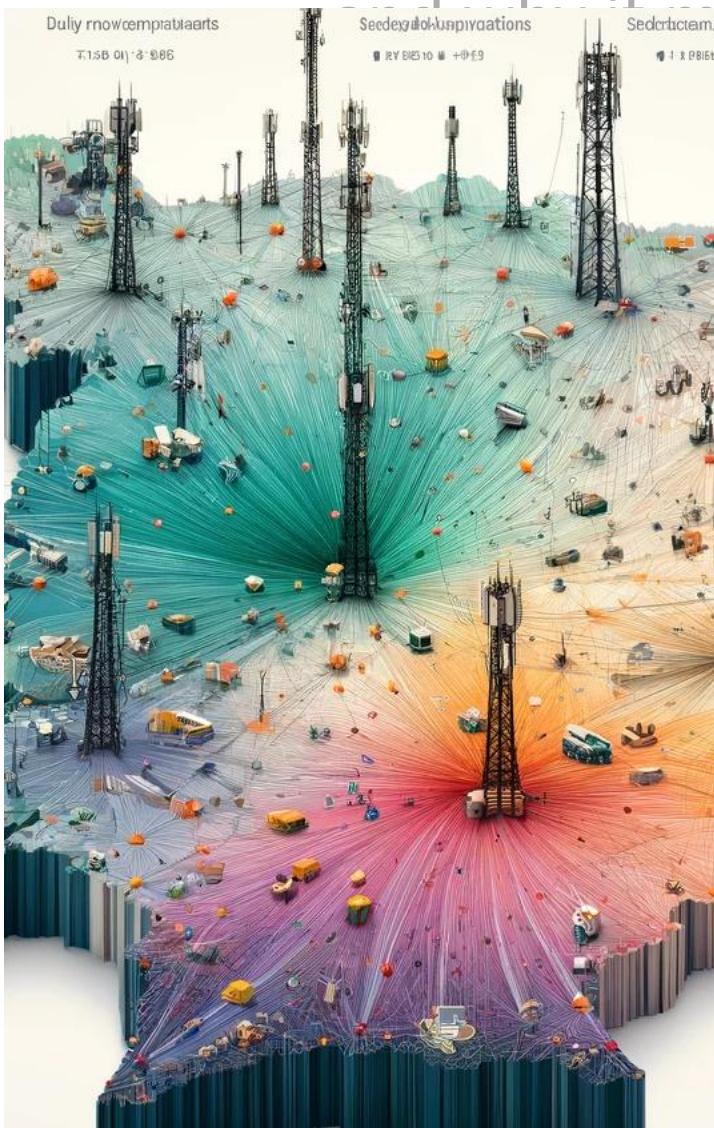
$$Y = MW \quad \text{and} \quad W = \xi^{-1} \{P_r(1 - P_w)\}(1 - P_{nr})^{-1}$$

$P_r$  = Proportion of foreign residents among travellers

$P_w$  = Proportion of workers among foreign residents

# What Could an Area Sample Look Like?

ight not work...



- **Sampling Design:**

- Sample geographical areas or antennae and time intervals
- Stratification into areas with different expected shares of non-residents
- Oversampling of areas with higher rates
- Either take all active SIM cards or sample SIM cards

- **Survey Design:**

- Questionnaire (simple adaptions from the resident version)
- Contact via MNO – e.g. SMS to phone or cell broadcast
- Expected response rate: very low

# What Is the Optimum?



- Multiple surveys?
  - Traditional surveys for residents + MNO data limited to residents?
  - Survey covering non-residents
    - Area sample?
    - Border survey?
- Ideally the measurement (position data) would be linked to the respondent. Theoretically, possible when starting from a SIM card.

# Conclusion



- **Survey necessity:** Quantify and reduce MNO data biases.
- **Implementation:** Promote adoption across ESS countries.
- **Goal:** Improved data quality through collaborative effort.

