



seriss

SYNERGIES FOR EUROPE'S
RESEARCH INFRASTRUCTURES
IN THE SOCIAL SCIENCES

www.seriss.eu
@SERISS_EU

POPULATION REGISTERS IN EUROPEAN COUNTRIES: A LITERATURE REVIEW

Angelica Maineri (Tilburg University) – a.m.maineri@uvt.nl

Ilziya Mindarova (Higher School of Economics (HSE))

Ruud Luijkx (Tilburg University)



This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 654221.



Table of content

- Population registers: what, uses
- Availability of PRs
- Access to PRs
- Quality of PRs
 - Quality criteria
 - Quality assessment
 - Case studies
- Data protection regulation
- Conclusion & Discussion





Population registers: what?

- “an individualized data system, that is, a mechanism of continuous recording, and/or of coordinated linkage, of selected information pertaining to each member of the resident population of a country in such a way to provide the possibility of determining up-to-date information concerning the size and characteristics of that population at selected time intervals” (United Nations, 1969, p. 1)





Population registers: uses

- Administrative purposes
- Produce demographic statistics
- Sampling frames?

SERISS WP2
Deliverables 2.1, 2.2, 2.5





Advantages of population registers as sampling frames

- Complete list of population...
 - Key requirement for representative samples of the population
- ... with contact information
- Auxiliary information
 - Non-response adjustment
 - Better stratification, complex sampling design

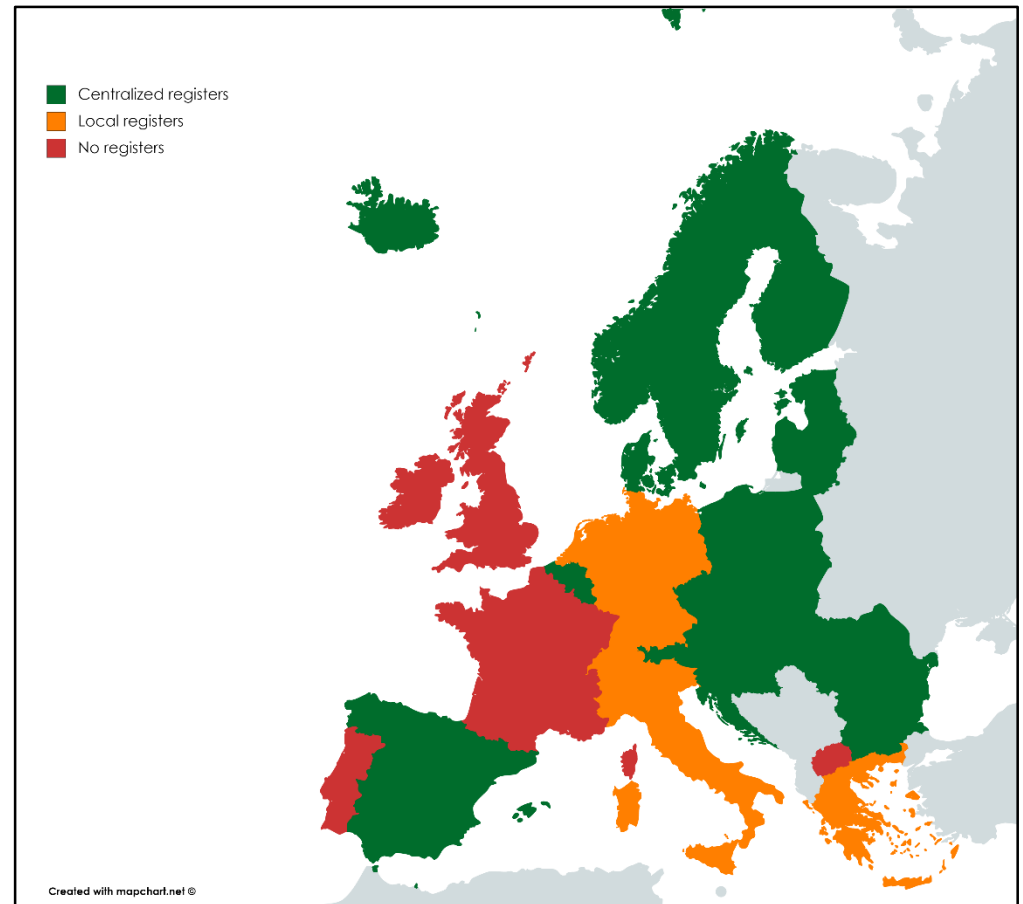




Availability of PRs

Sources:

Redfern 1989; Eurostat 1997; Krabina et al 2005; Poulain and Herm 2013





Access to PRs for sampling or research

- Legal constraints
- Digitalisation
- Cost

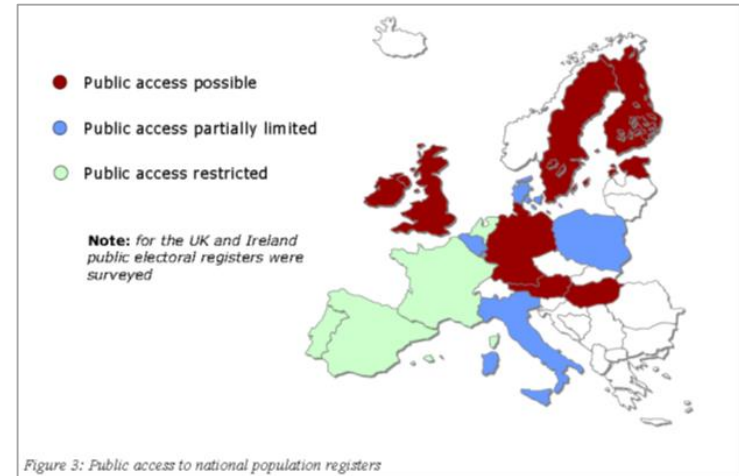


Figure 3: Public access to national population registers

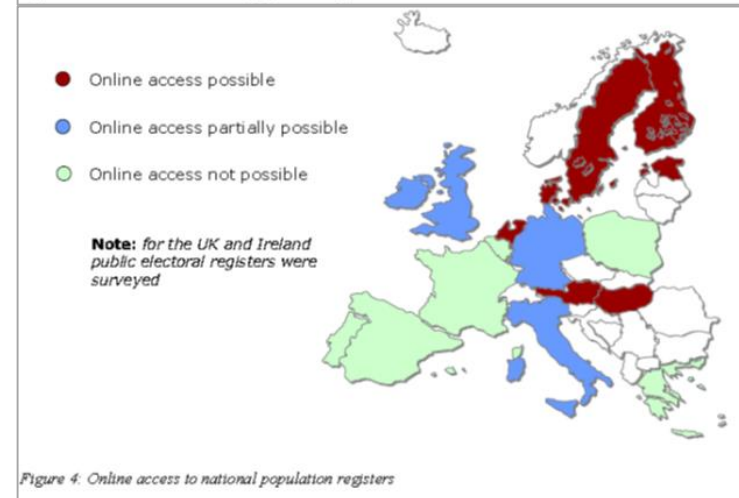


Figure 4: Online access to national population registers

Source: Krabina et al. (2005, p. 4-5)





Access – examples

- Finland
 - Apply for samples from the Longitudinal Census Data File and Longitudinal Employment File
- Statistikbanken (Denmark)
 - Access online to microdata
 - Only researchers affiliated with Danish institute/environment





Quality of registers

- Coverage issues
- Inaccurate information
- ...

➔ Affect quality of samples





Quality criteria for use as sampling frames (1)

Partly overlapping with 12 registers' quality criteria of Eurostat (2003)

- **Incompleteness, or undercoverage:**
 - Discrepancies between the actual target population and the one listed in registers.
 - Often, specific groups are not covered.
- **Duplicates, or overcoverage**
 - An individual or a household may appear several times in the register.
- **Out-of-scope, or overcoverage**
 - People, households or addresses which do not belong to the population of interest can be listed in the register.
- **Inaccuracy, or unreliability**
 - No regular updates, leading to a number of incorrect names and addresses.
 - Misclassifications: a woman may be registered as a man, or a person's age may be incorrectly registered, for example.





Quality criteria for use as sampling frames (2)

- **Difficulties of access and privacy issues:**
 - Much time or financial resources may be needed to get access to the register for sample drawing, or access may not be possible at all.
- **Lack of auxiliary information in the register:**
 - Not enough contact information is available in the register to find and contact people.
 - Lack of information about demographics and/or household composition of the persons in the register can make it hard to calculate design weights.
- **Complexity or poor usability**
 - When the register cannot be handled easily, is not well documented and logically/systematically organised, in a readable format, and if is all the information it contains is not coded in a consistent and understandable way, etc.
- **Clustering of sampling units**
 - The register may contains households instead of persons, or building blocks / postal code areas instead of individual addresses.





Sources of quality issues

- Underreported migration events
- Mistakes in documentary evidences
- Lack of self-reported amendments

- Differently distributed among sub-populations





European registers and quality assessment

- The literature:
 - Heterogeneity in the definition of registers
 - Country-specific works (not many in English)
 - Different fields
 - Some outdated
 - Not many accounts of quality checks (direct or indirect)





Nordic countries

- Long history
- Available PRs, often accessible
- Overall good quality and frequent quality assessment studies
 - Sweden: coverage of 95% immigration event, 91% emigration
 - Denmark: 99.2-99.4% births reported, 99.3-99.6% deaths





Focus: Norway

- 97.1% correct addresses in CPR (Thorsdalen, 2008)
- Falnes-Dalheim and Pedersen (2012): study on postal returns in a survey on 218000 immigrants
- 10% letters returned
- Higher incidence of uncorrect addresses among immigrant population





Focus: Finland

- Quality study: systematic and repeated measurement of register quality
- Questions at the end of the LFS
- In 2007: 98.8% correct addresses in PR (Hokka and Nieminem, 2008)





Western European countries

- Not available in all countries, not always accessible
- Only some information on quality
 - Switzerland: undercoverage of immigrants, no figures
 - Germany: registers by regions, quality may vary





Focus: Netherlands

- Gerritse et al. (2016): linked registers
- Capture-recapture estimation
- 0.5-1.1% undercoverage, 0.2% overcoverage





Focus: Belgium

- Poulain et al. (1992): 3Bbis study
- Test reliability of data collected via retrospective life history interviews
- Check against register information
- 4% migration not reported, 20% children moving out not reported





Southern European countries

- Seldom available, difficult to access
- Not many quality accounts
 - Italy: 5% bias on unemployment rate due to coverage error





Fair Information Principles (FIPs) (1)

- **Collection Limitation Principle**
 - limits to the collection of personal data, which have to be obtained lawfully and, if appropriate, with consent
- **Data Quality Principle**
 - data should be relevant according to the purpose they are collected for, and as much accurate, complete and updated
- **Purpose Specification Principle**
 - at the time of data collection, purposes have to be specified. Subsequent purposes should comply with those
- **Use Limitation Principle**
 - data should not be disclosed for other purposes, except under explicit consent of the subject or if decided by the authority of law.

OECD's guidelines (2013)





Fair Information Principles (FIPs) (2)

- **Security Safeguards Principle.**
 - Data should be protected (IT security!)
- **Openness Principle**
 - there should be an overall policy of openness.
- **Individual Participation Principle**
 - Individuals should be able to access data and do have some rights on them (e.g. challenge to obtain rectification).
- **Accountability Principle**
 - the data controller should be accountable for complying with these principles.

OECD's guidelines (2013)





EU General Data Protection Regulation (GDPR) (May 2018)

- Increased territorial scope
- Penalties
- Consent
- Data subjects rights
 - Breach notification
 - Right to access
 - Right to be forgotten
 - Data portability
 - Privacy by design
 - Data protection officers





Population registers as sampling frames?

- Availability, accessibility and quality vary by country
 - North-South divide
- Lack of comprehensive, European-level study
- Important to study what is done in practice of survey infrastructures





References

- See reports on the SERISS website
 - <http://seriss.eu/resources/deliverables/>

