MIRREM

Measuring Irregular Migration

www.irregularmigration.eu

Exploring the Use of Aggregate Air Passenger Data for Estimating Overstayer Inflows

MIrreM Briefing Paper

AUTHORS:

Luca Bernasconi

Ettore Recchi

Funded by the European Union

Co-funded by: UK Research and Innovation

Toronto Metropolitan University — Canada Excellence Research Chair in Migration & Integration

Deliverable Information:

Project Acronym:	Measuring irregular migration and related policies (MIrreM)
Project No.	101061314
WP	WP6 - Methods Innovation Lab
Deliverable Type:	Briefing Paper
Deliverable Name	D6.2 PS1 - Exploring the use of aggregate air passenger data for estimating overstayer inflows
Version:	1
Date:	07/04/2025
Responsible Partner:	University of Potsdam (UP) and European University Institute (EUI)
Contributing Partners:	University for Continuing Education Krems (UWK)
Authors	Luca Bernasconi (EUI) and Ettore Recchi (EUI-Sciences Po)
Reviewers:	Alejandra Rodríguez-Sánchez (UP), Jasper Dag Tjaden (UP), Jill Ahrens (UWK) and Albert Kraler (UWK)
Dissemination Level:	Public

Revision History:

Revision	Date	Author	Organisation	Description
1	22/09/2024	Luca Bernasconi and Ettore Recchi	EUI	Initial draft
1	30/09/2024	Alejandra Rodríguez Sánchez	UP	Reviewer feedback
1	02/10/2024	Jasper Dag Tjaden	UP	Reviewer feedback
1	14/10/2024	Luca Bernasconi and Ettore Recchi	EUI	Revised version
1	16/12/2024	Jasper Dag Tjaden	UP	Version 1
1	18/12/2024	Jill Ahrens	UWK	Review and Formatting check
1	02/03/2025	Albert Kraler	UWK	Review
1	02/04/2025	Adriana Harm	UWK	Formatting check
1	07/04/2025	Alejandra Rodríguez Sánchez	UP	Published version 1

Executive Summary

Overstayers —individuals entering a country with a valid visa but staying beyond the initially authorized period—are a blind spot of migration research. Past studies claim that overstaying is indeed the main gateway to irregular migration in Europe, but few estimates exist. This paper explores a new method to estimate the inflows of overstayers arriving via airline travel. We concentrate on the Schengen area as destination and on non-European countries as origins since commercial flights are the most likely transportation mode to Europe for travellers from these countries. Our method relies on aggregate information on incoming and departing passengers from all airports in the Schengen Area. First, we compute 'net air travel flows' as the difference between incoming and outgoing passengers. Second, we subtract regular flows. We rely on net migration figures from the QuantMig project to produce an estimate of overstay inflows through air travel in the year 2019, disaggregated by region of origin. The briefing discusses initial findings, advantages and limitation of this this approach.

Table of contents

LIST OF TABLES

LIST OF FIGURES

THE MIRREM PROJECT

MIrreM examines estimates and statistical indicators on the irregular migrant population in Europe as well as related policies, including the regularisation of migrants in irregular situations.

MIrreM analyses policies defining migrant irregularity, stakeholders' data needs and usage, and assesses existing estimates and statistical indicators on irregular migration in the countries under study and at the EU level. Using several coordinated pilots, the project develops new and innovative methods for measuring irregular migration and explores if and how these instruments can be applied in other socio-economic or institutional contexts. Based on a broad mapping of regularisation practices in the EU as well as detailed case studies, MIrreM will develop 'regularisation scenarios' to better understand conditions under which regularisation should be considered as a policy option. Together with expert groups that will be set up on irregular migration data and regularisation, respectively, the project will synthesise findings into a Handbook on data on irregular migration and a Handbook on pathways out of irregularity. The project's research covers 20 countries, including 12 EU countries and the United Kingdom.

TO CITE:

Bernasconi, L. & Recchi, E. 2024. *Exploring The Use Of Aggregate Air Passenger Data For Estimating Overstayer Inflows.* MIrreM Briefing Paper, Krems: University for Continuing Education Krems (Danube University Krems). DOI: https://doi.org/10.5281/zenodo.14809013

KEYWORDS

Irregular Migration; Statistics; Migration Flows; Visa Overstaying; Digital Sources; Flight passenger; Overstayers.

FUNDING ACKNOWLEDGEMENT

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

In addition, MIrreM benefit from funding provided by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee. The Canadian research component of this project is undertaken, in part, thanks to funding from the Canada Excellence Research Chairs Program of the Government of Canada.

1. INTRODUCTION

Visa overstayers—individuals entering a country with a valid visa but staying beyond the initially authorized short-term period—are a largely understudied population. Following one pioneer from 2009 (Clandestino, 2009; Triandafyllidou, 2010), overstaying is frequently cited as the primary form of irregular migration in the Global North (de Haas, 2023, p. 35; see also Siruno et al., 2024). Within Schengen countries, the Italian Home Office estimated that up to 75% of the stock of unauthorized migrants was constituted by overstayers in the early 2000s—but no clear methodological breakdown of the calculations was ever provided to back this statement (Fasani, 2008; Jandl, 2008). Finotelli and Sciortino (2013) go as far as to posit that visa overstaying is the "main elephant in the room" of the European migration debate. While governments have grown gradually wary of overstaying (European Commission, 2016), targeting it with increased surveillance (Hansen and Pettersson, 2021), evidence on the magnitude, trajectories, and drivers of the phenomenon remains scarce.

To further understand and measure overstaying, non-intrusive methodologies must be devised, using novel sources of data without breaching ethical guidelines. Leveraging the fact that a large portion of transnational travel is undertaken through commercial flights (Recchi et al., 2019), this briefing paper explores the potential of (aggregate) air travel flow data in estimating visa overstaying, by means of estimation resembling entry/exit systems (Rodríguez Sánchez and Tjaden, 2023; Jandl, 2008).

The aim of the paper is threefold. First and foremost, to describe a systematic methodology to generate estimates of visa overstaying flows through air travel. Second, to apply the method as a proof of concept and identify countries exhibiting larger flows to the Schengen Area. And last, to assess the strengths and limitations of the method itself.

2. CONCEPTS & DEFINITIONS

Obtaining migration flow data in general is complex (Willekens, 2019), let alone for irregular migration. Several reports find that methodologies in estimating irregular migration flows lag those that measure stocks (Siruno et al., 2024), and that these difficulties are only exacerbated by the issue of overstaying. While flow measurement of irregular migrants is often concentrated on border apprehensions (Siruno et al., 2024; Vespe et al., 2017), overstaying cannot be gauged in this way.

The most reliable measurement methodologies involve some form of individualized entry/exit data, wherein individuals are longitudinally tracked to see if they have left the territory. Older methodologies implemented by state authorities involved experimentations with double-entry card systems by which entry and exit records are matched (Jandl, 2008), but this was facilitated by the geographical situation (that is, insularity) of the countries involved (Australia and Japan). Technical advancements in biometric and biographic verification have allowed the emergence of databases that monitor compliance with entry/exit visas (Rodríguez Sánchez and Tjaden, 2023). On this basis, the European Commission has announced a future implementation of its own Entry/Exit System (EES) (Vespe, 2017). However, besides having yet to be rolled out in Europe, this method is not readily scalable, and can potentially be highly intrusive, thereby presenting ethical challenges (Rodríguez Sánchez and Tjaden, 2023). Moreover, even in the United States—the sole country to have implemented a biometric system and shared overstaying figures (Baker and Warren, 2024)—estimates have been shown to be erroneous on occasion (Warren, 2017).

Given these caveats, we try to find an aggregate-level methodology that is accessible to researchers and scalable across geographies. Focusing on the Schengen Area—the territory of interest to our study—Vespe et al. (2017) propose tourist visa issuances as a "significantly approximated" upper bound on visa overstaying: to become an "overstayer" you have to be a regular "stayer" first. This, however, is highly unreliable for several reasons, including the omission of individuals arriving from countries with visa-free travel agreements with the EU. Using aggregate-level—non-individualized flight statistics—we provide an accessible and more reliable alternative to the abovementioned type of systems. We lay out our methodology in the subsequent section, before critically assessing the limitations and evaluating its ethics, appraising it using established benchmarks.

Before presenting the method in detail, it is important to situate the population flows we are trying to measure within the extant theoretical literature on irregular migration. We follow the MIrreM taxonomy of irregular migration (Kraler and Ahrens, 2023), which builds upon previous scholarship. This classification scheme aims to provide a measurement framework at a specific point in time in a particular geographical unit of interest. Specifically, it links various inflows or *pathways into irregularity* with stocks of irregular migrants. In turn, inflows

are classified in a tripartite scheme: *demographic flows*, *geographic flows* and *status-related flows*. Overstaying—the object of the current paper—is identified as part of the latter category, since entering the territory and staying beyond the permitted time frame (typically 90 days for a tourist visa in the Schengen Area) involves a loss of legal status.[1]

[1] It is important to stress that the scope of this report is strictly focused on inflows: as such, we do not account for status-related outflows such as regularizations, since this is not an individual-based longitudinal study.

3. METHODS AND DATA

3.1 METHODS

The aim of the method presented in this paper is to estimate the size of the inflows of visa overstayers using novel macro-mobility data related to air travel. The method is inspired by entry-exit systems but uses a novel data source: registrations for commercial air travel. Airline reservation systems record digital traces of mobility that constitute a new, global-scale perspective on cross-border mobility (Recchi and Tittel, 2023).

The intuitive ground truth from which the methodology is derived is that net entries into the territory of interest should equal net legal migration, involving short visits between 1 and 90 days, as well as temporary and long-term migration movements. If the numbers do not match, the 'excess' consists of overstayers.

The rationale for the method is the following. We contrast the volume of incoming and outgoing passengers in each time span. Incoming air passengers are either national or non-national residents. National residents are expected to enter the country returning from a trip abroad; thus, their incoming and outgoing flows cancel out during a reasonable time frame (say, one year). Non-national residents, in turn, can enter with a short-term visa (or visa waiver agreement) or a long-term visa. Short-term visas and visa waiver agreements authorize 90 day stays in the Schengen area (European Commission, 2014). Thus, short-term visa passengers should also cancel out in 90 days: one trip in, one trip out[2]. What remains from the incoming-outgoing difference are either long-term visa holders (i.e. newly arriving legal immigrants) or overstayers. Once we deduce legal immigrants, the possible 'excess' number can only include newly arriving visa overstayers.

Analytically, our approach consists of two steps. First, we calculate the net flow of commercial airline passengers into the Schengen Area. We then subtract net migration figures from this quantity. More precisely, net entries, or *net air travel flows (NF)*, are the difference between inbound and outbound travel between the Schengen Area (*S*) and some third country (*O*) in time period *t*:

$$NF_{O\to S}(t) = F_{O\to S}(t) - F_{S\to O}(t) = F_{in} - F_{out} \qquad \text{(equation 1)}$$

[2] A possible bias is given by 'mixed-mode travel': the process by which an individual enters (or leaves) the Schengen Area by air and leaves (or enters) by means of another mode of travel. This modality is however very marginal statistically, especially when the destination is as large as the Schengen area.

In a second step, net regular migration (*NM*) is subtracted from the above quantity to find the *unaccounted net flow* (*U*) of passengers in a particular time interval. In absence of overstaying, the following subtraction should amount to zero:

$$U_{O\to S}(t) = NF_{O\to S}(t) - NM_{O\to S}(t) \qquad \text{(equation 2)}$$

A large value of *U* should, in fact, signal overstaying, with legal entries not being met with corresponding exits at an aggregate level for the given time frame. The air travel flow data at hand is monthly, but this may easily be aggregated according to needs as $\sum_{t_i<t} NF_{O\to S}(t_i)$, for it to be matched with yearly official migration data[3].

There are several complicating factors to contend with and account for in applying this methodology, some of which are specific to the geography of interest in this study. For instance, with the Schengen Area being the principal region of focus in the study, non-Schengen EU member states[4], in addition to the United Kingdom, are removed from the analysis due to the prevalence of land travel and the porosity of the borders involved (Deutschmann et al., 2023). Indeed, there is a strong possibility of mixed-mode travel (e.g. arrival by train, departure by air, or vice-versa) and circular tourism across several of these countries (Nemes and Happ, 2024) that is best accounted for by their exclusion from our estimates.

A further issue is the aggregation to yearly data for matching with the datasets necessary to measure net regular migration. To mitigate the cutoff during New Year's Eve holidays and volatility in the air travel data, we smooth the monthly data using a three-month centred moving average. Thus, for a given month *m*, we have

$$NF(m) = \frac{1}{3}\sum_{i=-1}^{1} F_{in,m+i} - \frac{1}{3}\sum_{i=-1}^{1} F_{out,m+i}$$

Further confounding issues are listed as exhaustively as possible in Table A.1, with accompanying potential remedial measures.

3.2 DATA

This study lends particular attention to the Schengen Area, with the years 2019-2023 as its time span of interest for reasons of data availability. The core dataset used—namely, commercial flight passenger data—was acquired from Sabre, a market intelligence firm specialized in the air travel industry. Sabre collects and compiles monthly aggregate passenger statistics based on access to a wide array of sources, including all available booking services, national governments, and aviation federations, among others. It is important to stress that the data at hand is not individualized; instead, monthly flow

[3] In trying to model Brexit-related migration into the European Union (EU), Sîrbu et al. (2024) log-transform the ratio of inflows and outflows, but since we are interested in the absolute value of the flows, we keep the raw differences as they are.

[4] Bulgaria, Cyprus, Ireland, Croatia (pre-2023), Romania, and the microstates.

aggregates are given at a highly granular, dyadic level. After some manipulations described in Appendix B, we obtain a dataset of 291,854 observations for 4,423 dyad-months. In other words, there are 4,423 monthly time series for each origin country ⇔ Schengen dyad, ranging from July 2018 to February 2024[5]. The Schengen Area is considered as a single unit for reasons laid out in the subsequent section.

It is crucial to note that the data is collected by origin and destination (henceforth O&D), rather than single trip legs. This means that, if an individual travels from country A to country B by way of a layover in country C, the trip from A to B is duly recorded in the aggregate statistics. Information about the individual legs, namely A → C and C → B is also available, but it is crucially important for the purpose of the current analysis that A → B be appropriately counted as well. As such, O&Ds are defined as trips for which layovers last no more than 24 hours[6]. This also entails that any stopover of more than 24 hours will break the journey into two separate O&D trips. In other terms, the traveller will be counted as coming from country C rather A.

In addition to the flight data, our analysis includes regular net migration figures. For our purposes and by Eurostat recommendations (Eurostat, 2024), immigration and emigration figures are, in principle, reported by country of previous and next residence, as opposed to nationality. Among a string of other issues, this is important since Schengen nationals departing from or returning to the Schengen Area must be appropriately included in our deductive method, which would not be the case if nationality-based data were to be combined with the flight statistics. There are, however, persistent issues of comparability and completeness in European migration statistics in spite of improvement efforts undertaken by Eurostat (Mooyaart et al., 2021). While Eurostat does publish immigration and emigration datasets organized by country of residence[7], this data is often difficult to compare (Lanzieri, 2018). In general, dyadic international migration data of sound quality is hard to obtain due to missing data, inconsistencies in reporting methodologies between states, and more.

Of the several attempts to improve estimates of migration flows in Europe (see Willekens, 2019 for a review), the Horizon Europe project QuantMig is one of the most comprehensive. QuantMig has sought to address these gaps by using novel Bayesian estimation methods (Aristotelous et al., 2022). We therefore refer to the estimates provided by QuantMig building on Eurostat. This data is dyadic only at the regional level[8]. Furthermore, it is solely available for the year 2019 and earlier. As such, overlap with the Sabre dataset is limited. However, the year 2019 is a sufficient case for a test of our method, as it is also not subject to the travel disruptions of the pandemic period (illustrated in Figure 1).

[5] Croatia is only included starting in 2023, when it formally joined Schengen.

[6] A breaking rule of 6 hours is used for much shorter-distance flights.

[7] *migr_imm5prv* and *migr_emi3nxt*, respectively.

[8] Specifically, it estimates the distributions of migration flows between European countries and defined regions of the world.

4. RESULTS

4.1 NET AIR PASSENGER FLOWS

We begin our analysis by focusing on the net flows of air travellers from non-EU and non-neighbouring countries to the Schengen Area on a dyadic level. Considering the years 2019-2023, discrepancies emerge across Schengen Area countries, with the COVID-19 crystalizing this volatility. Germany, Spain, the Netherlands, Portugal and France—all hubs that constitute major points of arrival into and departure from the EU—feature prominently across the four years, with large net air inflows relative to other Schengen countries (see Table 1).

Table 1: Net air passenger flows to Schengen countries, 2019-2023

Schengen Country	Net flows 2019-2023	Net flows as % of inflows	2019	2020	2021	2022	2023
Germany	1,543,460	1.79%	559,005	381,169	87,219	228,420	287,646
Spain	969,681	2.21%	105,074	10,146	233,353	393,447	227,661
Netherlands	748,341	3.41%	505,564	104,842	-72,570	140,238	70,266
France	669,527	0.98%	180,620	163,447	111,979	88,125	125,356
Portugal	503,606	4.22%	117,688	-10,554	88,451	180,686	127,334
Schengen	5,073,255	1.28%	1,601,736	981,999	626,509	895,842	967,170

The principal regions of origin diverge considerably among these countries with prominent net flow figures. For instance, Eastern Europe and some Asian countries make up for the largest net surpluses to Germany, while South America features disproportionately in the Spanish case.

However, the effects of free movement within the Schengen Area—which, in principle, allows travellers to enter the Union in one country and depart from another—potentially biases the results (as presented from the vantage point of arrival countries). As such, the Schengen Area is best considered as a single unit, wherein different points of arrival and departure are

matched and eliminate one another[9]. From this perspective, we observe a consistent discrepancy between arrivals and departures, contributing to a gradual increase in the cumulative net flows over the four years of interest, with a plateau during the pandemic years (Figure 1).

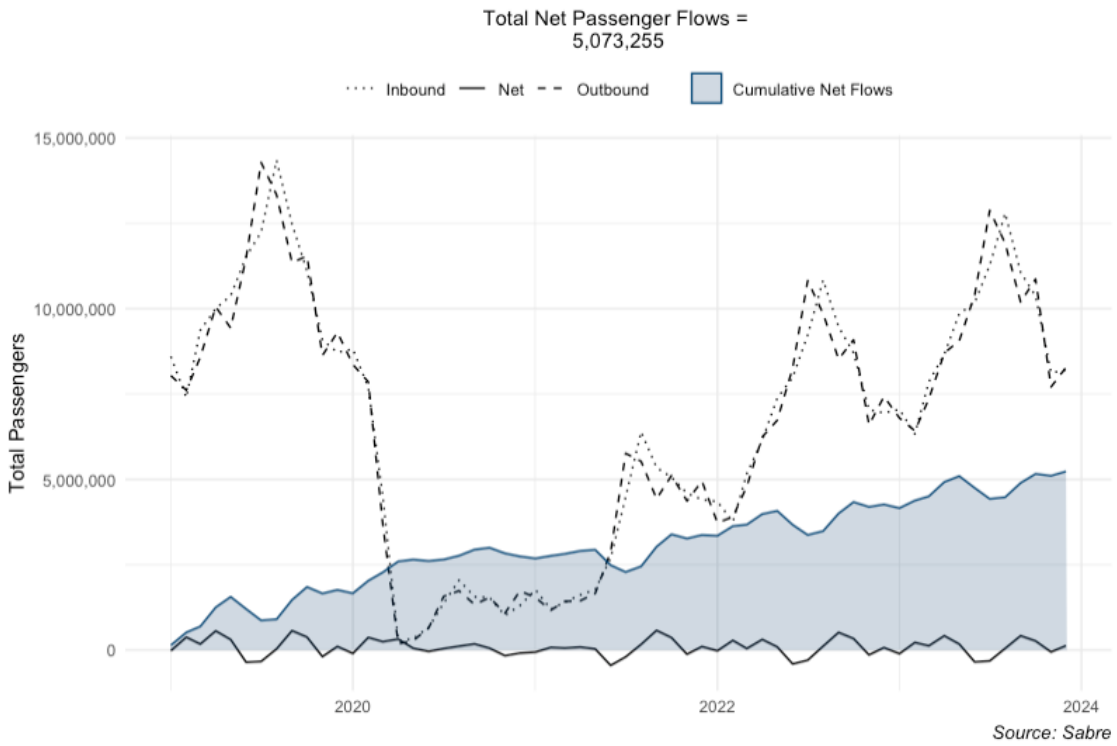

Figure 1: Air passenger inflows and outflows, Schengen Area, 2019-2023 (monthly data)

Due to the disruption generated by the pandemic, it is difficult to discern tangible patterns in net flows over time. Nonetheless, inflows tend to exceed outflows in magnitude, with seasonal peaks in each direction succeeding each other during touristic seasons (Figure C.1). Tourism-related inflows are followed by outflows at the end of the tourist season.

Still taking into consideration the net flows of air travellers, dyadic flows between the Schengen Area and extra-European countries of origin can be examined more closely. Non-Schengen EU member states, in addition to the United Kingdom, are removed from the analysis, for the reasons illustrated above. The countries of origin with the largest net flows are displayed in Table 2, with Colombia, Brazil and Indonesia occupying the top spots.

Broadening the level of analysis to regions of origin, South America, Eastern Europe and the Asian sub-continent are characterized by major net inflows, while North America, Far-East Asia (among which Japan and South Korea), and the Gulf have significant net outflows from the Schengen Area. Last, zooming into country dyads, we can find the top countries of

[9] This also mitigates the unobserved effects of strategic choices made by airlines to offer more or less flights depending on market prospects, which might affect dyadic passenger volumes.

destination for some of the major contributing countries mentioned afore. The next step is to move beyond net air travel flows to include net migration figures.

Table 2: Top countries of origin in terms of net air passenger flows

Country of Origin	Net 2019-2023
Colombia	503,500
Brazil	425,087
Indonesia	351,805
Turkey	343,598
India	331,964
Morocco	301,350
Philippines	185,233
Ukraine	182,801
Tunisia	176,607
Algeria	158,483

4.2 UNACCOUNTED NET FLOWS

Due to the aforementioned difficulties in applying official migration data to the method (equation 2), we use the median QuantMig estimates for immigration and emigration by previous country of residence to derive net migration figures for the Schengen area (Aristotelous et al., 2022). The QuantMig net migration data overlaps with commercial flight data for the year 2019, as already noted. This data is disaggregated by EU/EFTA country, thereby providing EU/EFTA ↔ Rest of the World (RoW)[10] figures, with RoW being separated by region. Figure 3 captures the gist of our two-step methodology, wherein net migration is subtracted from the net flow of passengers into the Schengen Area, yielding *unaccounted net flows*, which signal overstaying. Specifically, the graph focuses on the Schengen ↔ RoW dyad as a unit.

[10] As mentioned before, excluding the EU and the UK.

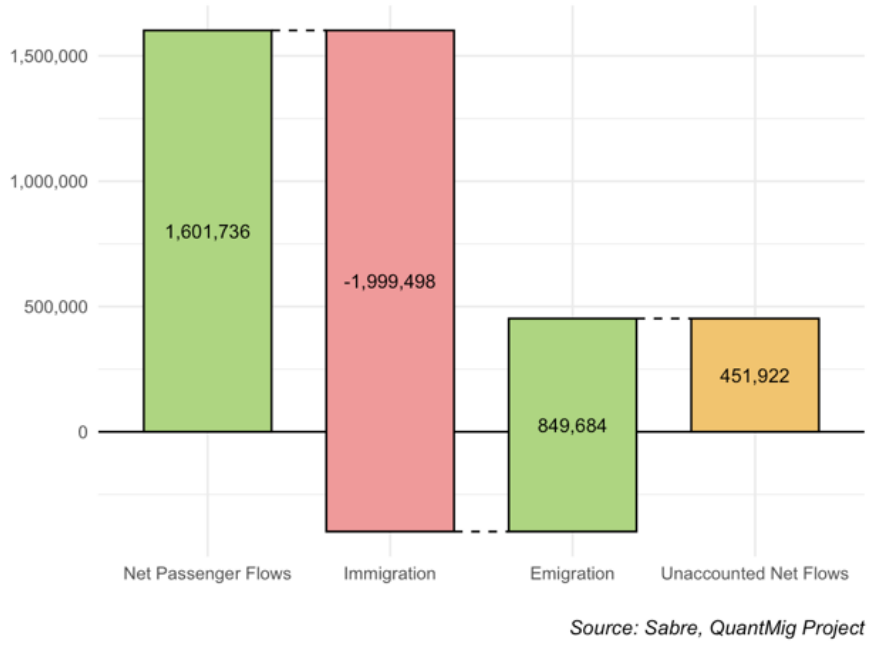

Figure 2: Breakdown of the deductive method for the Schengen area, using median QuantMig estimates of immigration and emigration (2019)

At the regional level, the principal regions with influxes are shown below (Table 3). South/Southeast Asia, a region defined by QuantMig scholars as encompassing countries in the Indian subcontinent and Southeast Asia, has the largest estimated unaccounted net flows. Sub-Saharan Africa and Latin America follow.

The total of visas issued—previously cited as an approximate upper bound on overstaying flows, due to lack of alternative data (Vespe et al., 2017)—is also included in the table. However, for regions with large countries that have visa-free travel regimes with the EU, the unaccounted net flows can exceed the number of uniform visas issued. This is particularly the case for Latin America. Major visa-free countries of the continent, like Brazil, have been identified as sources of overstaying by qualitative inquiries during the pandemic in specific localities (Slootjes, 2022). Large unaccounted net flows highlight the limitations of visa statistics as upper bounds for such regions.

Table 3: Top regions of origin by unaccounted net flows in 2019, QuantMig (QM) & Eurostat (ES) estimates

QuantMig Origin Region	Net Passenger Flows	QM Immigration (50th percentile)	QM Emigration (50th percentile)	QM Unaccounted Net Flows	ES Unaccounted Net Flows	Uniform Visas Issued
South/ Southeast Asia	432,123	262,054	84,900	254,969	147,763	2,039,499
Sub-Saharan Africa	349,153	220,223	71,115	200,045	134,285	806,656
Latin America	493,773	558,962	184,828	119,640	13,276	203,662*
Other Europe	257,510	363,562	169,444	63,392	-852,511	5,721,881
All regions	1,601,736	199,498	849,684	451,922	-1,086,754	14,718,082

This exercise highlights the dependency of the method on accurate immigration data. The Eurostat-based residuals[11], which identify the same principal top regions of origin in terms of unaccounted net flows, differ in magnitude. Additionally, since official data tends to significantly undercount emigration (Keilman and Aristotelous, 2021; Willekens, 2019)[12], values for regions tightly linked to the EU (notably, 'Other Europe') are strongly biased downwards, resulting in negative values for the RoW. Similar results occur for the years 2020-2023, for which Eurostat data are available and overlap with the flight data. These years are therefore left out; the analysis can only go as far as the quality of net migration data (that is, both immigration and emigration data) permits. With all these limitations and caveats in mind, considering Quantmig estimates of net migration the best available, our method leads to an estimate of about 450,000 newly incoming air overstayers in the entire Schengen area in 2019.

[11] Based on a mix of Eurostat datasets, which aims to mitigate missing values as much as possible.

[12] Emigration data is even completely unreported for a country as large as Germany.

5. DISCUSSION

Having gone through a brief application of the method, we critically assess it along the four dimensions proposed by Rodríguez Sánchez and Tjaden (2023) for appraising novel estimation methodologies: reliability, scalability, assumptions, and ethical issues. This evaluation will provide a comprehensive understanding of the challenges and limitations of the approach.

SUMMARY

This report explores a new method to estimate the inflows of overstayers arriving by airline. We concentrate on the Schengen area as destination and on non-European countries as origins, since flights are the most likely transportation mode to Europe for travellers from those countries. Our method relies on complete information on incoming and departing passengers from all airports in the Schengen Area. It unfolds in two steps. In the first step, we compute 'net passenger inflows' as the difference between incoming and outgoing passengers. In the second, we subtract net migration from 'net passenger inflows' to obtain 'unaccounted net flows' (i.e., our estimate of newly incoming overstayers).

ADVANTAGES

This method is inspired by entry-exit systems, which have been implemented by some States using traveller cards. In our application, we use aggregate data on air passenger volumes which are available in complete form through market intelligence companies. The method is theoretically and methodologically parsimonious, provided data is accessible.

RELIABILITY

Volatility is a persistent issue in the estimation of migration flows (Jandl, 2009; Siruno, et al. 2024). This also applies to our method because it is contingent on official migration and air travel data. Finding immigration and emigration data that is accurate and consistent remains a challenge in the EU (Mooyaart et al., 2021). Particularly, immigration and emigration by previous and next place of residence are not available. Even if complemented with citizenship-based figures (such as first permit statistics, like migr_resfirst from Eurostat), missingness is prevalent for immigration statistics, let alone emigration.

In addition to missingness, there are issues of consistency when applying the method to a set of countries like the Schengen Area. As mentioned above, the objective of the method is to assess the number of short-term visas that are violated, and therefore regular long-term visa inflows are deducted. These inflows include individuals on student visas, or asylum seekers, which are part of the QuantMig estimates (Aristotelous et al., 2022). Eurostat data, however, is inconsistent across reporting countries, with some countries omitting refugees

and asylum seekers, or even students (Keilman and Aristotelous, 2021; Lanzieri, 2018). This affects the deductive method, with the said categories of regularly entering immigrants being unsystematically deducted.

As exhibited in the comparison between QuantMig and Eurostat net unaccounted flows (Table 3), the magnitudes of overstaying signals can differ broadly based on the employed migration statistics. This is not an issue that is easily solved in the immediate future, unless the quality of official migration statistics improves.

One further caveat in applying this method is the estimation of uncertainty – that is, upper and lower estimation bounds. While this is highlighted as important and typically implemented in the case of stock estimations (see Van Hook et al., 2021 for the United States), it is more difficult to carry out inflow-based deductive methodologies due to the dearth of rate parameters that can be simulated along specific assumed distributions (such as birth or emigration rates). This is a challenge that will have to be addressed in future studies.

SCALABILITY

The method is highly scalable across geographies, provided air travel figures are available and overstayers are unlikely to travel by land. In fact, the Schengen Area's geographic situation (that is, including numerous land borders) and border permeability with neighbouring states render it one of the most difficult entities to which it can be applied. States with low border porosity or of insular nature constitute the most apt use cases, since entry and exit rely almost exclusively on commercial airlines. All in all, if the method is applicable to Schengen, it may be used anywhere, conditional on data availability on air passenger and regular migration volumes.

Furthermore, its use can be extended beyond mere overstaying in the destination country to the identification of countries of transit for irregular migration. Potential irregular migrants may fly to such a transit country—therefore recording an inflow without a matching outflow—and then attempt to cross land borders into their final destination. This is salient, as migrants have been found to face hardship and discrimination in intermediary countries to which they have flown to reside temporarily or migrate onwards (see for instance Lucio et al., 2023 for the Darién Gap; Pouessel, 2023 for Tunisia). Instead of measuring overstaying flows *per se*, then, the method could help identify novel, possibly perilous travel patterns that emerge as a consequence of increasingly restrictive tourist visa issuance and geographic disparities in their distribution (European Commission, 2024). Nonetheless, the quality of official migration statistics across the board once again hamper scalability also in this alternative application. For the current paper, for instance, attempts were made to retrieve immigration statistics from third-country statistical offices, like Brazil, to gauge the missing emigration numbers from the EU, without success (NEPO, 2024).

ESTIMATION ASSUMPTIONS

Beyond the assumption of reliable regular migration statistics, which has been discussed at length, the main assumption of this approach to measure overstaying consists in arrival by commercial air travel. This jeopardizes the method's usefulness for geographically proximate countries to the state of interest. However, the proportion of cross-border travel undertaken by means of airline travel has been shown to rapidly rise as distance increases (Recchi et al.,

2019). It is therefore most reasonably deployed by pragmatically removing neighbouring states that might induce bias, as we did in this paper with non-Schengen EU states and the UK. Several of the further implicit assumptions of the method are surveyed in Table A.1.

Perhaps of lesser importance is the assumption of absence of *stopover bias*. As mentioned in Table A.1, this refers to the possibility that a traveller performing a roundtrip incurs a *stopover*—namely, a layover of more the 24 hours, thereby "breaking" the trip—in one direction but not the other. This could be due to, for instance, structural differences in scheduling in one direction of travel relative to the other.

ETHICS

As a benchmark for ethical appraisal, we refer to Cyrus (2023), who comprehensively lays out guardrails for identifying potential methodological risks pertaining to digitized social research. Data protection, algorithmic biases, and research impacts are the foci of the ethical challenge framework, building on prior scholarship by Salah et al. (2022).

The chief area of concern is related to the protection of personal data. Insofar as the current approach is based entirely on aggregated, rather than individualized, longitudinal data, most of the stumbling blocks identified are avoided. This, in fact, is one of the appeals of using the method presented above. It circumvents ethical challenges associated with database systems enabling the identification of overstayers, long sought after by states (Rodriguez Sanchez and Tjaden, 2023).

The second component sketched out by Cyrus (2023) is loosely related to *dataism*, or the belief that all social behaviour can be traced using digital technologies, with little regard for the generation of the undergirding data itself. The current study is largely founded on airline registration systems as tracked by a private entity (Sabre), which sells its Global Distribution System to profit-maximizing actors of the aviation and travel industries, such as airlines, airports, or travel agencies. It therefore offers limited insight into the underlying data production mechanisms. This is an ethical limitation. Some metadata does offer insights into the data production process (e.g., the omission of commercial charter flights from the final dataset, and more). Despite this shortcoming, the global, far-reaching nature of the dataset at hand has the advantage of highlighting the diversity of mobility in time and space, moving beyond categories of nationality or citizenship, as stressed by Stielike (2023, p. 192) and Cyrus (2023, pp. 42-43).

The last issue identified is the potentially negative impact of the output for migrant communities. In this regard, it is crucial to stress that we conceive of this research as purely exploratory. The aim is primarily to outline and implement an approach to measure a particularly difficult-to-sample population, rather than to provide definitive, policy-informing figures. Our estimation exercise can nonetheless serve as an external benchmark once the EU might decide to implement its Entry/Exit System (EES).

CONCLUSION

The methodology outlined above offers a preliminary and innovative approach to measuring overstaying flows, which are severely lacking in the EU. While it contains certain limitations, particularly related to official migration statistics, its output can provide a foundation for future modelling of overstay flows. In future research, inferential modelling could focus on understanding the drivers and causes of visa overstaying rather than descriptively measuring

the phenomenon in raw terms. By treating flows as signals and modelling dyad-months, this approach could yield a clear picture of the underlying dynamics of overstaying, and of the complexities of migration patterns more broadly.

REFERENCES

Aristotelous G., Smith P., W. F., Bijak J. (2022). *Technical report: Estimation methodology. QuantMig Project Deliverable D6.3.* Southampton: University of Southampton. https://quantmig.eu/res/files/QuantMig%20deliverable_6_3.pdf

Baker, B., & Warren, R. (2024). *Population Statistics: Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2018-January 2022.* Department of Homeland Security's Office of Immigration Statistics. https://ohss.dhs.gov/sites/default/files/2024-06/2024_0418_ohss_estimates-of-the-unauthorized-immigrant-population-residing-in-the-united-states-january-2018%25E2%2580%2593january-2022.pdf

Cyrus, N. (2023). *Ethical Benchmarking in Irregular Migration Research* (MIrreM Working Paper No.3). Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.8389189

de Haas, H. (2023). *How migration really works: A factful guide to the most divisive issue in politics*. Penguin Books. Penguin.

Deutschmann, E., Gabrielli, L., & Recchi, E. (2023). Roads, rails, and checkpoints: Assessing the permeability of nation-state borders worldwide. *World Development, 164*, 106175. https://doi.org/10.1016/j.worlddev.2022.106175

ELIAMEP. (2009). *Clandestino Project: Final Report.* CLANDESTINO: Undocumented Migration: Counting the Uncountable. Data and Trends across Europe. https://emnbelgium.be/publication/clandestino-project-final-report

European Commission. (2014). *Proposal for a regulation of the European Parliament and of the Council establishing a touring visa and amending the Convention implementing the Schengen Agreement and Regulations (EC) No 562/2006 and (EC) No 767/2008* (COM/2014/0163 final - 2014/0095 (COD)). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52014PC0163

European Commission. (2016, 14 September). *State of the Union 2016: Commission Targets Stronger External Borders* [Press Release]. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/IP_16_3003

European Commission. (2024, 15 May). *Schengen Visa Applications Reach 10.3 Million in EU- and Schengen-Associated Countries* [Press Release]. Directorate-General for Migration and Home Affairs. https://home-affairs.ec.europa.eu/news/visa-applications-reach-103-million-eu-and-schengen-associated-countries-2024-05-15_en

Eurostat. (2024). *International Migration Statistic: Reference Metadata in Euro SDMX Metadata Structure (ESMS).* Eurostat, the Statistical Office of the European Union. https://ec.europa.eu/eurostat/cache/metadata/en/migr_immi_esms.htm#relatedmd1724925004995

Fasani, F. (2008). *Italian Country Report for CLANDESTINO.* Undocumented Migration: Counting the Uncountable. https://www.eliamep.gr/wp-content/uploads/2017/12/clandestino_report_italy_final_3.pdf

Finotelli, C., & Sciortino, G. (2013). Through the Gates of the Fortress: European Visa Policies and the Limits of Immigration Control. *Perspectives on European Politics and Society, 14*(1), 80–101. https://doi.org/10.1080/15705854.2012.732390

Hansen, F., & Pettersson, J. (2021). Contradictory migration management? Differentiated security approaches to visa overstay and irregular border crossings in the European Union. *European Security, 31*(1), 117–134. https://doi.org/10.1080/09662839.2021.1945038

Huang, S. (Sam), Bao, J., Jin, X. (Cathy), & Wang, J. (2023). COVID-19, Tourism, and China: Seeking Certainties in Uncertain Times. *Journal of China Tourism Research, 19*(1), 1–6. https://doi.org/10.1080/19388160.2023.2174644

Hubert, M., & Vandervieren, E. (2008). An adjusted boxplot for skewed distributions. *Computational Statistics & Data Analysis, 52*(12), 5186–5201. https://doi.org/10.1016/j.csda.2007.11.008

Jandl, M. (2008). Chapter 3: Methods Methods for estimating stocks and flows of irregular migrants. In *Report on Methodological Issues* (Clandestino Deliverable D3). Clandestino Project. https://migrant-integration.ec.europa.eu/sites/default/files/2009-07/docl_9061_620418687.pdf

Jandl, M., Vogel, D., & Iglicka, K. (2008). *Report on methodological issues* (Clandestino Deliverable D3). https://migrant-integration.ec.europa.eu/library-document/clandestino-project-report-methodological-issues_en

Keilman, N., & Aristotelous, G. (2021). *Expert Opinion on Migration Data* (QuantMig Project Deliverable D6.1). Southampton: University of Southampton. https://www.quantmig.eu/res/files/QuantMig%20Deliverable%206.1%20Expert%20opinion%20on%20migration%20data.pdf

Kraler, A., & Ahrens, J. (2023). *Conceptualising Migrant Irregularity for Measurement Purposes* (MIrreM Working Paper No. 2 (Version 3)). Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.7868237

Lanzieri, G. (2018, October 26). *Comparability of migration and residence permits data in the EU statistics*. UNECE – Conference of European Statistics, Work Session on Migration Statistics. https://ec.europa.eu/eurostat/cache/metadata/Annexes/migr_immi_esms_an_3.pdf

Lucio, J.V.P., Congram, D., Franco-Paredes, C., Mohareb, A.M., Sancho, J.A.S., Flores, A., Rodríguez, G., Flores, M.R.P. (2023). Mortality during migration through the Darién Gap in Panama in 2018-22. *Lancet 402*(10410), 1321-1322. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(23)01635-5/fulltext

Mooyaart J.E., Dańko M.J., Costa R. & Boissonneault, M. (2021). *Quality assessment of European migration data* (QuantMig Project Deliverable D6.2). The Hague: Netherlands Interdisciplinary Demographic Institute (NIDI-KNAW)/University of Groningen. https://quantmig.eu/res/files/QuantMig_Deliverable%206.2%20vf.pdf

Nemes, N., & Happ, E. (2024). Possibilities of Creating Cross-Border Tourist Destinations in East-Central Europe. *Geo Journal of Tourism and Geosites*, 55(3), 1066-1075. https://gtg.webhost.uoradea.ro/PDF/GTG-3-2024/gtg.55308-1280.pdf

Núcleo de Estudos de População Elza Berquó (NEPO). Imigrantes Internacionais Registrados no Brasil (com Registro Nacional Migratório – RNM). Unicamp. https://www.nepo.unicamp.br/observatorio/bancointerativo/numeros-imigracao-internacional/sincre-sismigra/

Piccoli, L., Dzankic, J., Ruedin, D., & Jacob-Owens, T. (2023). Restricting Human Movement During the COVID-19 Pandemic: New Research Avenues in the Study of Mobility, Migration, and Citizenship. *International Migration Review*, 57(2), 505-520. https://doi.org/10.1177/01979183221118907

Pouessel, S. (2023). *Un cosmopolitanism avorté? Parcours migratoires subsahariens et entrée en politique du racisme en Tunisie*. Arab Reform Initiative. https://www.arab-reform.net/fr/publication/un-cosmopolitisme-avorte-parcours-migratoires-subsahariens-et-entree-en-politique-du-racisme-en-tunisie/

Recchi E., Deutschmann E., & Vespe M. (2019). *Estimating transnational human mobility on a global scale* (MPC/RSCAS Working Paper Series 2019/30). EUI Migration Policy Centre [Global Mobilities Project]. https://cadmus.eui.eu/bitstream/handle/1814/62326/RSCAS_2019_30.pdf?sequence=1

Recchi, E., Tittel, K. (2023). The Empirical Study of Human Mobility: Potentials and Pitfalls of Using Traditional and Digital Data. In Bertoni, E., Fontana, M., Gabrielli, L., Signorelli, S., Vespe, M. (eds) *Handbook of Computational Social Science for Policy*. Springer, Cham. https://doi.org/10.1007/978-3-031-16624-2_23

Rousseeuw, P. J., & Hubert, M. (2018). Anomaly detection by robust statistics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8*(2), e1236. https://doi.org/10.1002/widm.1236

Salah, A. A., Korkmaz, E. E., & Bircan, T. (Eds.). (2022). *Data Science for Migration and Mobility* (1st ed.). British Academy. https://doi.org/10.5871/bacad/9780197267103.001.0001

Sîrbu, A., Goglia, D., Kim, J., Magos, P.M., Pollacci, L., Spyratos, S., Rossetti, G., & Iacus, S.M. (2024). International mobility between the UK and Europe around Brexit: a data-driven study. *Journal of Computational Social Science.* https://doi.org/10.1007/s42001-024-00277-4

Siruno, L., Leerkes, A., Hendow, M. & Brunovksá, E. (2024). *Working Paper on Irregular Migration Flows*, *MIrreM Working Paper No. 9*, Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.10702228

Slootjes, J. (2022). *The COVID-19 Catalyst: Learning from pandemic-driven innovations in immigrant integration policy (pp. 1–24).* Migration Policy Institute Europe. https://www.migrationpolicy.org/sites/default/files/publications/mpie-pandemic-innovation-integration-gov_final.pdf

Stielike, L. (2023). I.2.8 Big Data. In T. Scharrer, B. Glorius, J. O. Kleist, & M. Berlinghoff (Eds.), *Flucht- und Flüchtlingsforschung* (pp. 189–194). Nomos Verlagsgesellschaft: . https://doi.org/10.5771/9783748921905-189

Triandafyllidou, A. (Ed.). (2010). *Irregular Migration in Europe: Myths and Realities* (1st ed.). Routledge.

Van Hook, J., Morse, A., Capps, R., & Gelatt, J. (2021). Uncertainty About the Size of the Unauthorized Foreign-Born Population in the United States. *Demography, 58*(6), 2315–2336. https://doi.org/10.1215/00703370-9491801

Vespe, M., Natale, F., & Pappalardo, L. (2017). Datasets on Irregular Migration and Irregular Migrants in the EU. *Migration Policy Practice VII*(2), 26-33. https://publications.jrc.ec.europa.eu/repository/handle/JRC107749

Warren, R. (2017). DHS overestimates visa overstays for 2016; overstay population growth near zero during the year. *Journal on Migration and Human Security, 5*(4), 768-779. https://doi.org/10.1177/233150241700500403

Willekens, F. (2019). Evidence-Based Monitoring of International Migration Flows in Europe. *Journal of Official Statistics 35*(1), 231-277. https://doi.org/10.2478/jos-2019-0011

ANNEX 1

The MIrreM Methods Lab conducted a review of 21 traditional and innovative methodological approaches for estimating irregular migrant stocks and flows. Each approach was assessed based on its core concept, data sources, definition and coverage of irregular migration, estimation assumptions, reliability, scalability, general assumptions, and ethical considerations.

Building on this review, we developed six innovative approaches that have the potential to advance research on irregular migration.

As part of the broader MIrreM project, the WP6 Methods Innovation Lab carried out the following six Pilot Studies (PS). Please find the MIrreM Briefing Papers about the other Pilot Studies linked below:

MIrreM Briefing Papers	Authors	DOI
PS1 - Exploring the use of aggregate air passenger data for estimating overstayer inflows	Luca Bernasconi Ettore Recchi	https://doi.org/10.5281/zenodo.14809013
PS2 - Measuring the participation of irregular migrants in the informal economy	Aslı Salihoğlu Carlos Vargas-Silva	https://doi.org/10.5281/zenodo.14809000
PS3 - Estimating irregular migrant stocks using social media data and machine learning	Alejandra Rodríguez-Sánchez Jasper Tjaden	https://doi.org/10.5281/zenodo.14808984
PS4 - Irregular migration: What can mortality reveal?	Johan Surkyn Tuba Bircan	https://doi.org/10.5281/zenodo.14808979
PS5 - Estimating irregular migration in the UK using a health care reform	Alejandra Rodríguez-Sánchez Jasper Tjaden	https://doi.org/10.5281/zenodo.14808948
PS6 - Measuring irregular migration stocks through social media surveys	Jasper Tjaden Alejandra Rodríguez-Sánchez	https://doi.org/10.5281/zenodo.14801999

ANNEX 2: Typology of Potential Errors and Complicating Factors

Table A. 1: Typology of potential error sources/complicating factors

Name	*Description*	*Biased inflow example*	*Biased outflow example*	*Reasoning & risk*	*Potential Remedial Measure*
Mixed-mode travellers	Travelers that employ different modes of transportation to enter and leave the Schengen Area.	Individual flies to Austria from Albania (prior to 2023) and returns by car.	Individual flies to Albania (prior to 2023) from Italy and returns by boat.	Primarily poses an estimation problem for states bordering the Schengen Area/the European Union.	Omit bordering countries from analysis.
Intra-Schengen cross-border travel (circular tourism)	Travelers (typically tourist groups, backpackers) entering the Schengen Area through one country and departing from another.	Individual flies from the USA to Germany, travels across Europe and returns from Italy. ⇒ Biased inflow to Germany	Same example, reversed ⇒ Biased outflow from Italy (to USA)	One can assume it is relatively common for tourists or business travellers to undertake such trips.	Treat Schengen Area as a unit: aggregate in- and outflows for the entire Area.
Stopover bias	Travelers that have a *stopover* (i.e. a 24+ hour layover) in one direction of their (O&D) trip but not the other.	Individual travels to the Maldives from Germany via Dubai (UAE). On the way to MV, they stop in Dubai but shortly (not breaking the O&D). On the way back, they sleep one night in Dubai. ⇒ Biased inflow from the UAE	Same example, reversed ⇒ Biased outflow to Maldives	Primarily an issue for distant tourist destinations, particularly hard-to-access islands which require complicated itineraries. Also potentials for migrants returning to their home countries temporarily via more complex itineraries. Not a major problem if stopovers are randomly distributed across directions for given dyads. However, there might be scheduling structures that induce bias, creating higher likelihood of stopover in one direction but not the other.	Testing for stopover bias: - Using legs incurred during trip, identify the proportion of passengers that use a particular *layover* (as opposed to a *stopover,* which breaks the trip) for a specific dyad. - Then, fit a GLM to test whether there is a systematic difference in the use of layovers between directions of travel. If this is the case, we might have stopover bias.
Lacking emigration data	Travelers that enter the Schengen Area for third countries to reside are recorded more consistently than travellers leaving the Schengen Area.	An EU dweller moves away and travels by plane in the process. If the departure is not recorded in official migration statistics, unaccounted net flows might be underestimated.		Emigration statistics are known to be lacking in the European Union (Keilman and Aristotelous, 2021; Willekens, 2019).	

Lacking immigration data (by country of previous residence)	Large number of missing values in immigration by previous country of residence fails to match Sabre's inflows.		An individual moves back into the Schengen Area for living and travels by plane in the process. If the arrival is not recorded in official migration statistics, unaccounted net flows might be overestimated.	Preliminary analysis does show a high number of missing values in both immigration and first permit dyadic data, which can lead to structural issues.	Corroboration via further data from national statistics bureaus.

ANNEX 3: Treatment of the Missing Values

Upon the merging of the inbound and outbound datasets, there are rows that may contain missing values in one direction but not the other. Since we consider net passenger flows at the monthly level, we pay special attention to these cases, wherein an inbound flow is unmatched with an outbound flow, or vice-versa. The goal is to investigate whether the missing values represent actual zeros or whether the observation should be removed for safety, to avoid imbalances in the net flows.

Of the initially existing 219,037 observations, there are 10,897 (4.97%) outbound observations without a corresponding monthly inbound value and 13,658 (6.24%) inbound observations without a corresponding outflow, amounting to 11.21% of rows having solely unidirectional information. However, the dyad-months for which this is the case have only low flow numbers. For instance, the median passenger inflow without a corresponding outflow is 3.07 (see Table B.1). Thus, one may safely assume that many of the missing are true zeros, in dyads with negligible passenger numbers.

Table B. 1: Typology of potential error sources/complicating factors

In month *m*:	Mean	Median
Unmatched Inbound Values	7.41	3.07
Unmatched Outbound Values	5.50	2.87

It is nonetheless important to examine the distribution of these dyads to check if any extreme values exist that may skew the results. First inflows without matching outflows are considered. A simple Tukey-style boxplot (Rousseeuw and Hubert, 2018) reveals the skewness in the distribution (Figure B.1). As such, skewness-adjusted methods proposed by Hubert and Vandervieren (2008) are used to find the *fence*—the value beyond which data points are labelled extreme. Here, it is 34.03. Values below this threshold are safely assumed to be true zeros and are imputed as such.

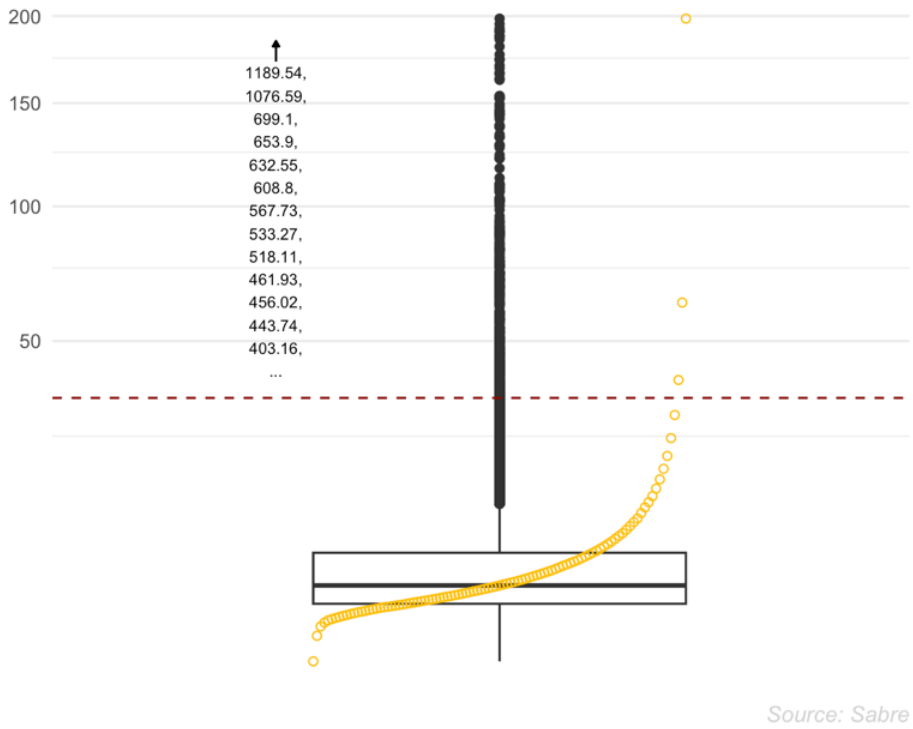

Figure B. 1: Unmatched inflows with skewness-adjusted fence in red

Further examination of the inflows without corresponding outflows shows that they are concentrated in 2020, year in which pandemic-related entry restrictions on transnational mobility peaked (Piccoli et al., 2023). Therefore, it is conceivable that outliers in the year 2020 are actual zeros and they are treated as such in our analysis. It is also noteworthy that once having adjusted for COVID-19, about 20% of the remaining outliers involve China and Myanmar, the former of which still had stringent international travel restrictions in 2022 (Huang et al. 2023). Therefore, these dyads too (for the years 2021 and 2022) are assumed to correspond to null outbound flows, and zero is imputed accordingly.

Having taken these *ad hoc* measures, we are left with 129 dyad-months containing inflows without outflows, which have a median of 49.7—small figures in the grand scheme of international travel. Most involve distant countries and Southeast Asia. They are listwise-removed conservatively.

A similar exercise is carried out for outbound flows which do not have a matching inflow in a given month. The outliers are generally smaller and the fence is 26.62 (see Figure B.2). Once 2020 values are duly assigned a zero, the remaining 185 missing dyad-months (corresponding median outflows = 40.6) relate primarily to Pacific countries for which zeros are very plausible. However, some of these remaining outliers are listwise removed for caution. This is a more conservative measure, particularly since several dyads contain

imbalanced results throughout the time series (persistent missing inflows but not outflows), which may be the result of stopover bias.[13] These time series are entirely removed.

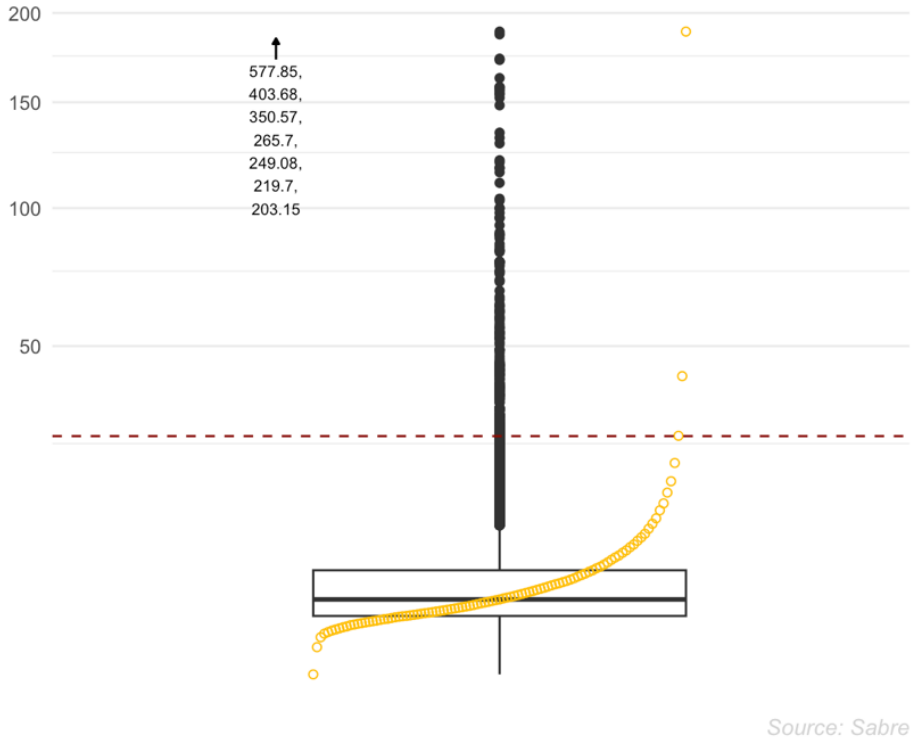

Figure B. 2: Unmatched outflows with skewness-adjusted fence in red

The *implicit* time series gaps are also scrutinized for the sake of completeness. Due to the processing above, these are gaps for which both outbounds and inbounds values are missing. Of the 46,405 implicitly-missing flows, more than 60% of the dyad-months concern origin countries from the Caribbean, the Pacific, and Sub-Saharan Africa, in combinations where it is particularly plausible that there are no flows whatsoever in a given month[14]. Furthermore, more than a third occur in 2020, after the month of March, when restrictions became widespread (Piccoli et al., 2023). Therefore, from a substantive point of view, it is most reasonable to leave these rows as missing, or implicitly treat them as zeros in our subtractive accounting.

[13] VU ↔ NO and BN ↔ PT stand out for instance, albeit with negligible total travel volume.

[14] As an example, the British Virgin Islands, Palau, Tonga, and Saint Kitts and Nevis are among the top countries with missing dyad-months with European countries. Travel from these micro-states to European countries is very sporadic.

ANNEX 4: Net Passenger Flows

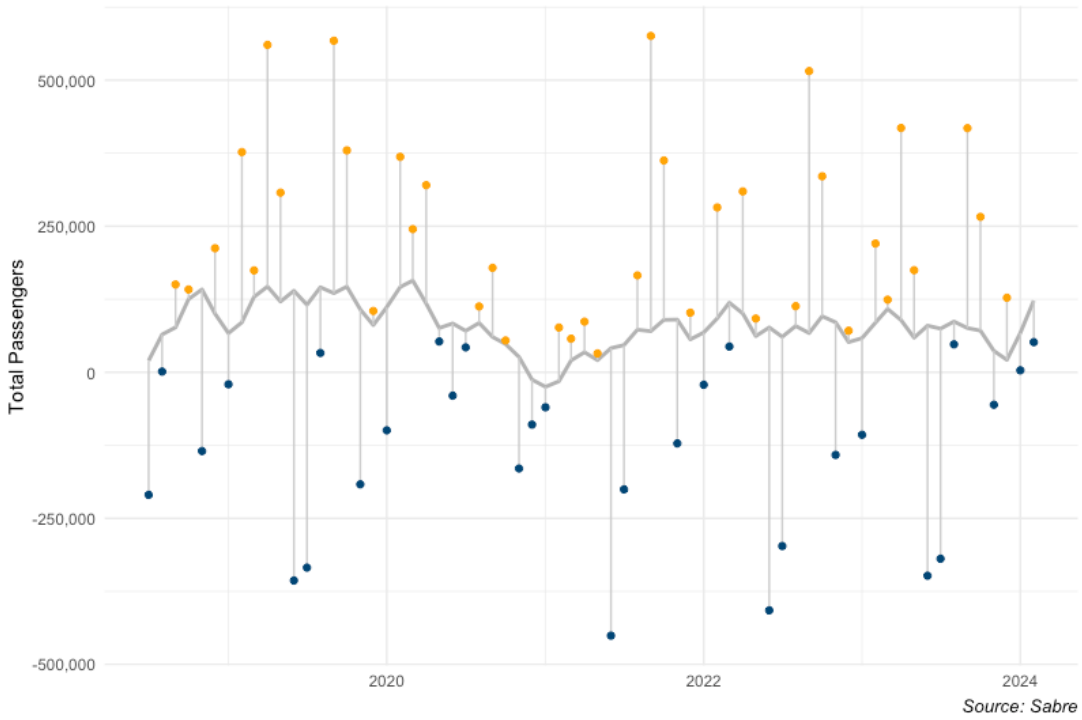

Figure C. 1: Monthly net passenger flows to the Schengen area, with centred moving average (k = 13). Colours differentiate between net flows above and below the average

ABOUT THE AUTHORS

Ettore Recchi is a part-time professor at the European University Institute's Migration Policy Centre (MPC) as well as a professor of sociology at Sciences Po Paris (CRIS) and Fellow of the Institut Convergences Migrations.

Luca Bernasconi is a research associate the European University Institute's Migration Policy Centre (MPC).

COPYRIGHT CLAUSE

THE MIRREM CONSORTIUM

University for Continuing Education Krems (Coordinator)

European University Institute

University of Osnabrück

University of Maastricht

University of Turku

Complutense University Madrid

Hellenic Foundation for European and Foreign Policy (ELIAMEP)

University of Milan

University of Potsdam

Platform for International Cooperation on Undocumented Migration (PICUM)

International Centre for Migration Policy Development (ICMPD)

Migration Policy Institute Europe (MPI-E)

University of Warsaw

Vrije Universiteit Brussel

Instituto Universitário de Lisboa

Associated Partners:

Toronto Metropolitan University

University of Leicester

University of Oxford