Administrative Data Linking: Enriching Administrative Data with Surveys

Tom Emery (NIDI/GGP)

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

- Work Package 6 on new forms of data
- Task 6.2 focuses on the use of Administrative Data linking and the Legal and Ethical Challenges
- GGP has already conducted such linking and so will provide a central case study within the work
  - What does data linking constitute?
  - What obstacles and challenges are there?
  - What opportunities are there?
Background

- Secure Site Access to administrative data is common
- Virtual Access Points are also increasingly prevalent
- Combined, multi-national access points are also being pursued (DWB)
Background

- GGP is a longitudinal, cross-national survey currently conducted in 19 countries
- Survey focuses on families and demographic behavior
- In several countries respondents are linked to population registers
- Less common is integration of administrative data in PUF
Example of Sweden

• Carried out by Stockholm University in collaboration with Statistics Sweden
• 9,688 Respondents aged 18-79 interviewed in 2012-13 (Response Rate of 54.7%)
• Phone survey with postal follow up and registry linkages
• Registry follow up provided every 3 years (2\textsuperscript{nd} Wave is being prepared for release now)
GGP Experiences

1. Improving Scientific Access to Population Registers
   • Survey Infrastructures work on Open Access Model
   • Population Registries have highly restricted access
   • Surveys can provide a ‘key’ of consent to allow for linking with surveys and even other data sources
GGP Experiences

1. Improving Scientific Access to Population Registers
   - Surveys and Registers have different definitions of Anonymity
   - Re-identification risk in social surveys is relatively high compared to administrative data
   - Successful linking depends on NSO accepting consent as circumventing this risk
GGP Experiences

2. Validating Population Registry Data
   • Registry and Survey data differ
   • Data linking helps identify why, when and how
   • 18.3% choose to correct their educational level in Sweden
   • Research on partnership and fertility histories in Netherlands
Administrative Data Linking in the GGP

• Challenges:
  • Data Protection
  • Respondent Consent
  • Data Privacy

• Opportunities:
  • Data Validation
  • Added Scientific Value to administrative data
  • Better quality survey data
Next Steps

• Other successful data linking projects and their experiences (ESS, SHARE and others)

• Unsuccessful data linking projects and their experiences (why did they fail?)

• Consolidating these into best practices of linking administrative and survey data
Thank you