

12th biennial IFC Conference: “Statistics and beyond: new data for decision making in central banks”

22-23 August 2024

Mobile phone data and the “travel” item in the Balance of Payments: the Bank of Italy’s experience¹

Andrea Carboni, Claudio Doria and Valeria Valletta,
Bank of Italy

¹ This contribution was prepared for the conference. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.

Mobile phone data and the "Travel" item in the Balance of Payments: the Bank of Italy's experience

Andrea Carboni, Claudio Doria, Valeria Valletta

Abstract

Despite the potential of Mobile Phone Data (MPD) as a resource for international tourism statistics, few countries use them in official statistics, due to the difficulty in translating telephony terminology into tourism's own definitions. Since 2020, the Bank of Italy has integrated MPD into the estimation process of the "Travel" item of the Balance of Payments, obtaining greater coverage of international borders and reduction in costs compared to traditional survey methodologies. This paper describes the advantages and challenges related to the use of MPD and their performance, through a comparison with administrative data sources.

Keywords: balance of payments, big data, mobile phone data, tourism statistics.

JEL classification: C55, C83.

Contents

Mobile phone data and the "Travel" item in the Balance of Payments: the Bank of Italy's experience.....	1
Introduction.....	2
1. The use of MPD in tourism research.....	2
2. The Bank of Italy's Survey on International Tourism	3
3. The path toward the use of Mobile Phone Data.....	3
3.1 Adoption of MPD and their benefits.....	4
4. The use of MPD by the Bank of Italy for compiling the Travel item of the Balance of Payments	5
4.1 The MPD used and the main characteristics of the MNO	5
4.2 How to make the most of MPD.....	6
5. Quality assessment of MPD.....	9
5.1 MPD Quality in 2023.....	10
6. CONCLUSIONS	12
APPENDIX	13
REFERENCES	14

Introduction

Tourism has a significant impact on social, environmental, and economic levels, representing an important share of the global GDP (Leng et al., 2016). Recently, tourism studies and compilation efforts have shifted from expensive surveys to using Mobile Phone Data (MPD) for their cost-effectiveness, continuous data collection, and broad geographic coverage.

Despite growing international attention (Demunter, 2017; UN, 2019; UN-ESCAP, 2022; Eurostat, 2014, 2023), the leverage of MPD for official tourism statistics remains very limited because of the strict privacy regulations and the challenge of translating telephony terminology into the definitions specific to official tourism statistics (Grassini & Dugheri, 2021; UN, 2019).

In this context, the Bank of Italy has gradually integrated MPD into the estimation process of the Travel item in the Balance of Payments (BoP), following an experimentation phase, started in 2018. This study describes this integration, highlighting the main characteristics of the data and their application.

1. The use of MPD in tourism research

MPD play a crucial role in various fields, including healthcare, socioeconomics and, disaster response (UN, 2019). In travel studies, Ahas and Mark (2005) were among the first to recognize the potential of MPD. Subsequently, interest in MPD grew, mainly to study travel behaviours and identify travel patterns (Phithakkitnukoon et al., 2015; Chen et al., 2018; Xu et al., 2021; Park et al., 2020), to assess the impact of various factors on tourism such as climate change (Kubo et al., 2020), and to define new indicators (Leng et al., 2016). Ahas et al. (2007) use MPD to identify seasonal travel patterns and nationality differences between summer and winter tourism areas. Ahas et al. (2008) describe MPD, highlighting privacy access challenges, data collection methods, spatial resolution, and representativeness issues. Raun et al. (2016) leverage on MPD to differentiate tourist flows in Estonia based on time, space, and composition, revealing seasonal patterns, diverse travel durations, and varied nationalities.

While MPD are extensively used in tourism research, their integration into official statistics is more recent. Notable examples include contributions from France (Cousin & Hillarieu, 2020) and Netherlands (Heerschap et al., 2014). The Estonian Central Bank, Eesti Pank, has been using Call Detail Records (CDRs) for the compilation of the Travel item of the BoP since 2008 (Eesti Pank, 2019, Saluveer et al., 2020). BPS-Statistics Indonesia has been using MPD (signaling data) for compiling official statistics on the number of inbound international travellers since 2016. MPD have allowed to achieve better coverage and timeliness compared to traditional sources (Lestari, 2018; Rahmadian, 2023; Noviyanyi et al., 2020).

However, MPD provide insights into travel flows but need to be integrated with traditional surveys for comprehensive tourism statistics, since MPD lack some crucial qualitative data (for instance, the distinction between business and personal travel) and, most notably, bear no information on expenditure, which is crucial for BoP statistics.

2. The Bank of Italy's Survey on International Tourism

The Bank of Italy has been conducting the Survey on International Tourism since 1996, with the main purpose of collecting information for the compilation of the "Travel" and "Transport - passengers" items of the BoP.

Using a stratified sampling plan¹, the survey takes place at international border points (ports, airports, railway and road crossings) and relies on interviews and counts. A subset of counts are "qualified", as they collect information on the country of residence or destination of the travellers² with a counting step based on the intensity of traveller flow and the specific crossing point characteristics³.

The purpose of the qualified counts is to estimate the size of the reference population, namely the total number of incoming non-resident international travellers and outgoing resident travellers, by country of residence or destination, respectively. Counts are supplemented with data provided by the authorities responsible for border points (e.g. airport authorities). These administrative data provide the "exact" population size of the reference universe but do not provide information on the residence of travellers.

The interviews capture expenditure and other relevant aspects for compiling the travel item of the BoP, as well as additional information useful for a broader understanding of the touristic phenomenon. Each interview takes place at the end of the journey and is administered to a single individual per group of travellers, who also responds on behalf of the other members.

The information acquired with the questionnaires is then grossed up to the reference population by integrating the findings from interviews with those of the counting operations and the administrative data sources.

3. The path toward the use of Mobile Phone Data

The process that led to the integration, in 2020, of MPD in the estimation of the Travel item took place in several stages.

¹ Using the following stratification variables: direction (inbound/outbound), mode of transport, border crossing, type of day (weekday/holiday), time slot, day of the week.

² The information is gathered either through brief interviews or by using license plates as a proxy for residence at the road border points.

³ For example, at a road crossing with a count step of 5, the nationality of the vehicle's license plate will be recorded once every five vehicles.

Initially, a pilot study was conducted in 2018 (Carboni et al., 2023) which saw the collaboration of the Bank of Italy with one of the main Italian Mobile Network Operators (MNOs). Subsequently, during the period when the COVID-19 emergency placed Italy in an almost complete lockdown, the absence of traditional information sources required the use of alternative ones, including MPD, limited to non-residents detected in Italy (Carboni et al. 2020).

3.1 Adoption of MPD and their benefits

The experience gained with MPD during the experimentation and the emergency phase led to their integration into official tourism estimation procedures by the end of 2020.

MPD have replaced part of the counting operations conducted at border crossings, with an overall reduction in the cost of the survey. A certain number of counts remains necessary, as a tool to double-check mobile data, especially for frontier points without administrative data sources; and for gathering evidence on specific secondary phenomena, which are difficult or impossible to capture with MPD, such as travellers from smaller countries without a resident MNO (like San Marino, Monaco, Andorra).

The main benefits of adopting MPD include:

- **Road passes:** Italy has dozens of road passes often located in remote mountainous areas. Inevitably, counting operations must be limited to a sample of relevant crossings, to a partial number of days each month and to a limited daily period usually ranging from 6 to 8 hours. Moreover, counting operations on roads in the absence of checkpoints⁴ cannot go beyond counting vehicles and the number of occupants, using the license plate as a proxy for residency status. Furthermore, information about road checkpoints is largely missing from administrative data sources.

On the contrary, MPD can potentially cover all road checkpoints and monitor them continuously. Therefore, an advantage is expected in terms of a better estimation both of the universe size and of traveller's residency distribution. Indeed, for residency determination, counting captures data for every *n*th individual, whereas MPD uses the nationality of the SIM card issuer for all detected SIM cards.

- **Port crossing points:** counting operations are conducted in boarding and disembarkation areas, which are subject to access restrictions for security reasons.

MPD can monitor ports without the need to access restricted areas and with equal effectiveness during both the boarding and disembarkation phases. Considering that only a minority of ports provide administrative data, MPD should lead to an improvement in both the number estimation and the qualification⁵ of the universe of international travellers.

⁴ Italy has free movement agreements with its neighbouring countries due to its membership in the Schengen Area.

⁵ Qualification provides the breakdown by country of origin for non-resident travellers and by country of destination for residents in Italy.

- **Airport crossing points:** all Italian international airports collect administrative data and, consequently, the size of the reference universe for airborne travellers can be considered as known. However, the administrative data do not provide any information on the residence of travellers.

The use of MPD should lead to a better qualification of the universe, especially for large airports characterized by numerous and distant gates, where counting operations, given the number of counting staff, can only be carried on for a subset of the departure and arrival areas.

- **Railroad crossing points:** counting operations are conducted directly on sampled trains. Administrative data is only available for international routes operated by the main railway operator, again with no information about the country of residence of the travellers.

The advantage of MPD in this case lies mainly in the fact that telephone data make it possible to monitor and qualify all trains crossing borders and not just a sample of them.

In addition to the advantages mentioned above, it was requested that the MNO ensure the MPD datasets have a structure similar to those produced by the counting operations. Nonetheless, unlike counting operations, that collect different types of information based on the mode of transport (e.g. the type of vehicle boarded on a ship, the class of a train carriage), MPD datasets maintain a consistent structure across all border crossings. This results in the loss of secondary information but increases the reliability of relevant information.

Ultimately, MPD present multiple advantages compared to traditional counting operations, including a better coverage and qualification of the universe of international travellers, with the sole exception of the loss of certain secondary information, not relevant for BoP statistical purposes.

4. The use of MPD by the Bank of Italy for compiling the Travel item of the Balance of Payments

4.1 The MPD used and the main characteristics of the MNO

MPD are currently provided to the Bank of Italy by Vodafone⁶, the same MNO that supplied data during the experimental phase and the pandemic period.

The information is derived from the data traffic that each SIM card, limited to human-type⁷ ones, continuously maintains with the operator's network. In short, when a SIM card issued in Italy disappears from the network near a border, it might indicate a resident travelling abroad. Conversely, when a SIM card issued by a foreign operator appears near a border, it might indicate a non-resident entering Italy (see Section 4.4 for details).

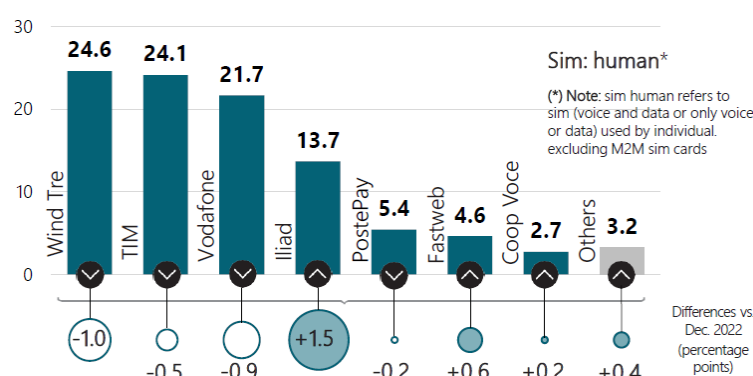
⁶ The data processing is performed by the company Motion Analytica S.r.l..

⁷ Human SIM cards are cards designed for use by human beings. It is the most common type of SIM card, and it is mainly used to connect mobile phones and other devices to cellular networks. On the other hand, machine-to-machine (M2M) SIM cards are used for communication between devices.

The Operator uses two main methods to extrapolate the data to the entire population (total SIM cards detected on Italian mobile networks): for Italian SIM cards, it relies on its market share; for foreign SIM cards, the proportion of users is determined by considering the likelihood of their devices connecting to the network of the mobile carrier in the specific cell near the border. Therefore, the market share of the operator is crucial to ensure an adequate expansion of the findings to the entire universe. In September 2023, the share of total SIM cards issued by Vodafone was 21.7% of the total human SIM cards issued in Italy (Figure 1).

The data are compliant with the Italian privacy law. With specific reference to the data transmitted to the Bank of Italy, they are irreversibly anonymized and aggregated into subsets that cannot be formed by less than 15 SIM cards. If this is not the case, the information is not provided.

Figure 1. Human SIM cards issued in Italy: market shares of the main Italian MNO.



Source: AGCOM⁸.

4.2 How to make the most of MPD

In order to use MPD for tourism statistics, it is necessary that the MNO be able to isolate the SIM cards and the information that contribute to the formation of the conceptual aggregates spelled by BoP compilation manuals. Listed below are the main aspects that were worked upon, in collaboration with the MNO, in this respect.

Identification of the destination country for Italians and of the origin country for foreigners

As already mentioned, an Italian resident travelling abroad is detected by a SIM card issued in Italy disappearing from the network near a border. To determine the destination country (or countries), the "Welcome message" received by each SIM card when connecting to a foreign operator's network is used.

A non-resident entering Italy is instead identified by a SIM card issued by a foreign operator appearing on the Italian territory near a border point. For the country of origin, the residence of the SIM card issuer is used as a proxy. This approximation

⁸ The Italian Regulatory Authority for Communications.

is of good quality in most cases. However, certain specific phenomena can be systematically distorted by this approach. A typical case is that of individuals who are resident⁹ in Italy but come from other countries and use a SIM card issued in their home country. In Italy, there are numerous and large communities of foreign residents, so this phenomenon cannot be overlooked. To correct this type of false-positive cases, an analysis of the SIM cards' behaviour over several months is required: if a SIM card issued abroad is detected on the Italian network for at least ninety days (not necessarily consecutive) within the twelve months preceding the month under examination, then the SIM card is associated with a resident in Italy.

Estimating Overnight and Same-Day Trips

In travel estimation it is crucial to distinguish between overnight and same-day travellers, as their spending profiles are dramatically different.

To associate a SIM card issued by a non-resident MNO with an overnight trip to Italy, it is necessary that at least 24 hours elapse between the time of entry into the Italian territory and the time of exit. SIM cards with a detection period of at least 6 hours and an exit from the Italian territory detected the day after the day of entry are also associated with overnight travellers.

For SIM cards issued by resident MNOs, SIM cards for which the signal is lost for at least 24 hours are considered to be associated with overnight travellers abroad. The 6-hour criterion is not applied as, due to the characteristics of certain Italian cross-border worker flows, this filter seems to lead to an overestimation of overnight travellers.

Determining the Province Visited or of Residence

Determining the Italian province visited (by inbound travellers) or of residence (by outbound Italian travellers) is not relevant for BoP compilation, but is extremely useful for tourism analysis.

As for determining the province visited by non-resident travellers, the "prevalent province" is defined as the one where the travellers spent the most time during their trip to Italy.

For Italian residents, the "residence" is defined as the municipality that most frequently in the previous twelve months has been the "prevalent night-time location", which in turn is defined as the municipality where the user spent the most time between 8 p.m. of the previous day and 8 a.m.. Although, MNOs collect information at a more detailed territorial level than the province, the data is then aggregated to this last dimension for two reasons: a) to contain the number of records that need to be masked in compliance with privacy regulations (see section 4.1); b) the impossibility, on the part of the MNO's network, to reconstruct with precision the administrative subdivision into municipalities, particularly for smaller municipalities.

The roaming effect near land borders

⁹ The concept of residence used here is economic rather than administrative, as outlined in the Balance of Payments and International Investment Position Manual Sixth Edition (BPM6) of the IMF. In this document, "The residence of each institutional unit is the economic territory with which it has the strongest connection, expressed as its centre of predominant economic interest."

A relevant problem concerns data roaming near land borders, where an Italian SIM card could connect to a foreign network, or vice versa, without an actual border crossing, thereby generating a false positive.

In order to avoid this type of error, specific "filters" were incorporated into the MNO's algorithms for monitoring land borders.

A first filter concerns the minimum duration for which a signal must be acquired (foreigners in Italy) or lost (Italians abroad) on the MNO's network. This minimum duration has been set to 30 minutes. Only if the signal exceeds this threshold, the SIM card movement is considered indicative of an actual border crossing.

The filter shows good effectiveness in most scenarios but further analysis is required for borders near urban centres, where a 30-minute threshold may prove to be inadequate and a shorter interval may be required. To address this issue, the possibility of introducing time filters differentiated according to the distance of the borders from residential areas is currently being considered.

A second filter concerns the daily frequency with which a SIM card connects to or disconnects from the MNO's network. If the frequency is high, it is assumed that the SIM card is providing wrong signals due to prolonged presence in areas near the border. The frequency filter prevails over the temporal one.

Finally, the third filter tracks the movements of SIM cards through virtual "corridors" of adjacent telephone cells, defined by the operator for each border crossing¹⁰. This applies to both foreign SIM cards connecting to the Italian network and Italian SIM cards disconnecting from it. If a SIM card progressively connects to cells further from the borders for foreigners or closer for Italians, it is considered a true positive, with priority over the previous filters.

Identifying arrivals and departures at airports

In order to identify arrivals and departures related to international flights and to exclude numerous other cases (like domestic travellers, airport personnel, etc.), it was necessary to define, in close collaboration between the Bank of Italy and the MNO, a number of criteria.

A set of consecutive filters allows each SIM card to be assigned a specific status, as summarized in the scheme below.

1. A SIM card was detected in the area adjacent to the airport, then inside the airport, and finally disappeared from the network:
 - a) If the SIM card is not detected again in the following hours, it is associated with a traveller (Italian or foreign) departing for an international destination.
 - b) If the SIM card is detected again in the following hours (at the same or at another Italian airport), the data is discarded.
2. A SIM card not present on the network "appeared" inside the airport and then was detected in the adjacent area: the SIM card is associated with an incoming (or returning) international traveller.

¹⁰ For instance, a virtual corridor may be designed so as to track the route of a motorway at a road border.

3. A SIM card not present on the network "appeared" inside the airport and then disappeared again from the network: the SIM card is associated with an international (resident or not resident) traveller in transit.

The described methodology, progressively refined over time, is able to provide estimates on the number of travellers for each airport that differ only slightly from the official administrative data.

5. Quality assessment of MPD

The quality assessment of MPD can be achieved through comparison with administrative data. In the absence of administrative data, survey data are used.

When discrepancies arise between MPD and administrative data for a specific border crossing, MPD are considered incorrect, and a revision process of the identification algorithms summarized in the previous section is initiated and repeated until the estimates derived from the MPD show a satisfactory convergence with the administrative data. A deviation of the estimate of less than 10 per cent is considered acceptable. This constraint is relaxed for border crossings with very low traveller flows, where percentage differences lose informational value. Once satisfactory estimates are obtained for the revised border crossing, they are closely monitored for several months to ensure that the convergence of the estimates towards correct values is genuine and not due to random effects.

In the absence of administrative data, the comparison is made with data from the survey. For each border crossing, MPD are considered of good quality if they approximately replicate the volumes and residence distributions of the survey data, without introducing predefined quantitative thresholds. If sufficient approximation is not achieved a dedicated counting session at the border crossing of approximately 12-18 hours is conducted, with the aim of gathering the most accurate information possible. The results are considered "correct" and are then compared with the MPD collected at the same border crossing during the same time frame to validate or potentially refine the MPD production algorithm.

An additional quality check for the MPD involves analysing data over a broader time frame than a single month. Over such a period, it is expected that the number of incoming foreign SIM cards will roughly equal the number of foreign SIM cards departing from Italy, once all border crossing points are considered (a similar and reciprocal expectation applies to SIM cards issued in Italy). This is based on the assumption that a non-resident visiting Italy will leave within a few days, weeks, or, in rarer cases, months. Nevertheless, monitoring individual border crossings has its validity, as significant imbalances between incoming and outgoing flows could indicate potential measurement issues.

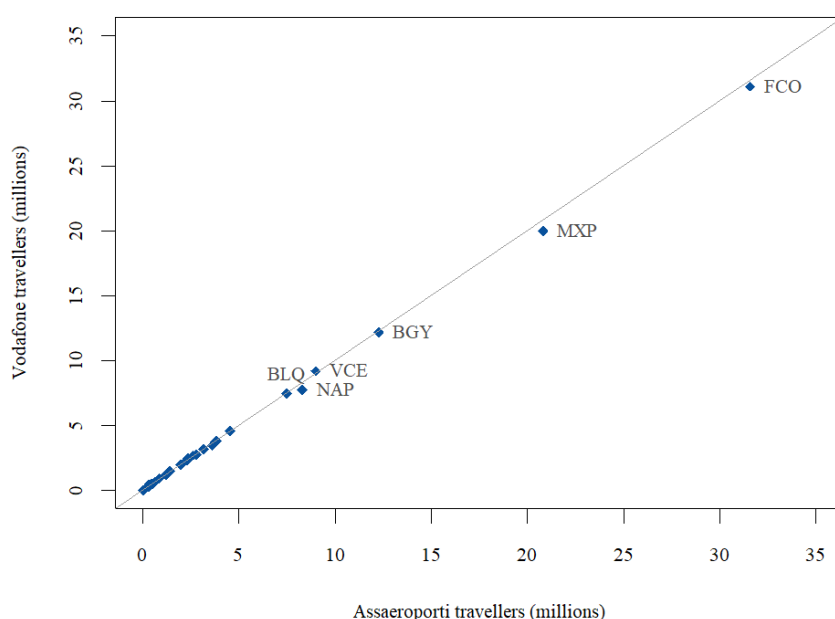
5.1 MPD Quality in 2023

Vodafone data and administrative data on the sampled airports¹¹ (source: Assaeroporti) are compared for the year 2023.

Administrative data provide information on the number of travellers on international flights, both inbound and outbound, with no distinction by country of residence. The comparable telephony aggregate includes the set of SIM cards detected entering and exiting the airports, obtained by adding both those issued in Italy and those issued abroad¹².

The number of international travellers recorded by the two sources for individual airports is very similar (Figure 3). The two major airports, Rome Fiumicino and Milan Malpensa, stand out as the largest, with the highest volumes of international travellers. In addition, there is a group of medium-sized airports, including Bologna, Naples, Bergamo, and Venice, each handling a significant number of international travellers, ranging from 7.4 to 12.4 million. The remaining 23 airports are classified as small, each managing fewer than 5 million international travellers.

Figure 3. Comparison of Vodafone and Assaeroporti data for all sampled airports (2023).



Source: MPD and Assaeroporti. See the Appendix for the list of IATA codes.

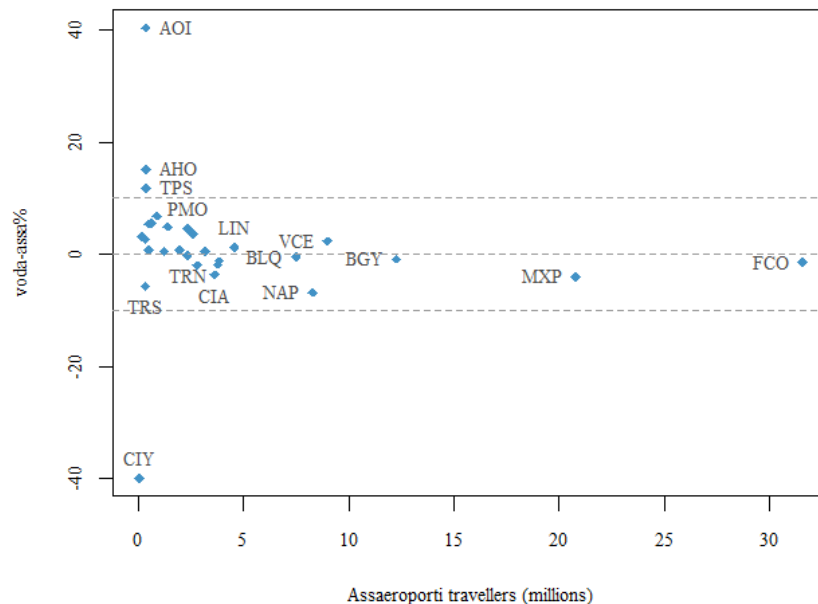
Deviations between the two sources in percentage terms (Figure 4) are almost always below 10 percent. The few cases where the deviation exceeds this threshold

¹¹ See the Appendix.

¹² These aggregates differ from that used in the BoP statistics, where only inbound SIM cards of resident MNOs and outbound SIM cards of non-resident MNOs have to be recorded.

refer to smaller airports with a very limited flow of international travellers for which, as already noted, percentage-based analysis loses its informative capacity.

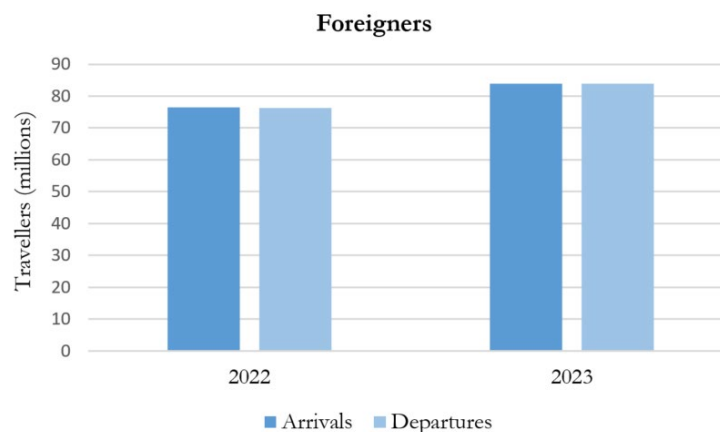
Figure 4. Percentage differences between Assaeroporti and Vodafone data (2023).

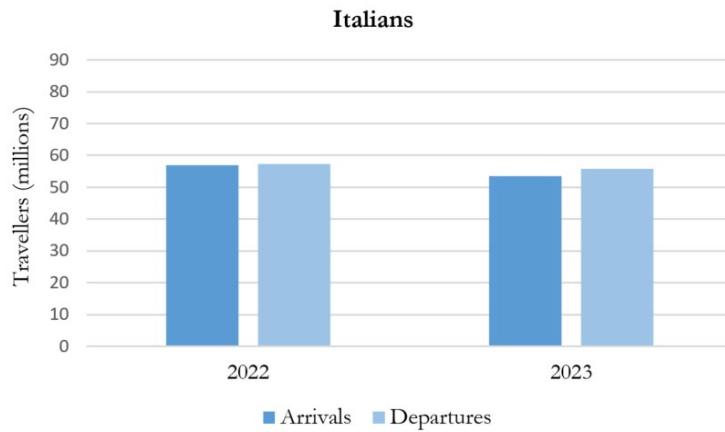


Source: MPD and Assaeroporti. See the Appendix for the list of IATA codes.

We next go back to the internal coherence reasoning that, over an extended period, the number of incoming foreign SIM cards should roughly equal the number of foreign SIM cards departing from Italy (and similarly for Italian SIM cards). The comparison of incoming and outgoing international traveller flows detected by MPD, separately for Italians and foreigners, shows a good similarity between the total values for both 2022 and 2023 (Figure 5). In particular, for foreigners, the number of arrivals is essentially equal to the number of departures in both years; a similar pattern holds for Italians in 2022, while for 2023 there is a difference of about 4 percent between the two flows, which is still modest.

Figure 5. Comparison between the number of international arrivals and departures (millions) for Italians residents and non-residents foreigners in 2022 and 2023.





Source: MPD.

6. CONCLUSIONS

Traditional tourism statistics have historically relied on surveys, such as the Bank of Italy's Survey on International Tourism, which faced limitations in coverage and adaptability due to the Schengen agreements and during emergencies like the COVID-19 pandemic.

To address these challenges, the Bank of Italy began integrating MPD into the BoP estimation process in 2020, replacing part of traditional counting methods. This shift has brought significant cost savings and improved data accuracy, particularly in complex border areas lacking administrative coverage. However, the adaptation of MPD for tourism statistics requires a close and continuous collaboration with the MNO in order to ensure alignment with official definitions and standards, highlighting ongoing challenges for methodological development and for data-quality assurance.

APPENDIX

Correspondence table for sampled airports: airports' names – IATA codes.

Name	IATA code	Nb. of travellers ¹³
Alghero	AHO	370,648
Ancona	AOI	354,941
Bari	BRI	3,185,392
Bergamo	BGY	12,275,947
Bologna	BLQ	7,509,198
Brindisi	BDS	880,806
Cagliari	CAG	1,241,537
Catania	CTA	3,853,279
Comiso	CIY	57,984
Firenze	FLR	2,631,538
Genova	GOA	512,734
Lamezia Terme	SUF	638,646
Milano Linate	LIN	4,566,040
Milano Malpensa	MXP	20,816,153
Napoli	NAP	8,311,651
Olbia	OLB	1,423,676
Palermo	PMO	2,375,718
Perugia	PEG	349,907
Pescara	PSR	494,082
Pisa	PSA	3,796,032
Rimini	RMI	212,347
Roma Ciampino	CIA	3,630,295
Roma Fiumicino	FCO	31,581,910
Torino	TRN	2,337,446
Trapani	TPS	400,965
Treviso	TSF	2,819,655
Trieste	TRS	346,462
Venezia	VCE	9,002,064
Verona	VRN	1,987,198

¹³ Source: Assaeroporti (2023).

REFERENCES

1. R. Ahas, A. Aasa, A. Roose, Ü. Mark and S. Silm, 'Evaluating passive mobile positioning data for tourism surveys: An Estonian case study', *Tourism Management*, 29(3), 2008, pp. 469-486.
2. R. Ahas, A. Aasa, Ü. Mark, T. Pae, A. Kull, 'Seasonal tourism spaces in Estonia: Case study with mobile positioning data', *Tourism Management*, 28(3), 2007, pp. 898-910, ISSN 0261-5177.
3. R. Ahas, S. Silm, O. Järv, E. Saluveer, and M. Tiru, 'Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones', *Journal of Urban Technology*, 17, 2010, pp. 3-27.
4. A. Carboni, C. Doria, C. Catalano, 'How can Big Data improve the quality of tourism statistics? The Bank of Italy's experience in compiling the "travel" item in the Balance of Payment', *Questioni di Economia e Finanza (Occasional Papers)*, Bank of Italy, 761, 2023.
5. A. Carboni, C. Doria, S. Zappa, 'La produzione statistica nell'emergenza COVID19: la stima dei "viaggi" in bilancia dei pagamenti', *Note Covid-19*, Bank of Italy, 19th June 2020.
6. Q. Chen, Z. Hu, H. Su, X. Tang and K. Yu, 'Understanding Travel Patterns of Tourists from Mobile Phone Data: A Case Study in Hainan', *IEEE International Conference on Big Data and Smart Computing*, 2018
7. G. Cousin, F. Hillaireau, 'Can Mobile Phone Data Improve the Measurement of International Tourism in France?', *Economie et Statistique / Economics and Statistics*, Institut National de la Statistique et des Etudes Economiques (INSEE), (505-506), 2018, pp. 89-107.
8. C. Demunter, 'Tourism Statistics: Early Adopters Of Big Data?', *Sixth UNWTO International Conference on Tourism Statistics*, 2017.
9. Eesti Pank (2019). International travel statistics. Retrieved 12 June 2019, from http://statistika.eestipank.ee/#/en/p/MAKSEBIL_JA_INVPOS/1410.
10. Eurostat. (2014) Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics: Consolidated Report. LU: Publications Office.
11. L. Grassini, G. Dugheri, 'Mobile phone data and tourism statistics: a broken promise?' *Annals of Mathematics and Artificial Intelligence*, 3(1), 2021, pp. 50-68.
12. T. Kubo, S. Uryu, H. Yamano, T. Tsuge, T. Yamakita, Y. Shirayama, 'Mobile phone network data reveal nationwide economic value of coastal tourism under climate change', *Tourism Management*, 77, 2020, 104010, ISSN 0261-5177.
13. Y. Leng, A. Noriega, A.S. Pentland, I. Winder, N. Lutz and L. Alonso, 'Analysis of Tourism Dynamics and Special Events through Mobile Phone Metadata', *Proceedings of Data for Good Exchange (D4GX) 2016*. New York, NY

14. T. Lestari, R. Rufiadi, E. Saluveer, S. Esko, S. Dimulyo, 'Indonesia's Experience of using Signaling Mobile Positioning Data for Official Tourism Statistics', 2018.
15. I. Noviyanti, et al. 'Towards Big Data as Official Statistics: Case Study of the Use of Mobile Positioning Data to Delineate Metropolitan Areas in Indonesia', 2020: 943 – 954.
16. S. Park, Y. Xu, L. Jiang, Z. Chen, S. Huang, 'Spatial structures of tourism destinations: A trajectory data mining approach leveraging mobile big data', *Annals of Tourism Research*, 84, 2020, 102973, ISSN 0160-7383.
17. S. Phithakkitnukoon, T. Horanont, A. Witayangkurn, R. Siri, Y. Sekimoto, R. Shibasaki, 'Understanding tourist behavior using large-scale mobile sensing approach: A case study of mobile phone users in Japan', *Pervasive and Mobile Computing*, 18, 2015, pp. 18-39, ISSN 1574-1192.
18. E. Rahmadian, Y. Virantina, D. Feitosa, A. Zwitter, A., T. K. Lestari, 'Big Data Governance: The Case of Mobile Positioning Data for Official Tourism Statistics In Indonesia', In *The 64th World Statistics Congress*, 2023.
19. E. Saluveer, J. Raun, M. Tiru, L. Altin, J. Kroon, T. Snitsarenko, A. Aasa, S. Silm, 'Methodological framework for producing national tourism statistics from mobile positioning data', *Annals of Tourism Research*, 81, 2020, 102895, ISSN 0160-7383.
20. UN.ESCAP, 'Using big data for official statistics: key considerations when using mobile phone data', 2022.
21. UN, 'Handbook on the Use of Mobile Phone Data for Official Statistics. Global Working Group on Big Data for Official Statistic', 2019.
22. Y. Xu, J. Xue, S. Park, Y. Yue, 'Towards a multidimensional view of tourist mobility patterns in cities: A mobile phone data perspective', *Computers, Environment and Urban Systems*, 86, 2021, 101593, ISSN 0198-9715.

Mobile phone data and the "Travel" item in the Balance of Payments: the Bank of Italy's experience

Andrea Carboni, Claudio Doria, **Valeria Valletta – Bank of Italy**
Twelfth IFC Conference - “Statistics and beyond: new data for decision making in central banks”
BIS Basel, 22 and 23 August 2024

Before and after MPD introduction

1996-2020: Survey on International Tourism

1.200.000 counting op.

estimation of the size of the reference universe

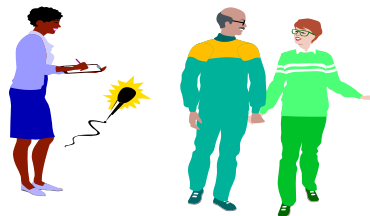
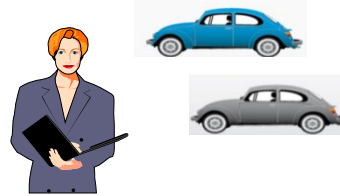
state of residence / state visited.

Administrative data

great precision on the overall number of travelers
no information on their residence or destination.

110.000 interviews

information on expenses, characteristics of the traveler, characteristics of the trip, etc...



After 2020: MPD replaced most of the counting op.

427.000 counting op.

Checking purposes, information on overlapping borders.

Mobile phone data

Country of residence (proxied by the company issuing the SIM) and number of travelers

MPD from Vodafone:

- 21,7% market share (2023), covers almost 99% of the Italian territory, good spatial and time resolution, data are compliant with Italian privacy legislation

Similar structure of MPD and datasets from counting operations – loss of secondary information

Administrative data

110.000 interviews

How MPD can improve the survey

Reduction of the total cost

MPD replaced a large part of the counting operations, that are very expensive.

Resilience to wide shocks

Due to Covid-19 movement restrictions, the survey had to be interrupted for several months. MPD were not affected by the pandemic.

Better estimation of universe size and residency distribution

Road: mountainous areas, lack of checkpoints, no administrative data. MPD monitor all crossings simultaneously in the continuous.

Seaports: access restrictions to boarding and disembarkation areas.

MPD monitor seaports without entering restricted areas. Only a minority of ports provide administrative data.

Airports: airports usually have good administrative data, MPD helps refine estimates of travellers' residency distribution, especially in large, complex airports with many distant gates.

From telephony terminology to tourism's definitions

Italian resident travelling abroad = SIM card issued in Italy disappearing from the network near a border.

Destination country (or countries) → Welcome message.

Non-resident entering Italy = SIM card issued by a foreign operator appearing on the network near a border.

Origin country → the residence of the SIM card issuer is used as a proxy.



Foreigners resident in Italy who use a SIM card issued in their home country.

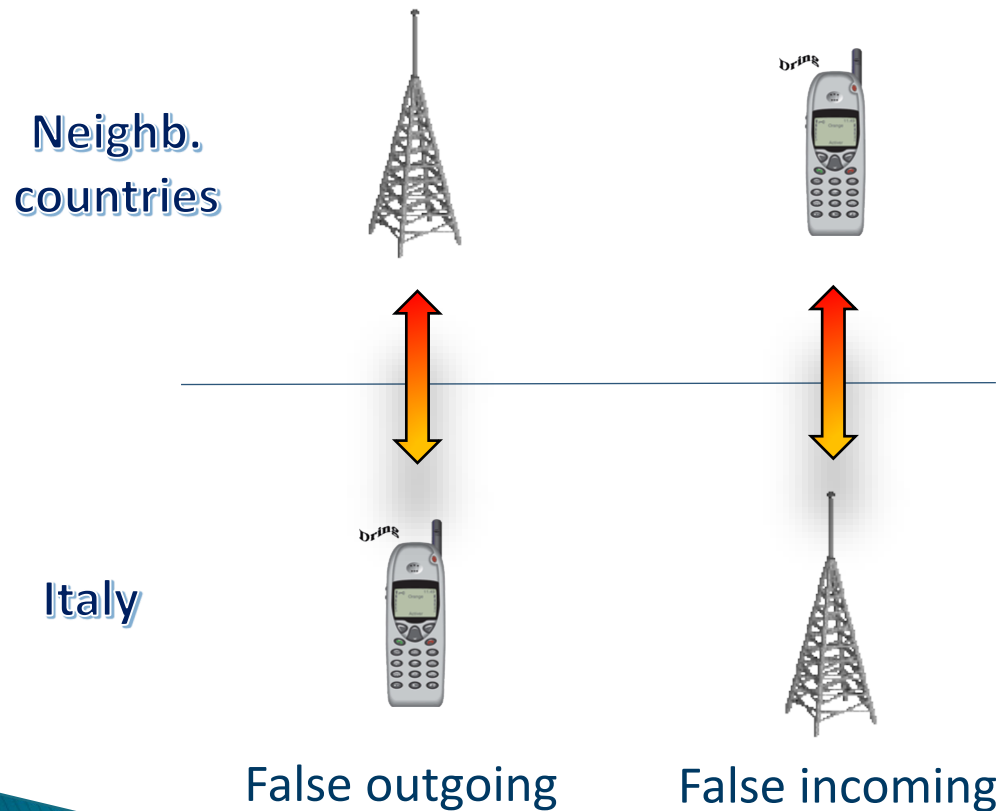
FALSE POSITIVE

if a SIM card issued abroad is detected on the Italian network for at least nine out of the twelve months preceding the month under examination, then the SIM card is associated with a resident in Italy.

From telephony terminology to tourism's definitions: roaming problem

Problem

In areas close to land borders a SIM could connect to a Network abroad without crossing the border.



Solution

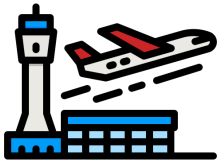
Time: a minimum amount of time (30 mins) is required to identify a SIM that has crossed the border:

- Non Italian SIMs: must be detected on the Vodafone network for at least 30 mins
- Italian SIMs: must disappear from Vodafone network for at least 30 mins.

Frequency: if a SIM connects and disconnects a large nb. of times in a single day, it is considered a false positive.

Corridors: if a foreign SIM connects to adjacent cells progressively more distant from the borders it is considered a true positive (conversely for Italians going abroad).

From telephony terminology to tourism's definitions



Arrivals and departures at airports

Departing SIM: the SIM must be detected while approaching the airport area;
then it has to disappear from the mobile network;

It has not to reconnect within a few hours in another Italian area

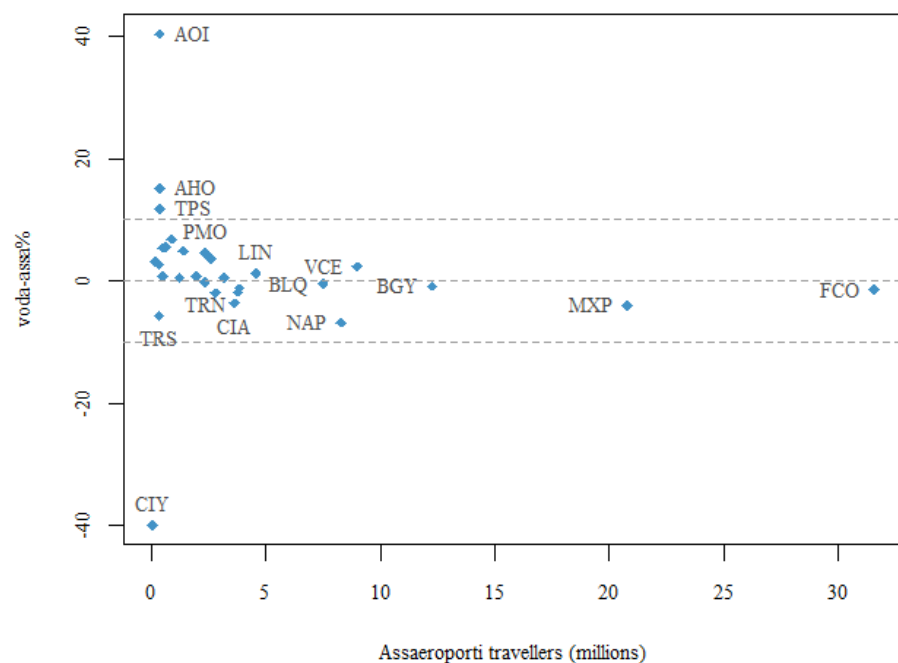
Arriving SIM: the SIM must not be present on the mobile network and must appear in an airport;

- If the SIM moves in areas in the proximity of the airport → arrival
- If the SIM disappears again without leaving the airport → transit

Do MPD match with administrative data?

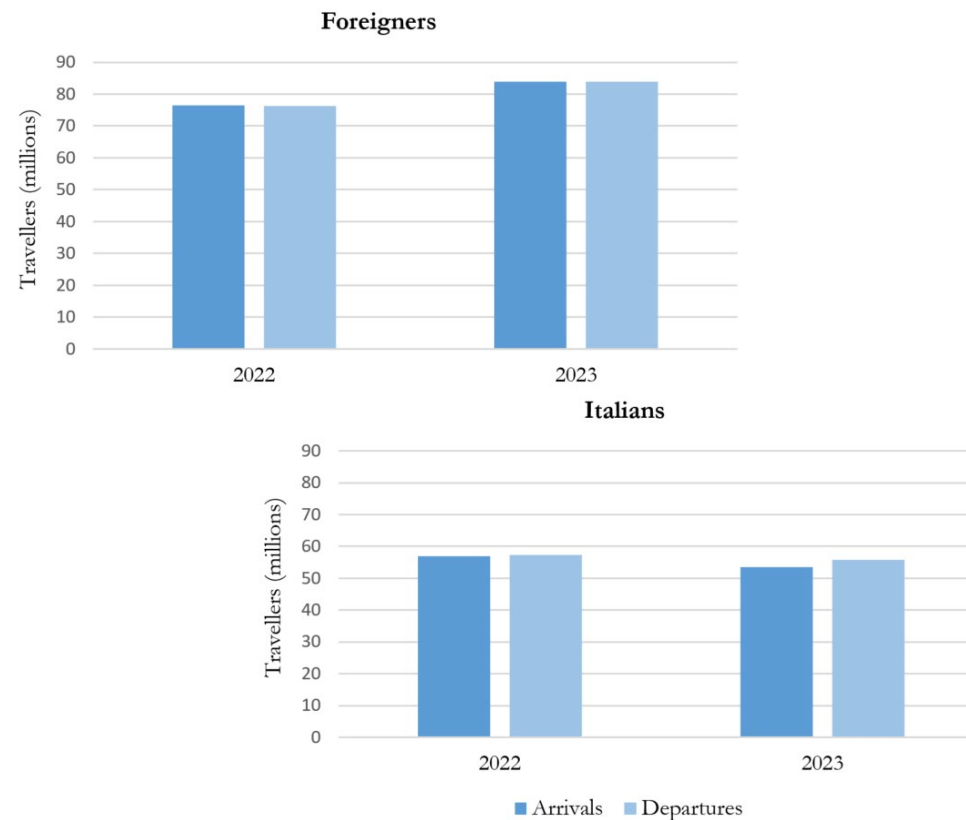
Differences between MPD and administrative data

Airports - 2023



International arrivals and departures (millions)

All crossings (2022, 2023)



Source: Assaeroporti, MPD

Thank you!

valeria.valletta@bancaditalia.it