Use of MNO data to improve "Beyond-GDP" measurement: MNO-MINDS project challenges

Marie-Pierre Joubert, Insee - France Istat team : Maria Grazia Calza, Maria Francesca D'Ambrogio, Immacolata Fera, Laura Tosco, Tiziana Tuoto, Tamara Zangla

The Project 101132744 — 2022-IT-TSS-METH-TOO is co-funded by the European Union

October 15, 2024 - CESS conference

Augmenting Beyond GDP Measurements with Data Integration

- Beyond GDP : Focus on societal well-being, sustainability, and progress beyond economic output (GDP).
- Mobile Phone Data : Provides real-time, granular insights into mobility and presence patterns.
- Combining Data Sources :
 - ▶ Traditional NSIs Data : Censuses, mobility surveys, administrative data.
 - New Data Sources : Sensors, Traffic data, Satellit imagery, Credit card transactions,...
- Together, these offer a more comprehensive, multidimensional view of well-being and economic development.



Enhanced Insights through Data Fusion

Granular Economic Indicators :

- ► Surveys + Mobile Data : Improve survey's coverage ⇒ regional economic insights.
- ► Credit Card + Mobile Data : Links real-time consumer behavior with presence patterns ⇒ spending trends.
- Administrative + Mobile Data : Track informal economies, underemployment trends.
- Social and Environmental Well-Being :
 - Satellite + Mobile Data : Track environmental sustainability, climate resilience.
 - Credit Card + Mobile Data : Analyze real-time social inequality and economic disparities.
 - Mobility Surveys + Mobile Data : Improve urban planning, access to services.



Producing official statistics based on MNO data



Source : ESS Task Force on MNO data Position paper - september 2023



ESS Net MNO MINDS

- Goal : development of methods for integration of MNO and non-MNO data
- Istat (IT) as coordinator, along with 9 NSI beneficiaries : Stat (AT), Destatis (DE), INE (ES), INSEE (FR), CBS (NL), SSB (NO), INS (RO), SCB (SE), INE-PT (PT)
- ▶ WP1 coordination, management, dissemination
- WP2 landscaping analysis of possible non-MNO data sources to be integrated with MNO data
- WP3 development of methodologies and open-source tools for integrating MNO and non-MNO data sources
- ▶ WP4 proof-of-concept of an **ad-hoc survey** to improve MNO data





ESS Net MNO MINDS

Milestones

- Started 1st November 2023 End 31st of October 2025
- Kick-off-Meeting : 5th December 2023, Istat, Rome
- Sprint Event 10 11 June 2024 in Vienna Statistics Austria
 - 20 consortium members from the 10 participating countries
 - 11 non-NSIs experts, from MNO compagnies, data aggregator compagnies, the flowminder fundation,...

• Closing conference : 2,5 days in automn 2025.



Which data are the most relevant to combine with MNO data?

Primary data scoring

- ► Technical cost of handling the dataset
- Easiness of access to a temporally and geographically harmonized data source.
- Easiness of access to **detailed data**
- Range of possible use cases
- EU availability of these data
- Accuracy and robustness of the information available on these data
- Non-sensitivity of these data

Secondary data scoring

Quality requirements derived from ES Code of Practice



Focus of the adequacy with the use cases

- Improving population's coverage : are these data complementary to MNO data? Do they allow to improve MNO data population representativeness? Or the geographical and spatial accuracy? Are there difficulties to due to different irreconcilable concepts (definitions, time periods, ...)?
- Improving precision of analyses : do these data allow to provide preciser analyses (ex : socio-economic indicators, indicators of the mobility type, indicators of the commercial activity,...)
- Broadening the scope of issues covered by official statistics : do these data, combined with MNO data, allow to cover new topics?



Focus on the access to data

- Who owns the data? Public administration, one company, several companies?
- Are there limitations to the amount or aggregation level of data that can be accessed? What is the nature of this limitation? Legal, technical, financial, ethical, other?
- > Are there potentially **competing uses** (operator publishing similar statistics, etc.)?



Most Promising Data Sources to Combine with MNO Data

Most Promising Data Sources

Census; Population **Register** (fiscal data, ...); Combination of survey and register; Transportation **Surveys**

Promising Sources but which Require Substantial Work or have Limited Accessibility

Vehicle, bicycle and pedestrian **sensors**; Vessel (boat) **traffic** data; **Pollution** data; **Satellite** data; Electronic **invoices**; **Tourism** Household and Border Surveys; Tourism **platform data**; **Credit Card Transaction Data**



Other Data Sources to Combine with MNO Data

Less Relevant Sources

Google Maps Popular Time; Smart Meters; Connected Vehicles; Social Media

Sources Still to Be Analyzed by the ESSNet

Tickets sold at mass events ; Tourism accommodation statistics ; Land use and land cover registers



Reference frame for methods to combine MNO and non-MNO data



M1: M-executor methods; M2: M-enabler methods.



Three main approaches for infering uncertainty

Randomisation

- A survey is conducted under some known sampling design and the uncertainty of the resulting estimation is considered to be dominated by the associated sampling error.
- eg. Tourism Statistics in Indonesia : Combination of macro MNO-counts and a sample-survey to estimate device-to-person adjustment factor

Quasi randomisation

- Postulate a model of the observation mechanism of the MNO data as if they had been obtained by designed randomisation, and the same mechanism is applicable to all the attributes associated with detected devices.
- eg. comparison between inhabitants numbers based on MNO-Home and usual residence concepts and post-stratification model.

Three main approaches for infering uncertainty

Super-population modelling

- Unlike quasi-randomisation, which builds a selection model applicable to multiple outcome variables, a super-population model is tailored to specific outcome variables, such that different models are needed for different outcomes generally.
- eg. Assume that the target distribution (eg. related to the population of residents) is the same as a distribution derived from mobile devices directly. Then inputes the distributions required for disaggregation by those derived from geotagged MNO data.



An ad-hoc survey to improve MNO data

Current challenges with MNO data

Device-User Ambiguity :

The person paying for the service (contract holder) may not be the primary user (e.g., family plans, business contracts), leading to inaccurate socio-demographic data.

Multiple Devices/SIM Cards :

Many users have multiple devices or SIMs, creating potential over-coverage and duplication. Deduplication methods exist but are prone to errors in scenarios like group travel or work phones.

Varying User Behavior :

Usage patterns differ widely across users (e.g., age, employment, lifestyle), affecting the accuracy of derived statistics. Some users turn off phones at night, during work, or on vacation.

eurosta

An ad-hoc survey to improve MNO data

Questionnaire design

- Device and SIM Card Usage : questions on the number of devices, their purposes, and details on SIM card usage.
- **Contractual Relationships** : Information on the contract owner, duration with current service provider, and previous providers.
- User Demographics and Lifestyle : Questions on age, gender, household composition, employment status, commuting habits, and usage during different activities and travels.





Conclusion

- Combining MNO and non-MNO data offers a significant opportunity to enhance the accuracy, precision, and scope of official statistics.
- It allows for better population coverage, more granular analysis, and the ability to tackle new issues not previously covered.
- The ESS Net MNO MINDS project provides a structured framework for developing methodologies, integrating different data sources, and exploring innovative solutions.
- To address challenges like device-user ambiguity and multiple device use, additional research and ad-hoc surveys will play a crucial role in refining MNO data usage.
- By overcoming these challenges, we can ensure high-quality, reliable statistics that serve the evolving needs of policymakers and society.



Thank you for your attention

Marie-Pierre Joubert - Insee, SSP Lab, *marie-pierre.joubert@insee.fr* (presenting author) Tiziana Tuoto - Istat (ESSNet MNO-MINDS leader) All participants from ESSNet MNO MINDS

