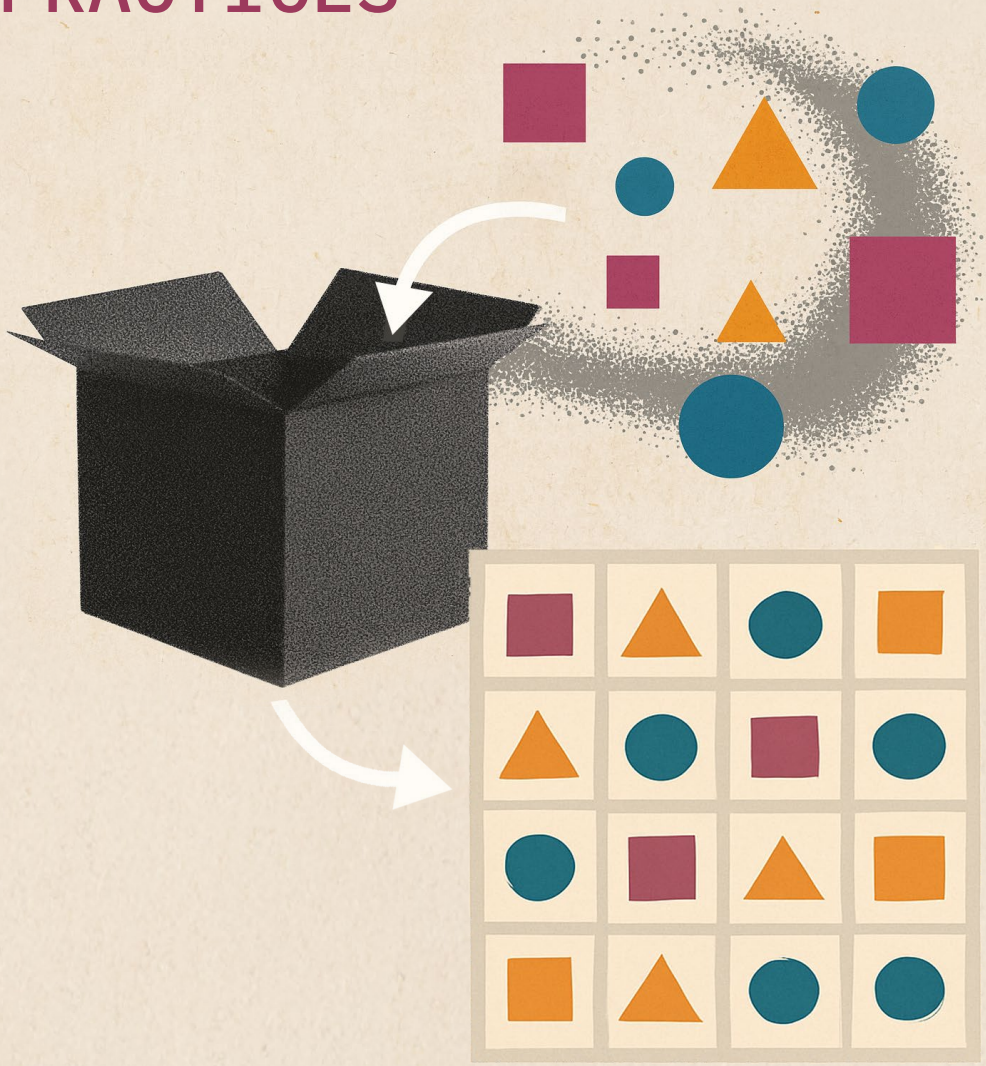
HANDBOOK ON IRREGULAR MIGRATION DATA

CONCEPTS, METHODS AND PRACTICES



Edited by Denis Kierans & Albert Kraler

Acknowledgements

The Handbook has benefited from the collegial support, critical feedback and contributions coming from the MIrreM project team, which form the backbone of the Handbook. We also thank the members of the MIrreM's advisory board who not only gave critical feedback but also opened doors to relevant stakeholders outside the project. Moreover, we are indebted to the wider 'community of practice' engaged in this field for often critical yet helpful feedback on the Handbook and getting us involved in different fora where we could test our ideas.

We would especially like to thank the contributors to this Handbook who have been instrumental in extending the scope of this Handbook beyond the confines of the MIrreM project, and allowing us to cover a much broader ground than we would have been able to do and turning this into a truly collective endeavour. We also wish to thank colleagues who have taken time to review drafts of chapters and textboxes, including Julia Descamps, Frank Laczko, Arjen Leerkes, Fran Meisner, Rocco Molinari, Laura Peitz, Ann Singleton, Randy Stache, Georgina Sturge, Jennifer Van Hook, Peter Walsh and Teddy Wilkin.

Last but not least, the Handbook would not have been possible without the support from Adriana Harm and Franziska Klauser at the Department for Migration and Globalisation at the University for Continuing Education (Danube University) Krems. Special thanks also to Chloë Bouvier and Imanol Legarda of MIrreM's project partner PICUM (Platform for International Cooperation on Undocumented Migrants).

Images: Philipp Lammer & Visuality

MIRREM
Measuring Irregular Migration

info@irregularmigration.eu
https://irregularmigration.eu/

University of Krems Press

Editors: Denis Kierans & Albert Kraler

Cover Illustration: Franziska Klauser & Philipp Lammer

Layout and Design: Franziska Klauser & Philipp Lammer, based on a template created by Chloë Bouvier

Publisher: University of Krems Press, © 2025

ISBN: 978-3-903470-24-8

DOI: https://doi.org/10.48341/g31s-vq79

Product safety according to EU regulation: bibliothek@donau-uni.ac.at

Suggested citation:

Kierans, D. & Kraler, A. (eds.). **Handbook on Irregular Migration Data. Concepts, Methods and Practices.** Krems: University of Krems Press, 2025. https://doi.org/10.48341/g31s-vq79

FUNDING ACKNOWLEDGEMENT

This book has benefited from financial support from the European Union's Horizon Europe research and innovation programme, UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee and the Canada Excellence Research Chairs Program of the Government of Canada.

Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union, the Research Executive Agency, UKRI and the Government of Canada.

Funded by the European Union

Co-funded by:

UK Research and Innovation

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

Table of Contents

Foreword

Frank Laczko

What is the scale of irregular migration across Europe? What have been the recent trends in irregular migration? How good are the data on irregular migration? What data are needed to improve our understanding of irregular migration? These are some of the key questions addressed in this new book based on research conducted in 20 countries in Europe and North America.

Irregular migration is a topic which receives a vast amount of policy, media and public attention. Yet reliable, timely and comparable data on the subject are often hard to find. Even when data are available, they may be misinterpreted and misused by policymakers and the media who do not understand fully how migration statistics are produced. There is a lack of guidance on how best to measure irregular migration.

In response to this challenge, the European Commission launched the MIrreM project in 2022 to strengthen the understanding and use of irregular migration data across Europe. The aim of the project is not just to produce more data, but to support more informed and transparent policy conversations, helping to ensure that decisions reflect evidence, not assumptions. This Handbook is one of the key outputs of this project.

The Handbook provides a user-friendly resource for navigating irregular migration data – highlighting what is available, how to interpret it, and where the limits lie. It speaks to policymakers, journalists, researchers, and advocates for those who want to use data more responsibly and effectively in a domain often dominated by uncertainty and speculation.

Irregular migration intersects with border management, asylum systems, labour markets, and social integration. Yet the data underpinning these discussions are often patchy, politicised, or poorly understood. This book provides tools to critically assess available estimates and encourages a more nuanced debate around what irregular migration numbers can (and cannot) tell us.

This Handbook builds directly on the CLANDESTINO project (2007-2009), which offered one of the first systematic attempts to estimate irregular migration in Europe. One of the key headline figures from the MIrreM project is the estimate that there were between 2.6 million and 3.2 million irregular migrants living in 12 European countries over the period 2016-2023. However, the quality of data on irregular migration in many countries is poor or outdated. Indeed, 5 countries studied by MIrreM have not produced any estimates in

recent years. Countries also tend to collect data on irregular migration in very different ways making comparisons difficult.

This Handbook offers guidance on how to interpret statistics on irregular migration. It offers a framework for navigating complexity rather than eliminating it – recognising that some uncertainty is inevitable, but that it can still be managed thoughtfully. The book clarifies complex concepts and the technical aspects of irregular migration data. Examples of data innovation are highlighted in the book and there is a discussion of the potential of using non-traditional sources of data to understand irregular migration trends. The book provides examples of insights gained from analysing data produced by the private sector and through the analysis of social media data. The Handbook provides examples of how irregular migration data are used in practice — from policymaking to service provision — helping to anchor abstract concepts in the real world. The book suggests practical tools for interpreting irregular migration data, supporting more informed and responsible use of estimates and indicators. The book frames data as a process, not just a product, drawing attention to how data are shaped by legal categories, institutional priorities, methodological decisions, and real-world constraints.

What comes next?

This publication is a step forward, not the destination. The long-term goal is to foster a more integrated and strategic approach to irregular migration data – one that combines the rigour of official statistics with the innovation of alternative data sources. The MIrreM project has taken important steps in this direction, but sustained progress will require ongoing collaboration across governments, civil society, academia, and the private sector – particularly in Europe, where MIrreM found irregular migration data especially uneven. It is hoped that this Handbook will serve as a contribution for that continued work.

Executive Summary

Denis Kierans and Albert Kraler

Irregular migration is a persistent feature of mobility to and within Europe, yet the evidence base remains fragmented, inconsistent, and often misunderstood. The very notion of 'irregular migration' is vague, ambiguous, and ultimately a legal and policy category that requires careful, context-sensitive interpretation. This Handbook distils lessons from the Measuring Irregular Migration (MIrreM) project and contributions from colleagues in research, government, and civil society. Focused on Europe, it offers practical guidance on how to compile, interpret, and use irregular migration data, bringing together conceptual clarifications, ethical safeguards, methodological advances, and examples of good practice.

This Handbook is intended for policymakers, statisticians, journalists, researchers, and practitioners. It can be used to: understand the current state of knowledge; recognise opportunities and pitfalls when working with data; identify promising approaches for producing estimates; learn from practical examples; and inform strategies for improving responsible use of data in policymaking.

Concepts and definitions

Conceptual clarity is essential because definitions shape what is measured and compared. Irregular migration is not a fixed fact but a policy category that varies across countries and over time.

- In this Handbook, 'irregular migration' refers to the phenomenon, 'irregular migrants' to people in that situation, and 'migrant irregularity' to the condition of lacking legal status under national law. Because definitions differ, comparability is limited.
- Terms such as 'illegal', 'undocumented', and 'irregular' carry connotations that shape perceptions and policies. The Mixed Migration Centre shows that rigid categories like asylum seeker or economic migrant often fail to capture overlapping motivations and vulnerabilities.
- The MIrreM taxonomy distinguishes irregularity as a subset of precarious immigration status and separates pathways into and out of irregularity from stocks and flows. For example, asylum seekers may enter a country irregularly but gain a legal right to stay once granted protection.
- The taxonomy also captures how individuals' legal status changes over time as they move into, through, and out of irregularity.
- Definitional choices shape what is visible in data and policy debates. The IOM Missing Migrants Project, for instance, records deaths in transit, highlighting border risks while not capturing deaths linked to irregular status after arrival.

Ethics and trust

Irregular migration data often concern vulnerable people whose rights and safety can be affected by how data are collected and used. Ethical safeguards and trust-building are essential to mitigating risks to individuals and groups and improving data quality.

- Safeguards need to go beyond GDPR compliance to avoid reinforcing stigma, fuelling harmful narratives, or deterring service use. Spain's municipal population register (*padrón*) is an example – it improves trust and service access while generating reliable local data.
- Privacy-compliant data linkage can be achieved, as shown by Austria's Austrian Micro Data Centre, which connects administrative datasets for longitudinal analysis without exposing personal identities.
- Trusted enumerator models such as the Mixed Migration Centre's 4Mi survey, based on community-based data collection across more than 30 countries, can reach migrants in transit and destination contexts often missed by other surveys.
- Innovative recruitment strategies are demonstrated by Germany's Migrationsmuster nach Ablehnung im Asylverfahren (MIMAP) survey, which combines ethnography with app-based respondent-driven sampling to reach rejected asylum seekers. Findings suggest that family ties and work opportunities often influence onward migration more than enforcement pressure.

Estimates and indicators

Estimates provide critical context for policy debates, but their quality, scope, and timeliness vary. Indicators help track aspects of irregular migration, but they are partial and need careful interpretation.

- Between 2016 and 2023, an estimated 2.6 to 3.2 million irregular migrants are thought to have lived in 12 European countries (including the UK). This represents less than 1% of the total population and between 8% and 10% of the non-EU-born population.
- No clear change since 2008 can be detected in the overall size or share of the irregular migrant population across these countries.
- National variation is significant: compared to 2008, estimates appear to have increased in Austria, Germany, and Spain; remained stable in Belgium, France, Italy, the UK, and the United States; and declined in Finland, Greece, Ireland, the Netherlands, and Poland.
- Flow estimates remain scarce, limiting the ability to track short-term changes or assess the impact of policy shifts.
- Even imperfect figures can be useful if their limitations are understood. MIrreM's quality assessment framework offers a way to judge whether an estimate is fit for purpose, while triangulation across indicators helps to verify trends and detect anomalies.
- Linking border apprehensions and asylum decisions shows that 55.4% of apprehensions at EU external borders between 2009 and 2021 concerned people likely to be recognised as refugees or beneficiaries of subsidiary protection.

Innovative data sources and approaches

New methods, digital traces, and register data are expanding the evidence base on irregular migration. They can provide fresh insights but require validation, and in most cases complement rather than replace traditional sources.

- Linking unconventional sources can reveal hidden populations. For example, mortality data in Belgium have been used to estimate the irregular migrant population, and matching consular registers with national records in the United States has shown under-registration of foreign nationals.
- Online search analysis can track migration-related search terms, providing near real-time insight into migration intentions, though these do not always translate into actual movements.
- Germany's Central Register of Foreigners provides longitudinal data on residence histories and status changes, offering insight into pathways out of irregularity such as voluntary or forced return and regularisation.
- Italy's "Signs of Life" approach triangulates tax, school, and utility data to detect likely irregular residence, showing the potential of combining sources.
- In Chile, census data are integrated with administrative data on irregular entry, stay, and student enrolment – an estimation method with potential scalability.
- Informal work remains a blind spot, though Labour Force Surveys and others can be adapted to generate plausible estimates at relatively low cost across contexts.

From data to policy

High-quality data do not automatically translate into better policymaking. Uptake depends on trust, accessibility, and the willingness of policymakers to engage with evidence.

- Policymakers often hesitate to use irregular migration data, seeing it as partial, uncertain, or politically contested. In some contexts this reflects methodological caution, in others it allows symbolic measures to take precedence over substantive policymaking.
- MIrreM presents a framework for improving uptake by harmonising definitions, providing clear documentation, producing user-friendly outputs, building capacity among policymakers, and establishing trusted data-sharing arrangements.
- Examples of effective use include: UK Home Office operational statistics, which provide regular figures with clear guidance; the German Federal Office for Migration and Refugees (BAMF), which uses research collaborations to analyse migrants' living situations; Italy's Regional Observatory on Integration and Multi-ethnicity (ORIM), whose survey outputs inform local integration and labour market policies; and the Mixed Migration Centre's 4Mi dashboards, which provide accessible data on migrant journeys.

Conclusion

This Handbook brings together concepts, findings, methods, and case studies with the aim of offering a clear, practical understanding of irregular migration data. It shows that definitions are being refined, methodologies are advancing, national statistical offices are increasingly engaged, and collaboration between them and researchers is expanding. Ethical considerations are gaining prominence, with more attention to safeguarding rights and preventing misuse. At the European level, Eurostat provides a basis for more systematic reporting, offering a pathway to greater harmonisation and comparability.

Irregular migration will never be fully knowable. Uncertainty is inherent in a phenomenon shaped by mobility, strong incentives to remain "under the radar," and shifting legal and policy frameworks. The aim is not to eliminate uncertainty, but to manage it responsibly, grounding policy and debate in evidence that is as reliable, transparent, and context-aware as possible.

The building blocks for better irregular migration data are already in place. The task now is to connect them, embed them in sustained systems, and invest in their long-term development. This requires leadership at the European level, notably from Eurostat, the Directorate General for Migration and Home Affairs of the European Commission (DG Home) relevant agencies of the European Union, specifically the European Border and Coast Guard Agency (Frontex), the European Union Agency for Asylum (EUAA), and the EU Agency for the Operational Management of Large-Scale IT Systems (eu-LISA), working with national statistical offices, ministries, researchers, and international organisations.

With such coordination, knowledge exchange, and long-term funding, improvements in irregular migration data can be sustained rather than episodic, supporting more informed and credible policymaking.

Preface

Albert Kraler and Denis Kierans

The aim of this Handbook is to provide critical guidance on quantitative measures of irregular migration, with a focus on Europe. Irregular migration is a phenomenon conceptualised in different ways, and that involves a partly hidden and therefore hard-to-count population. Quantifying it involves greater uncertainty and requires more careful interpretation and contextualisation than many other areas of migration statistics.

A central challenge is that the very notion of 'irregular migration' is vague and ambiguous. It is not an 'objective fact', but ultimately a time-bound legal and policy category – a 'social fact', created by human convention (Searle, 2011). The category is closely tied to modern states' role in controlling, shaping and categorising different forms of human mobility.

As a result, irregular migration is difficult to pin down for two key reasons. First, people in an irregular situation are a 'hard-to-reach' and often hidden population. Although, as this Handbook shows, many are in fact captured by official data, they nonetheless remain a 'hard-to-count' group. Second, the categories used to describe migrants in an irregular situation – or more broadly, migrants with a precarious residence status – are themselves contested. They are disputed not only politically and socially, but also in administrative and legal practice.

Who counts as having crossed a border irregularly, or as unlawfully staying within a given territory, is not a straightforward question. The answer depends on complex assessments made by states and other actors, each with their own interests in shaping how the term and its consequences are defined.

The difficulty of measuring irregular migration is therefore at least twofold. First, it is a conceptual and epistemological challenge. What exactly is being measured? On what bases are categories defined? And, not least of all, what types of knowledge shape what is being measured?

Second, it is a methodological challenge. How can particular conceptualisations of irregular migration be measured, especially when direct observation is not possible? What approaches are available to estimate different aspects of irregular migration, depending on the context and data sources? And to what extent can existing statistical indicators tell us something meaningful about its scale and dynamics?

Yet despite these challenges, the key message of this Handbook is that it is possible to make scientifically sound statements about the quantitative dimensions of irregular migration, and to tackle the conceptual, epistemological and methodological issues it presents. This requires efforts from all those involved in producing, using and disseminating these data. In particular, it depends on bridging the gap between conceptual debates, the practical demands of measurement, and the policy decisions that shape and are shaped by them.

The aim of this Handbook is to support that process. It does so through a series of focussed chapters, complemented by short textboxes that illustrate concrete examples of practice. Taken together, these contributions set out practical options for addressing the quantitative dimensions of irregular migration, while acknowledging the limitations and uncertainties that inevitably remain.

The idea behind the Handbook

This Handbook was developed as part of the Horizon Europe Project "Measuring Irregular Migration and Related Policies" (MIrreM)[1]. It is one of two Handbooks produced by the project, the other focusing on regularisation policies (Ahrens et al. 2025).

Both of the MIrreM Handbooks were inspired by earlier guidance produced by the Expert Group on Refugee, IDP and Statelessness statistics (EGRISS)[2]: one on refugee statistics and another on statistics on internally displaced persons (European Commission and Eurostat, 2018; European Commission, Statistical Office of the European Union and United Nations Organisation, 2020). The EGRISS Handbooks were eventually drafted by a core writing team at the secretariat, but were the product of a multi-year process, drawing on a group that included experts from national statistical offices, international organisations and Eurostat, as well as a range of independent experts.

Taking inspiration from this approach, MIrreM envisaged two working groups: one on irregular migration data and another on regularisation. Each was to involve a core of committed members from relevant organisations, supported by a broader network of stakeholders who would provide guidance along the way. In practice, however, it proved more productive to work on this Handbook with different expert communities at different points in the project. Rather than creating standing working groups, we engaged partners flexibly, working more closely with them when collaboration was most practical and productive.

Stakeholders consulted included academics developing methodologies and approaches to measuring irregular migration, as well as those critically engaging with categorisations and quantitative measurements. Civil society organisations and International Organisations contributed important insights from research and the field. National Statistical Offices (NSOs) played a significant role throughout the MIrreM project, reflecting their role in producing population estimates that account for the entire resident population, including irregular migrants. Immigration authorities and their affiliated research centres took part, drawing on their operational knowledge and expertise in migration data analysis.

Engagement at the European level was key, given the role of EU institutions in shaping migration policy and data collection and use. It included Eurostat, the Joint Research Centre of the European Commission and its Knowledge Centre on Migration and Demography (KCMD)[3], the European Union Agency for Fundamental Rights (FRA), the European Border and Coast Guard Agency (FRONTEX), and the European Union Agency for Asylum (EUAA) as well as individual experts from the European Commission's Directorate-General for Migration and Home Affairs (DG Home).

These stakeholders were involved in MIrreM's work from the outset, including through a series of workshops with national and local stakeholders in partner countries and several workshops at the European level.

1 More information on the project, including all its publications is available at www.irregularmigration.eu.

2 https://egrisstats.org

3 The Joint Research Centre embarked on a related project on measuring the effectiveness of return policy, drawing also on insights from MIrreM (Belmonte et al. 2025).

Concrete work on this Handbook began with a workshop in Brussels in April 2024, which brought together around 50 stakeholders to discuss and refine the concept. Exchanges continued through 2024 and 2025, including at the 2024 IMISCOE conference and the 2025 International Forum for International Migration Statistics in Malmö. They also included consultations with the United Nations Economic Commission for Europe (UNECE)'s statistical division and task forces under the Conference of European Statisticians (CES), Eurostat's task force on implementation of guidelines for a harmonised population base.

For this Handbook, the MIrreM team reached out to a wide range of experts to contribute chapters and textboxes. The strong response to these invitations reflects, we hope, a shared commitment to advancing discussions on the quantitative aspects of irregular migration, and to linking what is possible in practice with policy and decision-making.

The remainder of this Handbook is structured into twelve core chapters, complemented by a Foreword, an Executive Summary, and a series of thematic textboxes. It moves from definitional and ethical questions to methodological approaches and practical applications, offering researchers, policymakers, and practitioners a practical resource for navigating one of the most complex and politically sensitive areas of migration data.

Below is a brief overview of the Handbook by section:

Chapter 1: Introduction – Making the case for better data on irregular migration

Introduces the motivation behind the Handbook, describing common problems in how irregular migration is defined, counted, and debated, and setting out the case for a more structured and reflective approach to data.

Chapter 2: What is irregular migration?

Explores legal and administrative definitions of irregularity and presents the MIrreM taxonomy as a tool to differentiate between forms of irregular migration across stocks and flows.

Chapter 3: Ethics and Data on Irregular Migration

Examines the ethical risks associated with collecting and using data on irregular migration and proposes safeguards to reduce harm, ensure accountability, and promote responsible data practices.

Chapter 4: What are good quality data on a phenomenon that is hard to measure?

Defines what quality means in the context of irregular migration data and introduces MIrreM's assessment framework, focussed on transparency, accessibility, comparability, and the handling of uncertainty.

Chapter 5: Innovations in methodological approaches to estimate irregular migrant stocks and flows

Reviews key methods for estimating irregular migrant populations, including their assumptions, strengths, and limitations, and offers guidance on when and how each can be applied.

Chapter 6: Data Traces and the Inevitable Visibility of Irregular Migration

Analyses how irregular migrants appear in conventional and alternative data sources, challenging the assumption of invisibility and highlighting how visibility is shaped by institutional and legal contexts.

Chapter 7: Register data sources on migrant stocks

Assesses how administrative registers, such as Germany's AZR and Spain's padrón, can help derive indicators of irregular residence, while also noting gaps, biases, and data quality challenges.

Chapter 8: Getting into the flow - what do we know now, 15 years since CLANDESTINO?

This chapter takes stock of how irregular migration flows are measured, noting changes over time in the availability and accessibility of flow indicators, particularly at EU level, but also persistent challenges related to validity, scope, and interpretation.

Chapter 9: Irregular migration and informal work

Proposes a method to estimate the overlap between irregular residence and informal employment using labour force survey data.

Chapter 10: Surveying irregular migrants: challenges and approaches

Reviews strategies for including irregular migrants in survey research, including regularisation surveys, retrospective trajectory data, and targeted sampling approaches.

Chapter 11: Towards the More Effective Use of Irregular Migration Data

Explores how institutional, legal, and political factors shape the use of irregular migration data, identifying key barriers to uptake.

Chapter 12: Progress, limits, and the need for sustained effort

Summarises the Handbook's core insights and outlines practical steps to improve the production, interpretation and application of irregular migration data across Europe.

References

Ahrens, J., Kraler, A., Legarda, I. & Levoy, M. (eds) (2025), Handbook on Regularisation Policies: Practices, Debates and Outcomes. Krems: University of Krems Press. https://doi.org/10.48341/chqk-ey86.

Belmonte, M., Pingsdorf, J., Cortinovis, R., Nedee, A. & Tintori, G. (2025). Managing Irregular Migration and Return: Lessons from National Practices. Luxembourg: Publication Office of the European Union.

European Commission and Eurostat (2018). Expert group on refugee and internally displaced persons statistics – International recommendations on refugee statistics. Luxembourg: Publications Office of the European Union. https://data.europa.eu/doi/10.2785/517815.

European Commission. Eurostat. and United Nations Organisation (2020) International Recommendations on Internally Displaced Persons Statistics (IRIS): March 2020. LU: Publications Office. https://data.europa.eu/doi/10.2785/18809.

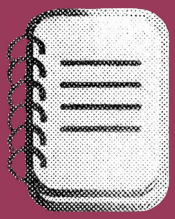
Searle, J. R. (2011): Making the Social World: The Structure of Human Civilization. Oxford: Oxford University Press.

Chapter 1

Introduction – Making the case for better data on irregular migration

Denis Kierans and Albert Kraler

Introduction – Making the case for better data on irregular migration



Key points

- Irregular migration data are often of low quality and misinterpreted by those who use it, when they are used at all. This Handbook provides clear and easy-to-understand guidance on how to improve the quality of these data and an understanding of them.
- Despite ongoing public and political interest, there has been a notable lack of investment in improving the methods and capacities for generating irregular migration estimates, particularly in European National Statistical Offices (NSOs). By synthesising key findings from the MIrreM project and highlighting good practices and promising innovations, this Handbook seeks to help bridge that gap.
- Ultimately, this Handbook makes the case that these challenges can only be met – and much risk mitigated in the process – through strengthened leadership on, coordination around and long-term investment in a Europe-wide infrastructure capable of producing, disseminating and fostering responsible and appropriate use of irregular migration data

Irregular migration is a subset of overall migration, typically only making up a small share of migration stocks and flows. Yet it warrants individual attention when it comes to methods for collecting, analysing and using data. We present this Handbook to assist ongoing and future efforts with the hope that it adds to this field of research in three ways.

The first is by bringing much needed guidance on interpreting different types of data on irregular migration, some of which are delivered on an almost daily basis in policy debates and the media. These data may be border apprehensions, interceptions at sea, deportations, or migrant deaths. Often, they are released and reported on without much contextual information and lack detail about their quality, the assumptions that underpin them and what the data actually show (Kraler & Reichel, 2022). Definitions are fuzzy, terms are conflated. Flow data may be presented as stocks, or vice versa.

From time to time, estimates on the number of irregular migrants present in a particular country, a group of countries, or another area make it to headline news. Again, this is typically with little attention to the quality of the estimate or the context in which the estimate was produced, such as the population group covered, the reference year or the methodology used. Sometimes these data are specifically collected to inform policy debates in

response to the presence (or perceived presence) of irregular migrants in a specific area.

When data are used in these debates, it matters not just whether they are accurate, but whether they are well understood and used appropriately. As discussed in Chapter 11, the responsible use of migration data depends as much on interpretation and communication as on technical quality. This Handbook is intended to support both: offering tools for better measurement and clearer thinking about what these numbers do and do not tell us.

In short, the production of these data and their use in different types of debates are here to stay. We hope this Handbook brings clarity to some of these recurring policy, operational and social challenges. Many of the technical problems that emerge in this area – e.g., small populations, partial visibility, reliance on administrative proxy data instead of or in addition to traditional data sources – are shared by those working on other 'hard-to-reach' groups. As such, the insights offered here may also be relevant to researchers and practitioners working in related areas.

Box 1.1: A history of interest: Irregular migration data in Europe

Albert Kraler and Denis Kierans

Irregular migration has been an issue of high salience in Europe since at least the 1990s, when migration flows to Western Europe surged following the collapse of Communist regimes in Eastern Europe and the displacement following the violent break-up of Yugoslavia. Beyond these major turning points, the primary receiving countries in Europe had already experienced a longer-standing increase in asylum-related inflows from beyond Europe, traditionally the main source of refugees in Europe. At the same time, legal migration increased considerably, facilitated in Europe by freedom of movement policies in the European Union and the Eastern enlargement. These developments fuelled a broader interest in migration, which in turn led to increased efforts to improve migration statistics at the national, European-wide and global levels (Kraler, Reichel, & Entzinger, 2015).

The political interest in irregular migration also went hand-in-hand with more systematic administrative data collection, such as on apprehensions, smuggling and deportation. At the European level, the first such effort was the Centre for Information, Discussion and Exchange on the Crossing of Frontiers and Immigration (CIREFI) data collection initiative, launched in 1996 to support, and initially conducted on a confidential basis (See Kraler & Jandl, 2006). This formed the basis for the Enforcement of Immigration Legislation (EIL) Statistics collected by Eurostat under the 2007 Regulation on Migration Statistics. Yet there was also growing interest in irregular migrants who have not come into contact with state authorities, but constituted an important part of the migrant population – and workforce – especially in Southern EU Member States.

A study commissioned by the European Commission in 1991 appears to have been the first to examine the scale of the irregular migrant population in a European comparative perspective (Werth & Körth, 1991). Another study commissioned by Eurostat a few years later placed greater emphasis on conceptual and methodological aspects (Delaunay & Tapinos, 1998). In some ways, this laid the groundwork for the first systematic European effort to collect, assess and produce estimates for a larger number of European countries, and to elaborate an estimate of the overall irregular migrant population in the EU as a whole: the CLANDESTINO project (CLANDESTINO, 2009).

Building on this foundation, MIrreM refines the CLANDESTINO methodology and adds important new elements. One is an exploration of innovative methods (Chapter 5). Another is a sustained effort to involve relevant stakeholders, raising awareness about the opportunities and limitations of data on irregular migration and encouraging more better practices in collecting, analysing and using these data.

Other recent initiatives address this issue as well. At the European level, negotiations on a new Regulation on Population and Housing Statistics have led to the creation of a "task force on implementation guidelines for a harmonised population base", which considers irregular migrants alongside other hard-to-count groups. Under the Conference of European Statisticians (CES), hosted by the United Nations Economic Commission for Europe (UNECE), two further task forces – 'Measuring Hard-to-Reach Groups in Administrative Sources'[1] and 'Defining and Measuring New Forms of International Migration'[2] – have collected practices from NSOs on how to account for irregular migrants in population statistics, including methodological approaches (UNECE, 2025).

References:

CLANDESTINO. (2009). Undocumented migration: Counting the uncountable. Data and trends across Europe (Final report). Foundation for European and Foreign Policy (ELIAMEP). https://emnbelgium.be/sites/default/files/publications/CLANDESTINO-final-report.pdf

Delaunay, D., & Tapinos, G. (1998). La mesure de la migration clandestine en Europe: 1. Rapport de synthèse 2. Rapport des experts. Eurostat. https://www.documentation.ird.fr/hor/fdi:010016012

Kraler, A., & Jandl, M. (2006). Statistics on refusal, apprehensions and removals: An analysis of the CIREFI data. In M. Poulain, N. Perrin, A. Singleton, & THESIM Project (Eds.), THESIM: Towards harmonised European statistics on international migration (pp. 271–285). Presses universitaires de Louvain.

Kraler, A., Reichel, D., & Entzinger, H. (2015). Migration statistics in Europe: A core component of governance and population research. In P. Scholten, H. Entzinger, R. Penninx, & S. Verbeek (Eds.), Integrating immigrants in Europe (pp. 39–58). Springer. https://doi.org/10.1007/978-3-319-16256-0_3

United Nations Economic Commission for Europe. (2025). Measuring hard-to-reach groups in administrative sources (Draft). Conference of European Statisticians, seventy-third plenary session, Geneva, 16–18 June 2025. https://unece.org/sites/default/files/2025-04/HardToReachGroups%20Consult.pdf

Werth, M., & Körth, H. (1991). Immigration of citizens from third countries into the southern member states of the EEC: A comparative survey of the situation in Greece, Italy, Spain, and Portugal. Office for Official Publications of the European Communities.

The second aim of this Handbook is to bring attention to the need for more high-quality European-focussed research on irregular migration data. Indeed, there has been a relative neglect in Europe of the irregular migrant population by demographers in general and by National Statistical Offices (NSOs) specifically.

While irregular migration data 'suddenly were everywhere'[3], irregular migrants as a population group and subject of demographic analysis are conspicuously absent from the statistical work of

1 https://unece.org/statistics/documents/2023/11/working-documents/terms-reference-task-force-hard-reach-groups

2 https://unece.org/statistics/documents/2024/02/working-documents/task-force-defining-and-measuring-new-forms

3 Paraphrasing Kathleen Newland's (2010) observation on the ascendancy of migration as a key concern on the international level in the early 2000s.

many European governments, despite their salience in public and political discourse. As a result, our knowledge on the demography and socioeconomics of irregular migrants is limited, and often biased. In contrast, demographers and sociologists in the United States have for decades produced regular estimates of the irregular migrant population and its demographic characteristics, supported in part by the availability of population-wide surveys (see chapter 5). The US also has a higher proportion of irregular migrants relative to its total (3% in 2022) and foreign-born (25%) population than European countries which in some respects makes these estimation exercises more feasible (Kierans & Vargas-Silva, 2024).

That said, the relatively small size of the irregular migrant population in Europe – estimated at less than 1% of the total population and between 8% and 10% of the total foreign-born population since 2008 (see Chapter 4)[4] – should not be interpreted as grounds for inaction. Nor is Europe devoid of good practice, as evidenced by the many case studies featured this Handbook. However, compared to the US, quantitative estimates of irregular migration are infrequent, and – with a few exceptions – limited to assessing the overall scale of the irregular migrant population, with limited to no detail about demographic or socioeconomic characteristics.

Part of the reluctance within European statistical institutions may reflect discomfort with publishing estimates that carry high uncertainty and diverge from the conventions of register and census based population statistics. But this caution comes with its own risks. In the absence of official figures, governments leave a vacuum that can be filled by unreliable or agenda-driven figures. These numbers can have outsized influence, and may, ironically, further discourage NSOs from stepping into the debate and improving the state of the art.

We hope this Handbook encourages more NSOs to take up efforts to produce reliable and well communicated information on irregular migration. To this end, this Handbook introduces several approaches that may be helpful, including capture-recapture methods, model-based simulations, residual estimation, and innovative uses of administrative irregularities.

The third contribution of this Handbook is building a case for investment into the infrastructure needed to support long-term improvements in irregular migration data. A recurring theme across MIrreM's work is that improving irregular migration data and their use is not only a technical matter, but an institutional one as well (see Chapter 11).

Eurostat, has already taken important steps coordinating irregular migration flow data through its enforcement of immigration legislation (EIL) statistics. It is well positioned to play a European-wide convening role around stock estimates. Longer-term funding, knowledge exchanges between NSOs and researchers, annual national updates and standardised reporting templates are all relatively low hanging fruits, which have the potential to transform the irregular migration data landscape in Europe for the better.

Conclusion

This Handbook does not offer a blueprint for improving the quality and use of irregular migration data in every context. But it does offer tools, examples and a case for long-term investment in the infrastructure needed to produce, disseminate and support the responsible use of these data. In doing so, we hope to reduce the risks of misuse, foster greater consistency and transparency, and ultimately improve the capacity of governments and institutions to engage meaningfully with one of the most contested issues in European migration policy.

4 The foreign-born population excludes those born in countries covered by free movement agreements.

References

Kierans, D. & Vargas-Silva, C. (2024). The Irregular Migrant Population of Europe. MIrreM Working Paper No.10. Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.13857073

Kraler, A., & Reichel, D. (2022). Migration statistics. In P. Scholten (Ed.), Introduction to migration studies (pp. 439–462). Springer. https://doi.org/10.1007/978-3-030-92377-8_27

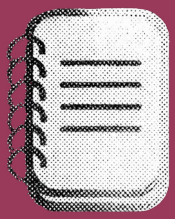
Newland, K. (2010). The governance of international migration: Mechanisms, processes, and institutions. Global Governance: A Review of Multilateralism and International Organizations, 16(3), 331–343. https://doi.org/10.1163/19426720-01603004

Chapter 2

What is irregular migration?

Albert Kraler, Tuba Bircan and Ann Singleton

What is irregular migration?



Key points

- This chapter explores how 'irregular migration' is defined and why the concept is contested, showing the tension between using existing categories for measurement and critically interrogating them.
- It highlights that terms such as 'irregular', 'illegal', or 'undocumented' are not neutral but historically and politically charged.
- The chapter explains that 'irregular migration' may denote different phenomena, legal status, border crossings, or policy violations, and stresses the need for precise definitions.
- It shows that irregularity is not fixed but shaped by laws, administrative practices, and political contexts, varying between states and over time.
- Understanding irregular migration requires both snapshots of populations and trajectories of status change. The MIrreM taxonomy maps pathways into and out of irregularity, while making visible the limits of classification.

This chapter addresses an irresolvable challenge: how to discuss 'irregular migration' in a reflexive way, whilst necessarily using language and terminology that reproduces contested narratives and categories. It is in itself an area of study in need of the 'demigranticization' advocated by Dahinden (2016). The chapter addresses this challenge by exploring how 'irregular migration' is conceptualised, used, and measured and by proposing an approach that allows quantifications without falling into the pitfall of reifying problematic categories.

At first glance, it appears to describe a clearly defined phenomenon, often equated with 'undocumented', 'clandestine', 'unauthorised', 'unlawful' or 'illegal' migration (see on the terms used Box 2.1 below). Yet, in practice, the term is used in divergent and often ambiguous ways. It features prominently in academic, policy, and media discourse, but rarely with consistent meaning.

Box 2.1: "Words matter"[1] – Terms used to describe irregular migration.

Albert Kraler

A wide range of terms are used to speak about 'irregular migration'. Until the 2010s, 'illegal migration' — and related terms such as 'illegal migrant' or 'illegal alien' (the latter predominantly used in the United States) — were the most widely used (Paspalanova, 2008) . In years since, they have been criticised for their association with criminality and the harmful effects of these associations on migrants (PICUM, 2017). As far back as 1975, a UN General Assembly Resolution recommended that UN agencies instead use terms like 'non-documented' or 'irregular migrant workers', reflecting the focus at the time on labour migration. Similar language has been adopted in subsequent international fora.

The 1994 International Conference on Population and Development in Cairo spoke of 'undocumented or irregular migrants', while the International Labour Conference and later the European Union Agency for Fundamental Rights (FRA) used terms such as 'irregular status' and 'migrants in an irregular situation' to draw attention to the legal dimension of status without stigmatising individuals (European Union Agency for Fundamental Rights, 2011). In 2009, the European Parliament passed a resolution calling on EU institutions and Member States to stop using the term 'illegal immigrants', pointing to its negative connotations, and instead to refer to 'irregular' or 'undocumented' migrants or workers. Since then, the European Commission has started to use 'irregular migration', although 'illegal migration' is also still used. In legal contexts, the more precise term 'unlawful entry and stay' is used at the European level (European Migration Network 2025). Other language, such as 'clandestine migration', 'clandestine migrants', 'unauthorised migration' and 'unauthorised migrants' remain in circulation. Unlike most terms that either convey a negative or neutral connotation, the French term 'sans papier' (migrants without papers) has a pro-migrant and activist connotation, reflecting the lasting legacy of the French sans-papier movement of the 1990s (Freedman, 2008).

Researchers have drawn attention to a shift in media and academic discourse towards terms seen as more neutral, such as irregular, undocumented[2] or unauthorised — as opposed to clandestine or illegal (Spencer & Triandafyllidou, 2022:192). Yet even these more neutral terms are contested. Their meanings and uses can shift over time, particularly when they become politicised. For this reason, it is important to use terms with care. While they may appear straightforward, their meanings are not fixed and can vary depending on the context and audience.

In this Handbook, 'irregular migration' or ' irregular migrants' are used, as the currently most widely used terms. Whilst considered the most neutral terms (cf. Squire 2010:4), their use still reproduces narratives that in themselves 'irregularise' and 'other' people defined as migrants. 'Migrant irregularity' is used when referring to the condition of lacking a legal status (cf. Chauvin Garcés-Mascareñas 2012 speaking of "migrant illegality" in a similar way). We recognize that migrant irregularity is not a fixed trait, but is produced by state driven processes, captured by the term 'irregularisation'. We prefer the term 'irregular migrant' or 'irregular migration', as the Handbook is primarily concerned with the measurement of outcomes of processes of irregularisation. This said, we are also interested in processes, even in the more limited perspective of legal status trajectories, that is changes of legal status over time (see in particular chapter 7 and 10 for such perspectives).

1 "Words matter" was the motto of a campaign by the Platform for International Cooperation and Undocumented Migrants (PICUM) launched in 2010, see https://picum.org/words-matter-2/.

2 In the United Kingdom, for instance, the term 'undocumented' has come to be seen as appropriate in the context of the Windrush scandal, which involved people with a legal residence, but no documentation to prove it (The authors thank Peter Walsh for this observation).

Importantly, terms such as 'irregular', 'illegal' or 'undocumented' are not neutral descriptors. They are embedded in historical legacies of statecraft, border control, and postcolonial governance. The production of migration categories has often served to reify racialised boundaries of belonging and to legitimise differential access to rights. Even more 'neutral' alternatives, such as 'undocumented' must be used reflexively, acknowledging that terminology can both reflect and reproduce the hierarchies it seeks to name.

References

De Genova, N. P. (2002). "Migrant 'Illegality' and Deportability in Everyday Life." Annual Review of Anthropology 31(1):419–47. https://doi.org/10.1146/annurev.anthro.31.040402.085432.

European Migration Network. (2025). "Irregular Migration." EMN Asylum and Migration Glossary. https://home-affairs.ec.europa.eu/networks/european-migration-network-emn/emn-asylum-and-migration-glossary/glossary/irregular-migration_en.

European Union Agency for Fundamental Rights (2011). Fundamental Rights of Migrants in an Irregular Situation in the European Union - Comparative Report. European Union Agency for Fundamental Rights. https://fra.europa.eu/sites/default/files/fra_uploads/1827-FRA_2011_Migrants_in_an_irregular_situation_EN.pdf.

Paspalanova, M. (2008). "Undocumented vs. Illegal Migrant: Towards Terminological Coherence." Migraciones Internationales 4(3):79–90. https://www.scielo.org.mx/scielo.php?script=sci_abstract&pid=S1665-89062008000100004&lng=es&nrm=iso&tlng=en.

PICUM (2017). "Words Matter. Why 'Undocumented' or Irregular'." https://picum.org/wp-content/uploads/2023/08/Words_Matter_Terminology_FINAL_March2017.pdf.

Spencer, S., & Triandafyllidou, A. (2022), "Irregular Migration." Pp. 191–204 in Introduction to Migration Studies, IMISCOE Research Series, edited by P. Scholten. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-92377-8_12.

Squire, V., ed. (2010). "The Contested Politics of Mobility: Politicizing Mobility, Mobilizing Politics." Pp. 2–25 in The Contested Politics of Mobility. Routledge. https://doi.org/10.4324/9780203839829.

We use the term here not to suggest it is an analytically clear category, or to affirm its supposed normative neutrality, but because of its widespread institutional use across statistical, legal, and public domains. Crucially, we do not treat irregularity as a fixed category. Rather, we understand it as a politically constructed condition, shaped by legal frameworks, administrative decisions, enforcement practices, and the broader discourses that surround migration.

Different actors invoke the term to refer to different things. Demographers may use it to denote population segments not captured in official records. Lawyers focus on violations of entry or stay conditions. Policymakers and civil servants use it to delineate eligibility for return or regularisation programmes. Meanwhile, politicians and media figures often mobilise the term symbolically, to invoke crisis, disorder, or humanitarian need. These usages reflect not only different operational logics but also shifting political agendas. They reproduce the concept and narratives without necessarily clarifying or explaining the material realities of human mobility.

Public perceptions of irregular migration often draw on powerful visual imaginaries. A simple image search using the term 'illegal migration'[3] yields predictable results: crowds of racialised individuals, people in boats or crowded along border fences. Search for 'refugees' or even 'migration' show similar patterns, the resulting images being little different. Results for 'migration' are perhaps a little more varied but still show significant overlaps. These representations align with what De Genova (2002) called the 'border spectacle', a dominant imaginary in which different legal categories are blurred, and irregular migration is primarily associated with visible, racialised mobility. Recent studies confirm that media and political narratives often conflate terms like 'asylum seeker', 'refugee' and' irregular migrant', reinforcing a stereotypical figure of the 'unauthorised other' (Rheindorf et al., 2025; Smellie & Boswell, 2024).

Such narratives are not only visual and discursive, but are embedded in how data are produced, reproduced and communicated. As discussed in Chapter 1 of this Handbook, widely cited indicators such as apprehension figures or asylum statistics are shaped by a logic of securitisation and humanitarian concern. These indicators, even when presented as neutral metrics, contribute to the framing of irregular migration as a problem of control or protection.

The ambiguity of the term 'irregular migration' also spills over into legal, policy and scientific debates. Its everyday use tends to blur important distinctions, applying casually to a wide range of situations. In scientific contexts, this ambiguity is particularly problematic, as it undermines efforts to define and measure the 'phenomenon' with precision. Any attempt to quantify irregular migration therefore requires clear conceptual foundations and careful definitions that acknowledge the fluidity and diversity of legal statuses

This chapter aims to provide such a foundation. It does so in three steps. First, it introduces three demographic and sociological concepts: flows, stocks and trajectories, that help us understand how legal status changes over time. Second, it examines how migrant irregularity can be defined and classified. Third, it presents the MIrreM conceptual framework: a structured taxonomy for identifying and analysing different 'classes' of migrants in precarious legal situations and for tracing the pathways into and out of irregularity.

Flows, stocks and trajectories

In narrow administrative terms, 'irregular migration' is typically defined as the cross-border movement or presence of individuals outside the authorised channels established by states. However, this view risks oversimplifying what is, in reality, a complex and often reversible condition. Migrants do not simply enter or remain 'irregularly'; rather, they may move through a range of legal statuses over time, including lawful, provisional, suspended, or ambiguous forms of stay. Legal status is fluid, and its boundaries are shaped by administrative decisions, legal uncertainty, and institutional practices.

To account for this complexity, this chapter adopts an analytical framework based on three interrelated concepts: flows, referring to movements into or out of a legal status; stocks, denoting the population with a given status at a particular point in time; and trajectories, which capture transitions across different statuses over time. These are familiar terms in demographic and statistical analysis, yet they are far from neutral. The very notion of a stock of irregular migrants, for example, draws on a population logic historically rooted in the nation-state and has been critiqued for its tendency to objectify migrants and reinforce racialised imaginaries of mobility.

3 The search was performed using a 'private' browser window to exclude user specific results.

In line with such demographic framing, the IOM Glossary of Migration defines irregular migration as the "movement of persons to a new place of residence or transit that takes place outside the regulatory norms of the sending, transit and receiving countries" (Sironi, Bauloz, & Emmanuel, 2019). But the term is also frequently used to describe the presence of migrants who have entered regularly and later lost their legal status, raising the question of where irregular movement ends and unauthorised residence begins. This distinction is far from trivial. When do "people on the move" cease to be on the move (Pijnenburg & Rijken, 2021)? Legal status further complicates the picture, especially when migrants' status changes not as a result of movement, but due to administrative actions, procedural delays, or expiration of documentation.

Moreover, irregular movement itself does not automatically imply irregularity; conversely, lawful movement does not necessarily imply lawful stay. For example, migrants may enter a country on a visa or permit for work, study, or tourism, and later overstay or breach the conditions attached to their stay. Others may cross borders without authorisation but subsequently apply for asylum and be granted protection. In both cases, the boundaries between regular and irregular status are blurred.

The concept of mixed migration articulates this complexity in the context of asylum related migration (see Box 2.2 on mixed migration, below).

Box 2.2: Applying a mixed migration lens to irregular migration

Roberto Forin

The term mixed migration emerged in the 1990s amid growing attention to the Migration-Asylum Nexus and was later adopted as a policy concept during the Global Consultations on International Protection launched by UNHCR in 2000. It was introduced to better capture the reality of overlapping refugee and migratory flows that defy clear-cut categorisation and often move along the same routes using similar means (Van Hear, 2011). These dynamics challenge states' and mandated UN agencies capacity to apply distinct legal and policy responses and risk leaving people in mixed flows without adequate protection if they do not neatly fit into established legal categories.

MMC definition of mixed migration

According to the Mixed Migration Centre, mixed migration refers to cross-border movements of people including refugees fleeing persecution and conflict, victims of trafficking and people seeking better lives and opportunities. Motivated to move by a multiplicity of factors, people engaged in mixed migration have a range of legal statuses as well as a variety of vulnerabilities. Although entitled to protection under international human rights law, they are exposed to multiple rights violations along their journey. Mixed migration describes migrants travelling along similar routes, using similar means of travel—often travelling irregularly and wholly or partially assisted by migrant smugglers.

Why is a mixed migration lens useful for understanding irregular migration?

The concept of mixed migration is essential for understanding the complexity of irregular and onward migration today, including in the European context. It highlights the reality that people rarely move for just one reason—such as war, economic reasons, or environmental stress—but for a combination of factors that are often intertwined. These motivations defy neat categorisation and reflect the complexity of contemporary mobility.

From a protection perspective, a mixed migration lens shows that regardless of their legal status—whether they are refugees, asylum seekers, or undocumented migrants—people on the move are forced to travel using irregular means and often face similar risks and vulnerabilities. These may include violence, exploitation, detention, and trafficking.

A mixed migration perspective also challenges simplistic distinctions between "forced" and "voluntary" migration. Not all people who are forced to move are entirely without agency, just as those who begin their journeys voluntarily may lose agency along the way. Recognising this continuum between choice and constraint helps us understand the lived realities of migration more fully.

Finally, while it is essential to fully acknowledge the specific rights of refugees under the 1951 Refugee Convention and its 1967 Protocol, the mixed migration lens places equal emphasis on the rights and protection needs of all people on the move, regardless of status. This inclusive framing is vital for developing research, policies, and interventions that reflect the complex and evolving nature of contemporary mixed and irregular migration.

References

Van Hear, N. (2011). Mixed Migration: Policy Challenges. Oxford: Migration Observatory. https://migrationobservatory.ox.ac.uk/resources/primers/mixed-migration-policy-challenges/

One way to approach these complexities is through the 'population balance model' also referred to as or 'demographic accounts' which distinguishes between in- and outflows and the resulting population stocks. As Figure 2.1 illustrates, this model provides a structured way to conceptualise how individuals move into and out of irregularity, whether through border crossings, overstays, regularisations, or status loss. While inherently simplified and limited by the reproduction of problematic concepts this approach offers an important tool to expose complexities and to clarify the discussion: linking legal definitions, administrative data, and demographic analysis in a more coherent and transparent manner.

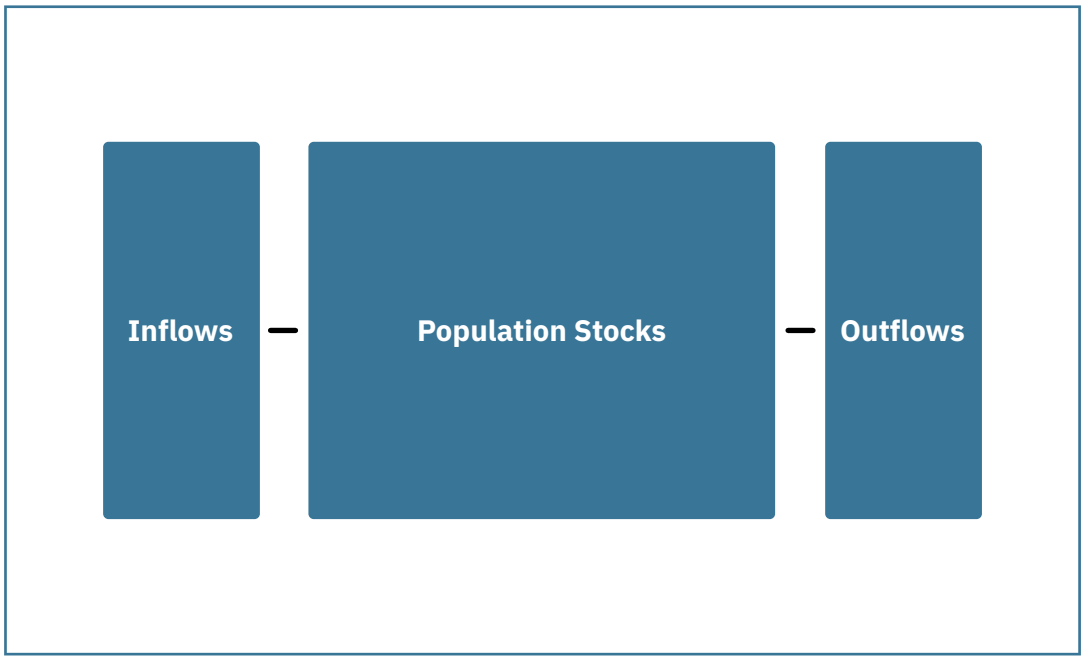


Figure 2.1: Population stocks and flows

Inflows increase population stocks while outflows decrease them. Importantly, a stock and flow perspective focuses on a particular territory (usually a country) and requires a precise definition of the population to be measured, which in turn dictates the flows that are to be considered (and those which are to be disregarded). Time is a key aspect here that determines when an inflow becomes part of the population stocks, or conversely, when an outflow is considered to reduce the stocks.[4] In the context of international migration, the conventional time criterion for long-term migration is one year (Kraler & Reichel, 2022), although many countries also include temporary migrants in their national population estimates, that is, migrants that have been staying for at least 3 months but less than a year. In relation to irregular migrants, their legal status adds another layer of complexity: Even if they need to be considered part of the resident population according to statistical conventions if they meet the residence requirements, they do not form part of the 'de jure' population – that is, [define 'de jure']. Change of legal status in turn constitutes also a specific type of flow, complementing natural population movements (births and deaths) and migration (in- and outmigration) as main pathways in and out of the population of migrants in an irregular situation.

Yet many of the regularly collected indicators on irregular migration lack any specification on duration of stay. In a similar vein, oft-used flow indicators record only events (such as a detected irregular border crossing), but do not link those events to a specific person in a given period of time. Both aspects make available measures problematic as measures of population dynamics and lead to higher uncertainty. In relation to irregular migrants, there are also different pathways (see Box 2.3) into irregularity, making an account of population movements even more complex. We will revert to this model of the 'population balance' in relation to irregular migrants further below. Suffice it to note here that flow indicators do not necessarily relate to stocks in this context, but even then the model of the population balance helps to clarify population dynamics and the type of robust statements that can – or cannot – be made.

Box 2.3: Pathways in and out of irregularity

Albert Kraler

Flows into and out of irregularity can also be conceptualised as pathways into and out of irregularity. This terminology moves away from a demographic conception and highlights the process of becoming, or 'unbecoming' irregular.

Individuals can become part of the population of migrants in an irregular situation by birth (a demographic flow), through irregular entry (a geographic flow), or by loss of a residence status , including (lawfully staying) asylum seekers, whose claim is rejected (a status-related flow). Similarly, individuals cease to be part of the population of migrants in an irregular situation by death (a demographic flow), outmigration (return or onward migration, both geographic flows) or by acquisition of another legal status , for example through regularisations (a status related flows).

The 'population balance' is a static concept: it allows for the definition of population stocks and in- and outflows within a given time period. Yet as scholarship on migrant irregularity has emphasized, irregularity is not a "static condition, but a dynamic space" in which the legal status is negotiated (rephrasing Chauvin Garcés-Mascareñas, 2012, 253 see also; Kraler & Ahrens, 2023, 21f). Other scholars have described migrant irregularity as 'fluid' (see for example Triandafyllidou & Bartolini, 2020). This dynamic and 'fluid' nature can only be captured by explicitly considering legal status trajectories over time (cf. Beauchemin, Descamps, Dietrich-Ragon, 2023; Descamps, 2024; Goldring, 2022; Jasso et al., 2008, see also chapter 7). A trajectory perspective sheds light on changes of legal status over time, on pathways into and out of irregularity as well as repeated cycles of irregularity and how this is linked to (im)mobility, employment or housing trajectories, or indeed other aspects. A trajectory

4 See the UN Recommendations on Statistics on International Migration and Temporary Mobility on using this framework for the collection of migration statistics more generally (United Nations Secretariat. Department of Economic and Social Affairs. Statistics Division 2025).

perspective also helps to overcome the limitations of a "presentist" perspective. For example, Jasso et al. (2008) were able to demonstrate that almost a third of all persons granted permanent residence in the United States in 2016 (around 900,000 persons) had experienced periods of irregularity previously, suggesting a considerable regularity of irregularity, but above all demonstrating the extent to which irregular migrants were able to regain a legal status even in the absence of an explicit regularisation policy.

Yet we also acknowledge that our analytical framework misses important quantifiable aspects of irregular migration, that are nevertheless relevant to assess policies addressing irregular migration. The issue of migrant deaths in transit is a case in point: while not relevant to describe the population of migrants with precarious legal status, and pathways into or out of a precarious legal status in Europe, it is an important measure of mortality risks, and more broadly, violence at the EU's external borders (See Carling, 2007). Yet, how migrant deaths are conceptualised is also contested: which deaths should be considered, and which should not? (see Box 2.4).

Box 2.4: Defining "missing migrants"

Julia Black

Since 2014, the International Organization for Migration's Missing Migrants Project has documented more than 75,000 deaths and disappearances during migration worldwide, but many more remain undocumented and largely invisible. The population of "missing migrants" is challenging to define, given the politicization of the topic and the lack of visibility of the largely irregular movements in which deaths and disappearances during migration occur.

IOM's Missing Migrants Project was created in response to the October 2013 shipwrecks off the coast of Lampedusa which claimed more than 300 lives. Perhaps because of its inception in the trans-Mediterranean space, it includes only deaths which occurred in the process of international migration, as well as those who go missing during maritime crossings and who are presumed dead. This definition is aimed at identifying the risks that occur during transit, but necessarily excludes many other types of missing migrants, such as deaths of labour migrants, deaths in detention or reception centres, and deaths related to internal displacement. It also excludes the hard-to-measure population of missing persons who have lost contact with their families during their migration journey.

Other datasets, including those from UNITED (UNITED for Intercultural Action, 2025), ICRC (IRC, 2022), and the Border Deaths Database (T. K. Last, 2015; see also T. Last et al 2017), use different definitions in the production of their data that include or exclude these sub-groups of missing migrants. Much of the variance in these definitions stems from the interpretation of state boundaries. A narrow definition of "missing migrants" includes only those deaths that take place at state border crossings as viewed on a map. A broader definition includes those that are linked to any "manifestation of state-made boundaries in any space," (Cuttitta Last 2019) such as suicides linked to lengthy asylum application processing times.

The production of data is key to policymaking—notably, the word "statistics" is derived from "state"—as well as forming public opinion on migration and many other topics. Different definitions of "missing migrants" make certain population groups visible, while skipping over others entirely. Different definitions of "missing migrants", and the data they entail, illuminate specific aspects of the risks of migration. These different definitions may be used constructively by data producers and users to illustrate how policy and practice contribute to preventable deaths and disappearances of migrants.

References

Cuttitta, P., & Last, T., eds. (2019). Border Deaths: Causes, Dynamics and Consequences of Migration-Related Mortality. Amsterdam University Press. doi:10.2307/j.ctvt1sgz6.

IRC (2022). "Counting the Dead: How Registered Deaths of Migrants in the Southern European Sea Border Provide Only a Glimpse of the Issue." https://missingpersons.icrc.org/library/counting-dead-how-registered-deaths-migrants-southern-european-sea-border-provide-only.

Last, T. K. (2015). "Deaths at the Borders Database." https://research.vu.nl/en/datasets/deaths-at-the-borders-database.

Last, T., Mirto, G., Ulusoy, O., Urquijo, I., Harte, J., Bami, N., Pérez, M. P., et al. (2017). "Deaths at the Borders Database: Evidence of Deceased Migrants' Bodies Found along the Southern External Borders of the European Union." Journal of Ethnic and Migration Studies 43(5): 693–712. DOI:10.1080/1369183X.2016.1276825.

UNITED for Intercultural Action (2025). "List of Refugee Deaths." https://unitedagainstrefugeedeaths.eu.

Defining irregular migration and precarity of status

What is migrant irregularity? From a legal perspective, irregular migration is defined by its opposite – what in a given state and at a given point in time is defined as legal migration, or more precisely, what the conditions are for admission and residence. Irregular migration thus is a residual category whose meaning may vary considerably over time and space.

During the period of 'guest worker' recruitment, for example, post-entry regularisation was quite common across Europe, as the majority of labour migrants were recruited through informal channels outside the formal frameworks established by labour recruitment agreements and entered on tourist visa or in fact lacked any authorisation. For instance, in 1968, 82 per cent of residence permits issued in France were issued to migrants already present in the territory, highlighting both the massive scope of informal recruitment and the scale of post-entry regularisations (Descamps, 2024:5). Today, this option is no longer available or used in most countries and application from abroad has been established as the default requirement for obtaining a residence permit.

Another example of the changing meaning of irregularity is the expansion of free movement rights within the European Union since the Treaty of Rome in 1957, its extension to family members, students and other categories and its geographical extension by successive waves of EU enlargement. While EU citizens also need to comply with certain residence requirements when moving to other EU Member States, they enjoy a wide-ranging right to movement and settlement in other EU Member States, until Brexit exposed the consequences for those citizens who had not obtained permission to stay. Otherwise non-compliance with rules is usually only sanctioned with mild penalties, for example with a fine in the case of the requirement to obtain a 'registration certificate'[5] (a type of

5 See for a summary of rules for EU citizens moving to another EU Member State https://europa.eu/youreurope/citizens/residence/documents-formalities/registering-residence/index_en.htm.

residence permit that documents, rather than authorises lawful residence). , only under certain circumstances – notably lack of means, a criminal conviction or on grounds of public security can – EU citizens be expelled or issued with a residence ban.

A third example concerns the differential conditions of entry based on visa regimes. Citizens from some states may enter visa-free, thus avoiding the risk of unlawful entry, while others require prior authorisation, making them more vulnerable to irregularisation. These distinctions are not merely technical, and reflect deeper global hierarchies of mobility rooted in postcolonial relations and geopolitical inequalities.

These examples are striking reminders of the importance of context. They also highlight that irregular migration cannot be understood as a simple binary (regular vs. irregular), as migrant irregularity is often debated in public and policy debates. The binary approach often masks the complex and diverse experiences of migrants who do not easily fit into legal categories (Triandafyllidou & Bartolini, 2020). Moreover, immigration policy itself is highly differentiated within and between countries, foreseeing different rules for different categories of people, for example between those requiring a visa and those who do not, or EU citizens and third country-nationals. Some of these distinctions are fundamental in terms of migrants' legal status. EU citizenship is one of these key distinctions within the European Union and associated states.

Another important distinction is between irregular migrants 'known to authorities' (in the sense that identities and address are known) and 'undetected' irregular migrants (European Union Agency for Fundamental Rights, 2011). Irregular migrants known to authorities are migrants who have been apprehended, whose asylum claim was rejected or whose permit has been withdrawn and currently are awaiting return. Some of these migrants may be in a situation of unlawful stay and being known to authorities only for very brief periods of time until their return is effected. In other cases, return may be suspended and they may remain in this limbo situation of receiving some legal recognition of their stay, but in principle obliged to return for years, such as in the case of persons receiving a 'Toleration' (Duldung) status in Germany (see chapter 7). In other cases, migrants may abscond , thus turning into irregular migrants not known to authorities again.

Among individuals in irregular residence situations who are not known to the authorities, further distinctions can be drawn. Crucially, it is not the person who is 'irregular', but rather their legal status, an administrative condition produced through state processes. Irregularity arises from specific legal and procedural determinations, often shaped by gaps in documentation, delayed decisions, or breaches of immigration conditions. Referring to individuals as 'irregular migrants' risks essentialising a status that is contingent, contested, and often temporary. Within this group, we can differentiate between those who lack any authorisation of stay and those who violate the conditions of an otherwise valid permit. The latter may include, for instance, tourists or students who engage in unauthorised employment or who overstay their permitted duration of stay (see Chapter 9). In both cases, the condition of irregularity is not automatic: It is formally established only once a legal process has identified an individual being in breach of immigration rules.

Another category of interest are asylum seekers. In public debates, asylum related migration has long been associated with irregular migration. Indeed, given the absence of legal pathways for admission for refugees, the large majority of asylum seekers enter European states irregularly. Yet according to Article 31 of the Geneva Refugee Convention, unlawful entry is irrelevant in the case of refugees. Also, asylum seekers' stay is lawful during the time their claim is assessed. At the same time, if their claim is rejected, they become unlawfully staying.

The status of asylum seekers therefore is of a special kind. In the MIrreM project, we have included asylum seekers in a broader category or 'class'[6] of 'provisionally staying migrants', alongside other categories of migrants, notably migrants with a suspended return decision, or migrants awaiting the outcome of the regularisation procedure. This

6 In MIrreM we use the term 'class' as we have sought to define mutually exclusive groupings of migrants within a broader taxonomy of migrant irregularity (Kraler and Ahrens 2023).

category reflects the fact that the residence rights of migrants subsumed in the category are limited and that there is a strong link to migrant irregularity, despite a temporary lawful stay. A key conclusion that we have drawn from this reflection on different types of irregularity and associated phenomena is that it is useful to place irregularity within a wider concept of legal status precariousness as an overarching category of analysis comprising irregular migrants narrowly speaking, those with a provisional right to stay, and finally, in the EU, EU citizens who have lost free movement rights (Vargas-Silva et al., 2025).

Migrants with a precarious legal status can be defined as those "individuals who lack regular immigration or residence status or, having a conditional or temporary status, are vulnerable to the loss of that status. They are therefore deprived of or run the risk of losing the most basic social rights and access to services." (Homberger et al., 2022). Table 2.1 provides an overview of the three main types of migrants with a precarious legal status we have distinguished in the MIrreM project, how we defined these, and concrete examples.

The MIrreM taxonomy of migrants with a precarious legal status

Table 2.1, below, focuses on stocks. Combining this perspective with a flow perspective, provides a scheme for analysing pathways into and out of irregularity and how these relate to different types of legal status precariousness, presented in Figure 2.2, overleaf. Importantly, this scheme only provides a snapshot at a given point in time – and within a given period of time in relation to flows. Nevertheless, it also provides a basis for conceptualising legal status trajectories over longer periods of time by considering how individuals move through different pathways and obtain or lose particular statuses, in a reiteration of the 'static' snapshot. The main purpose of the MIrreM taxonomy is a systematic mapping of available statistical indicators and estimates – and providing a conceptual framework for the collection of original data.

	Class	Definition	Examples
Irregular Migrants	Migrants without residence rights	Non-nationals (i.e. third-country in the EU) without any legal residence status in the country where they reside, including those whose presence in the territory – if detected – may be subject to termination through an order to leave and/or an expulsion order because of their activities.	• Non-nationals (i.e. third-country nationals in the EU) without any status • Non-nationals (i.e. third-country nationals in the EU) Persons engaged in an activity that violates the terms of their permission to remain in the country, which, if detected could result in the revocation of their permission to remain in the country and/or their expulsion from it. • Unregistered persons with false papers and identities • Persons issued with a return decision who do not return
Related classes to irregular migrants	Migrants with a provisional residence status or a reasonable claim to a provisional status	Non-nationals (i.e. third-country in the EU) who enjoy a provisional right to stay subject to a review of their case	• Persons whose removal has been formally suspended • Individuals awaiting status determination • Unaccompanied minors whose asylum claim has been rejected • Third-country nationals in the EU who are victims of trafficking or exploitation with a provisional permit to stay
	Mobile EU citizens with a revoked right to stay	Mobile EU citizens who have lost residence rights and no longer enjoy the right to movement and/or settlement in the EU and are liable to be removed	• Mobile EU citizens with a residence ban on public order or security grounds or criminal charges • Mobile EU citizens without a long-term residence and without sufficient means

Table 2.1: Migrants with a precarious legal status (Source: Kraler and Ahrens 2023, p.23f)

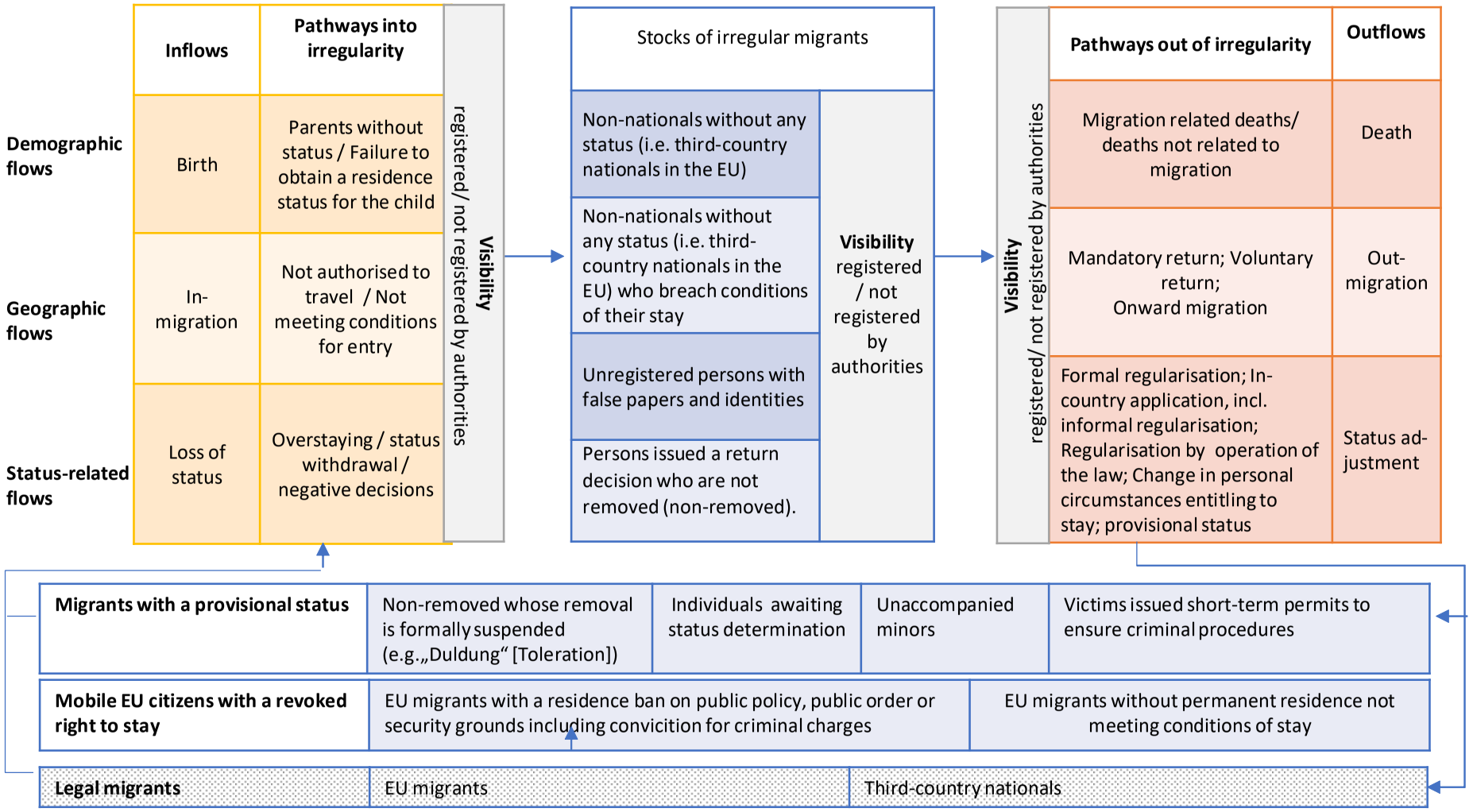


Figure 2.2: MIrreM taxonomy of migrants with a precarious legal status and pathways in and out of irregularity (Source: Kraler and Ahrens 2023, p.31.)

Conclusion

The term irregular migration occupies a central position in contemporary migration debates, yet its meaning is anything but settled. As this chapter has shown, it is a contingent, politically loaded, and administratively unstable construct. Its usage varies across institutional, national, and disciplinary contexts, often conflating legal status with racialised and gendered assumptions about social worth, deservingness, or security risk. It remains a term of operational importance for statisticians, demographers, and policymakers tasked with monitoring population movements, allocating resources, and designing policy responses.

From a scientific standpoint, treating irregular migration as a discrete, countable population is both analytically problematic and ethically fraught. People move in and out of irregularity through a range of legal, administrative, and life-course events. Their status may be ambiguous, temporary, or contested, conditions that are poorly captured by static categories. For this reason, this chapter has argued for a shift away from binary framings toward a trajectory-based understanding of legal status. This approach not only reflects the empirical realities of status transitions but also aligns with a more nuanced, longitudinal perspective on migration dynamics.

Equally important is the recognition that categories such as 'irregular', 'unauthorised', or 'undocumented' are not neutral descriptors. They are produced and reproduced within legal systems, institutional logics, and discursive fields that are themselves shaped by imperial and colonial histories of inequality, racialisation, and state power. The very effort to define and measure irregular migration thus becomes entangled with the politics of boundary-making, between citizen and non-citizen, insider and outsider, legitimate and illegitimate mobility.

At the same time, it is necessary to acknowledge the use of legal categories in migration governance. States regulate entry and residence, and these regulations inevitably generate distinctions, which in turn generate concrete outcomes. Scientific integrity requires that we do not take these distinctions at face value. Instead, we must interrogate the assumptions on which they rest, examine the consequences they produce, and remain attentive to their evolution over time.

The MIrreM framework proposed in this chapter is intended as a tool for navigating these tensions. It provides a structured yet flexible taxonomy that allows researchers, officials, and civil society actors to engage with the various phenomena of legal status precariousness in a more systematic and transparent way. It is not a final answer, but a starting point for methodological development, dialogue, data improvement, and policy reflection.

In short, irregular migration is not a property of individuals, but a product of institutional arrangements and political decisions. Measuring it (if this is possible) demands methodological rigour, definitional clarity, and above all, critical awareness. As social scientists, our task is not only to describe the world as it is but to understand the dynamics of social phenomena and make visible the ways in which categories, measurements, and narratives shape that world—and, in turn, to question whether they ought to.

References

Beauchemin, C., Descamps, J., & Dietrich-Ragon, P. (2023). “Sans papiers ou sans logement : les aléas des trajectoires des immigrés « installés » en France.” https://www.ined.fr/fr/publications/editions/document-travail/sans-papiers-ou-sans-logement-aleas-des-trajectoires-des-immigres-installes-france/#tabs-2.

Carling, J. (2007). “Migration Control and Migrant Fatalities at the Spanish-African Borders.” International Migration Review 41(2): 316–43. https://doi.org/10.1111/j.1747-7379.2007.00070.x.

Chauvin, S., & Garcés-Mascareñas, B. (2012). “Beyond Informal Citizenship: The New Moral Economy of Migrant Illegality.” International Political Sociology 6(3):241–59. https://doi.org/10.1111/j.1749-5687.2012.00162.x.

Dahinden, J. (2016). “A Plea for the ‘de-Migranticization’ of Research on Migration and Integration.” Ethnic and Racial Studies 39(13): 2207–25. https://10.1080/01419870.2015.1124129.

De Genova, N. P. (2002). “Migrant ‘Illegality’ and Deportability in Everyday Life.” Annual Review of Anthropology 31(1):419–47. https://doi.org/10.1146/annurev.anthro.31.040402.085432.

Descamps, J. (2024). “Can We See Their ID? Measuring Immigrants’ Legal Trajectory: Lessons From a French Survey.” International Migration Review 01979183241295995. https://doi.org/10.1177/01979183241295995.

European Union Agency for Fundamental Rights (2011). Fundamental Rights of Migrants in an Irregular Situation in the European Union - Comparative Report. European Union Agency for Fundamental Rights. https://fra.europa.eu/sites/default/files/fra_uploads/1827-FRA_2011_Migrants_in_an_irregular_situation_EN.pdf.

Goldring, L. (2022). “Precarious Legal Status Trajectories as Method, and the Work of Legal Status.” Citizenship Studies 26(4–5):460–70. https://doi.org/10.1080/13621025.2022.2091228.

Homberger, A., Kirchhoff, M., Mallet-Garcia, M., Ataç, I., Güntner, S., & Spencer, S. (2022). “Local Responses to Migrants with Precarious Legal Status: Negotiating Inclusive Practices in Cities Across Europe.” Zeitschrift Für Migrationsforschung/ Journal of Migration Studies 2(2):93-116 Seiten. https://doi.org/10.48439/zmf.179.

Jasso, G., Massey, D. S., Rosenzweig, M. R., & Smith, J. P. (2008). “From Illegal to Legal: Estimating Previous Illegal Experience among New Legal Immigrants to the United States.” International Migration Review 42(4):803–43. https://doi.org/10.1111/j.1747-7379.2008.00148.x.

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

Kraler, A., & Reichel, D. (2022). “Migration Statistics.” Pp. 439–62 in Introduction to Migration Studies - An Interactive Guide to the Literatures on Migration and Diversity, IMISCOE Research Series, edited by P. Scholten. Springer. https://doi.org/10.1007/978-3-030-92377-8_27

Pijnenburg, A., & Rijken, C. (2021). “Moving beyond Refugees and Migrants: Reconceptualising the Rights of People on the Move.” Interventions 23(2):273–93. doi:10.1080/1369801X.2020.1854107.

Rheindorf, M., Vollmer, B., Van Liempt, I., & Sigona, N. (2025). Understanding and Reframing Migration Narratives: Towards an Evidence-Based Policy Discourse in Europe. Policy brief. I-Claim.

Sironi, A., Bauloz, C., & Emmanuel, M. (2019). Glossary on Migration. No.34. International Migration Law. International Organization of Migration. https://publications.iom.int/system/files/pdf/iml_34_glossary.pdf.

Smellie, S., & Boswell, C. (2024). “Comparative Analysis of Migration Narratives in Political Debate and Policymaking.” doi:10.5281/ZENODO.10590197.

Triandafyllidou, A., & Bartolini, L. (2020). “Understanding Irregularity.” Pp. 11–31 in Migrants with Irregular Status in Europe, IMISCOE Research Series, edited by S. Spencer and A. Triandafyllidou. Cham: Springer International Publishing.

United Nations Secretariat. Department of Economic and Social Affairs. Statistics Division. (2025). Recommendations on Statistics of International Migration and Temporary Mobility. https://unstats.un.org/go/migrrec.

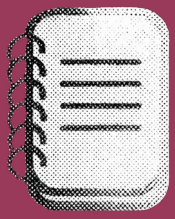
Vargas-Silva, C., Leerkes, A., Kierans, D., Siruno, L., & Kraler, A. (2025). “Tools for Collecting Information on Irregular Migration Estimates and Indicators.” Open Research Europe 5:176. doi:10.12688/openreseurope.20695.1.

Chapter 3

Ethics and data on irregular migration

Jill Ahrens

Ethics and data on irregular migration



Key points

- Embed ethics and rights-based approaches into migration research: Go beyond compliance with the European Union's General Data Protection Regulation (GDPR) and the Artificial Intelligence Act, as well as other technical guidelines to minimise harms such as surveillance, discrimination and misuse of data concerning vulnerable groups
- Ensure transparency and accountability: Acknowledge uncertainty, avoid uncritical use of categories, and prevent data misuse in shaping restrictive policies.
- Build trust and inclusive governance: Strengthen safeguards, clear communication and migrant engagement to support ethical and effective data collection and use.

Introduction

Ethical and data protection principles need to be integral to collecting, analysing and sharing data on irregular migration. This research area raises distinct ethics challenges due to migrants' vulnerability to migration enforcement, surveillance, and social exclusion. In politically charged contexts such as irregular migration, ethics requires going beyond mere compliance with laws and guidelines; proactive research ethics and integrity ensure responsible, substantive transparency, and accountable data practices while protecting participants' rights and broader social interests.

Here, transparency means not just providing access to information, but actively explaining methodological limitations, potential errors, assumptions, and the ways that data can – and cannot – be interpreted. At their core, research ethics aim to safeguard individuals through principles such as *autonomy, beneficence* (or 'do no harm' and maximise benefits), and *justice*. When working with irregular migration data, ensuring autonomy means ensuring negotiated and informed consent that takes fears about abuse and deportation seriously; beneficence requires careful assessment of the risks such as profiling or stigmatisation; and working towards justice demands recognising and mitigating power imbalances between researchers, policymakers, and migrant communities.

Research integrity builds on these ethical foundations through four principles which *The European Code of Conduct for Research Integrity*

(ALLEA, 2023, p.4) defines as:

- Reliability
- Honesty
- Respect
- Accountability.

Embedding these principles in research practice goes beyond the individual's responsibility, to include institutions and scholarly communities, and thus fostering a culture that prevents misconduct, and upholds public trust and promotes reflexive and risk-aware decision-making.

Professional standards in politically charged fields such as irregular migration include ensuring data quality, designing proportionate and privacy-conscious data collection, communicating findings with care, acknowledging others' contributions, and anticipating real-world consequences of research. In this context, transparent communication requires researchers to explicitly acknowledge uncertainty, document assumptions, and explain how findings should and should not be interpreted, rather than relying on vague claims of openness. Such an approach treats ethics not as a bureaucratic hurdle or checklist, but as an ongoing, reflexive practice shaping every stage of the data lifecycle. This is essential given the heightened risks of misuse, discrimination, and rights violations faced by irregular migrants, and the responsibility to avoid reinforcing existing inequalities through research (see Box 3.1, for the example of irregular migrant children).

Box 3.1: Making undocumented migrant children visible: A balanced approach to data collection, analysis and use

Marzia Rango, Naomi Lindt, Sebastian Palmas and Danzhen You

Collecting, producing and disseminating data and statistics on children who migrate without proper documentation or authorisation requires careful consideration. The lack of reliable data on migration that can be disaggregated by dimensions including age, sex and migratory status often renders this population statistically 'invisible', complicating efforts to uphold their most basic rights. However, if the generation of this evidence is not grounded in a child-sensitive, rights-based approach, undocumented migrant children can be potentially exposed to further rights violations, such as detention, deportation, family separation and human trafficking.

The well-being of undocumented migrant children is often undermined by their lack of legal status, particularly if they are unaccompanied or separated. The need to shed light on their deprivations and the risks they face, while also identifying and mitigating potential risks of harm that result from data collection, production and use must be thoughtfully balanced and informed by best practices (Sherr et al., 2025).

As enshrined in the UN Convention on the Rights of the Child, the best interests of children must be prioritised in all data work. To guarantee that the process of evidence generation for children is truly ethical, some core principles need to be adhered to at all times: Benefit, "Do no harm," non-discrimination, respect, justice or fairness, integrity and accountability (Rahman and Keseru, 2021). In practice, this means designing and adopting an approach centred on children's rights, which ensures that children's views are heard and their dignity respected, all while maintaining strict confidentiality and data protection protocols. Data collectors must carefully consider which data are needed to adequately represent an undocumented migrant child's circumstances and how the data will be collected, stored and used. Building trust and providing a safe environment for children to share their experiences are also crucial (Graham et al., 2013).

In an attempt to operationalise these principles, UNICEF, in collaboration with The GovLab at New York University, launched the Responsible Data for Children (RD4C) initiative.[1] This framework provides a comprehensive set of principles to guide data handling throughout its entire lifecycle –

from collection and storage to analysis and use:

- Participatory: Involving and informing children, their caregivers, and communities in the data process.
- People-centric: Prioritizing the needs and expectations of children, their caregivers, and their communities.
- Prevention of harms: Assessing and mitigating risks at every stage of the data lifecycle.
- Professionally accountable: Establishing institutional processes and roles to ensure responsible data practices are implemented.
- Purpose-driven: Ensuring data is collected with a clear objective that benefits children.
- Protective of children's rights: Upholding the rights of the child throughout the data process.

Additional resources relevant to the ethical collection and production of data and on children in vulnerable situations – such as undocumented migrant children – include the UNHCR-UNICEF Guidance Note on Responsible Disaggregation of Data on Refugee Children (UNICEF and UNHCR, 2023), UNICEF e-course on Ethics in Evidence Generation,[2] the Compendium on Ethical Research Involving Children (Graham et al., 2013) and the report Researching Sensitive Topics Involving Children (Sherr et al., 2025). A series of reports also address the ethical dimension of the use of new technologies and novel data sources for evidence generation for children (Berman and Albright, 2017; Berman et al., 2018a; Berman et al., 2018b; Rahman and Keseru, 2021).

The International Data Alliance for Children on the Move (IDAC) was launched in 2020 as a direct response to the need for better data on children on the move, particularly those who are the most vulnerable. More about IDAC, its mandate, events and resources are available at dataforchildrenonthemove.org.

References

Berman, G. & Albright, K. (2017). Rights and Ethics in a Big Data World.UNICEF Office of Research – Innocenti. https://doi.org/10.18356/7d33bc21-en

Berman, G., de la Rosa, S. & Accone, T. (2018a). Ethical Considerations When Using Geospatial Technologies for Evidence Generation. UNICEF Office of Research – Innocenti. https://doi.org/10.18356/60c0e27b-en

Berman, G., Powell, J. & Garcia Herranz, M. (2018b). Ethical Considerations in Using Social Media for Evidence Generation. UNICEF Office of Research – Innocenti. https://doi.org/10.18356/5bf16c88-en

Graham, A., Powell, M., Taylor, N., Anderson, D. & Fitzgerald, R. (2013). Ethical Research Involving Children. Florence: UNICEF Office of Research - Innocenti. https://www.unicef.org/innocenti/reports/ethical-research-involving-children

Rahman, Z. & Keseru, J. (2021) Predictive Analytics for Children: An assessment of ethical considerations, risks and benefits. UNICEF Office of Research – Innocenti. https://www.unicef.org/innocenti/media/5161/file/UNICEF-Predictive-Analytics-Working-Paper-2021.pdf

1 See https://rd4c.org/.

2 See https://agora.unicef.org/course/info.php?id=33813.

Sherr, L., Mathews, S., Hussein, M., Carter, K.L., & Bianchera, E. (2025). Researching Sensitive Topics Involving Children: Managing Risk, Distress and Trauma in Research Design and Data Collection. UNICEF Office of Research – Innocenti.
https://www.unicef.org/innocenti/reports/researching-sensitive-topics-involving-children

UNICEF and UNHCR. (2023). Responsible Disaggregation of Data on Refugee Children. New York and Copenhagen.
https://data.unicef.org/resources/responsible-disaggregation-of-data-on-refugee-children/

Ethics risks in data collection and use

Intrusive practices and group privacy

Collecting data about people in irregular situations carries significant risks of harm, especially through intrusive or disproportionate data processing. Under the General Data Protection Regulation (GDPR), personal data includes any information that can identify a person, requiring careful control over collection, storage, use, sharing, and deletion (Regulation (EU) 2016/679). Strict enforcement of GDPR compliance means that: personal data will not be shared with third parties, informed consent is mandatory, and any sharing within the project (e.g. for associated researchers or external colleagues) will follow carefully regulated agreements.

Yet, compliance with formal data protection law is only the starting point. Ethical research must also consider 'group privacy' (Floridi et al., 2018). Even anonymised and aggregated big data can enable profiling, reinforcing existing surveillance and discrimination. For migrants in an irregular situation, who are already subject to heightened scrutiny, combining or linking datasets can expose group-level patterns (e.g. concentrations in certain locations or demographic profiles) that risk further stigmatisation or enforcement action and that are also not covered in emergent regulations aimed at AI use (e.g. the EU AI act).

To mitigate these risks, it is advisable to adopt a 'dynamic approach to anonymisation' (Reed-Berendt et al., 2022). Rather than treating anonymisation as a one-off technical step, this approach recognises that identifiability can change over time or through the linking of datasets. Researchers must therefore remain vigilant, proactively assessing and reducing the risk of harmful inferences that can be made about vulnerable groups. This demands careful design of data access policies, technical safeguards, and ethical review processes, ensuring that individual and collective rights are protected at all stages.

Data sources and uncritical categories

A further ethics risk is that existing migration data sources and infrastructures have become 'invisible' or are taken for granted. Taylor and Meissner (2024) encourage researcher to uncover "a new form of metadata", namely that of data infrastructures, so as to understand who designed them, with what interests and with what assumptions about migration. Ethical practice requires resisting the role of passive data consumers, and interrogating the powers, politics and purposes built into the data systems that frame (ir)regular migration.

We should also avoid uncritically reproducing politically charged categories. Research on irregular migration often relies on legal-administrative labels that obscure lived experiences and intersectional inequalities. Such labels risk treating 'irregular' status as a dominant or 'master status' that overshadows other factors such as gender, ethnicity, racialisation, class, etc. This reification of legal categories can have real-world consequences, including legitimising restrictive policies and contributing to public fears or moral panics. As Bakewell (2008) warns for refugee studies, there is a danger of "co-producing" the problem we claim to study by adopting policy actors' assumptions uncritically.

Critical reflection on categorisation is therefore essential. Calls to 'de-migranticise' migration research (Dahinden, 2016) or adopt 'methodological denationalism' (Anderson, 2019) highlight the need to challenge taken-for-granted national or legal frameworks. Scheel and Tazzioli (2022) extend this critique by arguing that migrants are not a fixed group, but are continually shaped by the policies and practices of bordering and 'migrantisation'. Mohan et al. (2023) similarly urge researchers to reframe irregularity, foregrounding how migrant status is produced through institutional and political processes. This means treating migrant status as one variable among many, paying close attention to the processes of 'irregularisation' and the intersectional experiences of migrants. Conceptual attentiveness is not merely an academic concern but an ethical responsibility to avoid reinforcing the very inequalities and exclusions that research seeks to shed light on (see also chapter 2).

Misuse of findings

Data and findings are susceptible to misinterpretation or misuse that may inadvertently justify restrictive policies or surveillance strategies. Estimates of irregular migrant stocks and flows, or evaluations of regularisation schemes can shape public debates and policy decisions – sometimes in harmful ways. Findings can be misinterpreted, intentionally distorted, or weaponised to justify restrictive policies, surveillance technologies, or immigration enforcement measures that undermine migrants' rights.

Researchers have an ethical duty to anticipate these risks, to mitigate them and to reflect continuously on what constitutes responsible use of data (Cyrus 2023; Hendow et al. 2024). Reflexive research requires scrutiny not just of categories or analytical choices, but of the ways data might be applied in policy and public discourse. Estimates produced for analytical purposes may inadvertently aid the development of surveillance tools or influence how status determination procedures are designed in ways that limit social inclusion.

To mitigate these risks, researchers should adopt transparent and interpretively responsible communication, documenting uncertainty and methodological assumptions and providing metadata on data reliability and to always clarify what the data can and cannot show. When sharing estimates of irregular migrant stocks, it is more appropriate to publish ranges rather than point estimates alone, explicitly explaining underlying assumptions and limitations, and avoiding the impression of false precision. Ethical communication also involves being sensitive to language and framing, recognising that labels can stigmatise, and data can be appropriated to serve various purposes.

Topic bias and data gaps

Existing data on irregular migration often overlooks complex trajectories, status loss, and duration of stay in an irregular situation, creating systematic biases that must be transparently acknowledged and addressed. Bias arises not only from analytical choices but also from the fragmented and selective nature of available datasets, which are shaped by institutional priorities.

Hendow et al. (2024) argue that enthusiasm for new data sources must be tempered: data are not a panacea, and policymakers need clarity on what new data can and cannot reveal. They highlight persisting data 'blind spots' (e.g. patterns of overstaying, secondary movements, unverifiable returns, etc.), which hamper efforts to produce comprehensive EU-level estimates on the irregular migrant population. They call for data collection to be proportionate to its aims and in line with EU law, and more efforts to harmonise flow indicators and reduce double-counting. This needs to be supported by regular data exchanges and related efforts to anonymise data to address legitimate concerns over data protection and privacy.

Several studies emphasise how data collection and data use can affect the (in)visibility of irregular migrants. Jasso et al. (2008) combined administrative and survey microdata to show that administrative sources understate prior irregular experiences (e.g. entry without inspection, overstaying, and unauthorised work, etc.) and revealed important differences across origin countries, migrant categories, and within the wider population in an irregular situation (cf. chapter 2). Meanwhile, Descamps (2024), uses retrospective biographical survey data from the Trajectoires et Origines 2 survey to identify and qualify measurement biases (i.e. social desirability, recall errors, and non-proactivity) in migration status trajectory reporting, concluding that

such biases are relatively minor. She argues that migration status should more often be included in surveys, because this would enrich theoretical understandings of migrants' experiences and inform policy development. However, this assumes that migrants know their status and would report it openly and accurately.

This highlights the importance of reflexivity in interpretation, recognising whose experiences are represented and whose are overlooked. Such critical awareness should guide transparent communication to policymakers in order to prevent decisions based on incomplete or skewed evidence, which may further marginalise already vulnerable populations.

Safeguarding rights and responsible data use

Trust and mistrust

Trust is a foundational element in the collection and use of migration data. Descamps and Boswell (2018) show how institutional mistrust (e.g. fuelled by rivalries, lack of transparency, conflicting incentives, etc.) undermines coordination and data sharing. Mistrust between agencies can lead to fragmented systems, duplicated efforts, and ultimately weaker evidence for policymaking.

At the same time, migrants themselves may deeply distrust data collection efforts. Fear of surveillance, deportation, or misuse of personal information reduces willingness to participate or share accurate data (Kraler et al., 2015). This affects not only research quality but also the credibility of policy responses. However, when trust is established through robust safeguards and ethical practice, data collection and use can serve positive purposes. Responsible data use can inform the design of social inclusion programmes, improve service provision, and support policies that protect migrant's rights.

Researchers need to recognise that trust cannot be demanded but must be earned through ethical practice, including respecting autonomy, ensuring confidentiality, negotiating consent to participate, and demonstrating commitment to protecting research participants from harm. These principles must guide both data collection and the wider institutional relationships on which migration data systems depend (see Box 3.2, for an example).

Box 3.2: Addressing ethical challenges in surveying irregular migrants – The MIMAP survey on the im-/mobility of rejected asylum seekers

Randy Stache

When no sampling frame exists (e.g. when studying irregular migrants unknown to the authorities) or when particularly sensitive topics are being explored, conventional survey methods quickly reach their limits. Irregular migrants are hard-to-reach and hard-to-survey: The group is blurry and elusive (hard to identify, highly mobile with mistrust against authorities and researchers). The group also is socially and legally marginalised, vulnerable and typically lacks prior engagement with empirical research. Many are familiar with interviews only in the context of authorities, such as police or asylum proceedings. These conditions raise ethical challenges, including data protection, informed consent, and the positionality of researchers. In consequence, innovative and adaptive methodological approaches are needed.

One example for such an approach is the MIMAP project ("Feasibility study on the im-/mobility of rejected asylum seekers"). Conducted between 2022 and 2025 by the Research Centre of the Federal Office for Migration and Refugees in Germany, a part of the project focused on irregular migrants from Anglophone West Africa who had undergone an asylum procedure in Germany. It employed an innovative mixed-methods design, combining quantitative survey research with in-depth ethnographic fieldwork. Ten rejected asylum seekers were repeatedly interviewed and accompanied in their everyday lives. This ethnographic engagement facilitated trust-building and enabled the identification of key community individuals who acted as gatekeepers for the quantitative study. The survey applied Respondent-Driven Sampling (RDS), implemented via a custom-designed mobile application. The app hosted the survey, ensured full anonymity by collecting no personal data, and enabled participants to digitally refer the survey to up to three peers. Participants received a digital €10 shopping voucher both for completing the survey and for each successful referral.

To explore the sensitive issue of mobility aspirations (staying, returning, or migrating onward) the survey incorporated a factorial survey. Participants evaluated four hypothetical profiles of individuals with a 'tolerated' status, whose characteristics (e.g., length of stay: 1, 4, or 10 years) were experimentally varied. Respondents were asked to recommend whether each fictional individual should stay in Germany, return to the country of origin, or migrate to another country. The experimental variation enabled the identification of factors that shape im-/mobility aspirations. In line with the contextualizing qualitative interviews, the quantitative findings show that employment status, conditions in the country of origin, and the location of own children strongly influence (im) mobility aspirations. In contrast, migration enforcement policies such as deportation pressures and return assistance play minor roles (Stache et al., 2025).

Combining qualitative interviews and ethnographic trust-building with a respondent-driven sampling featuring an anonymous, app-based survey and a survey experiment, enabled the systematic investigation of sensitive topics among a highly inaccessible population – while maintaining ethical rigor and contextual depth.

References:

Peitz, L. and Stache, R. and Johnson, L. (2024) How to Survey Hard-to-Reach Populations: A Practical Guide to App-Based Respondent-Driven Sampling. Robert Schuman Centre for Advanced Studies Research Paper No. 2024/23, http://dx.doi.org/10.2139/ssrn.4901241

Stache, R., Johnson, L., Peitz, L., & Carwehl, A.-K. (2025). Bleibe- und Rückkehrabsichten von Geduldeten: Erkenntnisse aus einem Umfrageexperiment mit westafrikanischen Schutzsuchenden. (BAMF-Kurzanalyse, 2-2025). Nürnberg: BAMF. https://doi.org/10.48570/bamf.fz.ka.02/2025.d.2025.bleibrueckabsicht.1.0

Transparency, contestability and responsibility

Quantifying irregular migration can lend findings a veneer of objectivity and authority that masks their contingent, uncertain nature. Numbers often carry persuasive power in policy debates, but when poorly communicated or misinterpreted estimates can mislead decision-makers or the public.

Ethical responsibility demands that researchers clearly communicate the limits and assumptions of

their methods. Transparency here is substantive: it requires explaining potential sources of error, methodological assumptions, and the ways findings can and cannot be interpreted. For algorithmic methods, transparency can help other experts (and it is important to acknowledge this facet) to contest the assumptions and biases embedded into computational analysis. By doing so, policymaker and researchers help ensure that data supports informed, balanced policy decisions rather than fuelling sensationalism or punitive responses.

Considerations for data linkage and anonymisation strategies

Data protection law, especially the GDPR, imposes clear limits on how personal data may be collected, used, and shared. While these rules are crucial for protecting individual rights, they can also pose practical challenges for research, particularly in linking datasets across sources or countries.

It is necessary to respond to these challenges through careful anonymisation strategies. Pseudonymisation of individuals' identities is a standard practice, with participants given choices about the level of disclosure they are comfortable with. Researchers can use coded protocols for interviews, workshops, and surveys to minimise identifiability. Anonymisation should not be treated as a one-off exercise but as an ongoing obligation to protect participants' rights as data is processed, analysed, and shared. This also involves putting in place technical safeguards, for example: access controls that limit who can view or process data; and secure environments supported by encryption (see for an innovative example of pseudonymisation by design, Box 3.3).

Box 3.3: Linkage of administrative data in a data protection sensitive way – The case of Austria

Albert Kraler

On the national level, a wide range of statistical indicators on irregular migration are available from different administrative databases, including those on migration enforcement (apprehensions, return orders, rejections at the border, migrant smuggling, etc.), asylum databases, and residence permit databases. Despite some inherent limitations associated to their administrative purpose, the anchoring of measurement concepts in operational and legal categories and their specific scope linked to domain specific regulatory frameworks, administrative databases provide a rich source for scientific analysis. This is particularly true when they contain historical data and allow examining migrants' trajectories (chapter 7) or when they allow linkage of different databases (record linkage). In both cases, questions about data protection arise. For example, in compliance with the privacy regulations databases generally foresee a certain timeframe after which personal data needs to be deleted, if no longer necessary for the particular administrative purpose they are meant to serve. Sometimes, specific events will lead to the deletion of records from registers. For example acquisition of citizenship will result in the deletion of that person's records from residence permit registers). Similarly, record linkage can be restricted by law, as is the question of who has access to different types of data.

The case of Austria is a good example of database linkage and the preservation of historical records are possible in a data protection compliant way. In Austria, the pseudonymisation of register data for statistical purposes is achieved through the use of (encrypted) sector specific personal identifiers (verschlüsselte Bereichsspezifische Personenkennzahl Amtliche Statistik – bPK-AS). The bPK-AS is generated by the Stammzahlenregisterbehörde (Central Register Authority). It is a cryptographically derived identifier derived from the personal identifier used in a specific domain (for example social security, or the population register code) and a code for the domain.[3] It is unique to each individual

3 The principle of encryption used for the generation of the sector specific identifiers is described (in German) here : https://www.bundeskanzleramt.gv.at/agenda/digitalisierung/stammzahlenregisterbehoerde/bereichsspezifische-personenkennzeichen/beschreibung.html.
The encryption procedure is based on Central Register Authority Ordinance (Stammzahlenregisterbehördenverordnung) 2022, see https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20011934.

and serves as a key to match data from different registers with each other. Crucially, the bPK-AS is not reversible, meaning it cannot be traced back to the original personal identification number (Statistics Austria 2024). Statistics Austria uses these anonymised personal identifiers to link data from various sources – such as social insurance records, employment data, and education registers – through deterministic linkage and without revealing personal identities. While Statistics Austria gets updates from administrative databases in real time, it uses anonymized statistical mirror databases for statistical purposes (Fuchs et al. 2024). All register data is stored in a historicised way, allowing longitudinal analysis.

Since 2022, all statistical databases based on data collected by Statistics Austria itself (through surveys and other statistical reporting systems) as well as a wide range of administrative databases from different public bodies are assembled in the "Austria Micro Data Centre" (AMDC).[4] By mid-2026, all public administrative database – with the exception of security related databases – should be made available by the AMDC. In addition, researchers can link their own datasets to the AMDC by obtaining a sector specific identifier from the Central Register Authority for their own dataset, which in turn enables Statistics Austria to include this dataset in the AMDC, making it linkable to all datasets contained in the AMDC. A precondition for including a dataset in the AMDC is that researchers collect personal information (notably name, date of birth, place of residence) to enable pseudonymisation by the Central Register Authority. The AMDC is open for researchers in accredited institutions, which need to meet a number of criteria for accreditation (such as scientific purpose of the organisation, research quality, independence).

While immigration and migration enforcement related databases are not (yet) linked to the AMDC and therefore cannot be used to analyse legal status trajectories, the design of the system nevertheless can serve as a model for balancing data utility and privacy protection.

References

Fuchs, R., Göllner, T., Hartmann, S. & Thomas, T. (2024). "Fostering Excellent Research by the Austrian Micro Data Center (AMDC)" Jahrbücher für Nationalökonomie und Statistik, 244(4), 433-445. https://doi.org/10.1515/jbnst-2023-0043

Statistik Austria (2024): Standard-Dokumentation Metainformationen (Definitionen, Erläuterungen, Methoden, Qualität) zu den Registerbasierten Erwerbsverläufen. https://www.statistik.at/fileadmin/shared/QM/Standarddokumentationen/B_1/std_b_erv.pdf

Importantly, 'special categories' of personal data, such as 'race', ethnic origin or political opinions, carry heightened risks. The MIrreM project applies the 'data minimisation principle' by deliberately limiting data collection to what is strictly necessary, while ensuring individuals are fully informed of their rights and protections.

When applied carefully, these practices allow data to be used constructively: for example, to understand migration patterns, design inclusive services and improve resource allocation without compromising individual privacy.

4 https://www.statistik.at/en/services/tools/services/center-for-science/austrian-micro-data-center-amdc

Secondary use of data

MIrreM also uses existing datasets to estimate irregular migrant populations. Even if these are anonymised or aggregated, ethical issues remain. Researchers and policymakers must consider the conditions under which data were originally collected, and if this included informed consent, voluntariness and transparency, and how linking datasets may create new risks or reinforce surveillance logics.

To address this, researchers need to commit to clear documentation of data sources, ethical review of any secondary use, and a critical assessment of how data linking may affect the rights and perceptions of the populations concerned. Policymakers must be wary of normalising data practices that reinforce securitisation narratives, where migrants are framed primarily as risks to be managed rather than individuals with rights.

At the same time, responsible linking and analysing of secondary data can yield valuable insights for planning services, understanding the characteristics of migrant populations and evaluating the effectiveness of policies. This requires a careful balance between administrative utility and respect for fundamental human rights.

Inclusive governance and legal safeguards

The use of data about irregular migration should complement, not replace, engagement with migrants themselves or with civil society organisations that work with them directly. Policymakers should strive for inclusive governance in migration data systems, ensuring that policy proposals reflect diverse perspectives and do not solely rely on technocratic or quantitative assessments.

Policymakers must also consider the need for updated legal frameworks to regulate the use of linked or repurposed datasets, especially when applied to groups that may lack formal protections. This includes reviewing data protection laws and institutional safeguards to ensure that they cover the specific vulnerabilities associated with an irregular residence status.

When governance frameworks are inclusive and transparent, data can be used proactively to identify gaps in protection, target resources effectively and support interventions that benefit migrants and wider communities.

Conclusion

Research ethics in the context of irregular migration cannot be reduced to a checklist. Compliance with legal frameworks such as GDPR and the EU AI Act is necessary, but only as a baseline. What is required instead is an ongoing reflexive approach about the risks, responsibilities, and power relations involved at every stage – from research design and data collection to analysis and communication. For those involved in data collection and processing, such as researchers, statisticians, public sector officials and those working in migrant support organisations, this means embedding ethics awareness in all activities, recognising the rights and dignity of those whose lives are studied, and promoting transparency and accountability in the production and use of migration data. By treating ethics as an integral, continuous process, researchers can help ensure that their work contributes to more just, humane, and evidence-informed migration policy.

References

Anderson, B. (2019). New directions in migration studies: Towards methodological de nationalism. Comparative Migration Studies, 7(1), 1-13. https://doi.org/10.1186/s40878-019-0140-8

ALLEA. (2023). The European Code of Conduct for Research Integrity (Revised ed.). Berlin: All European Academies. https://doi.org/10.26356/ECOC

Bakewell, O. (2008). Research beyond the categories: The importance of policy irrelevant research into forced migration. Journal of Refugee Studies, 21(4), 432–453. https://doi.org/10.1093/jrs/fen042

Cyrus, N. (2023). Ethical Benchmarking for the Measurement of Irregular Migration. In MIrreM Policy Brief No.1. Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.10042022

Dahinden, J. (2016). A plea for the 'de-migranticization' of research on migration and integration. Ethnic and racial studies, 39(13), 2207-2225. https://doi.org/10.1080/01419870.2015.1124129

Descamps, J. (2024). Can We See Their ID? Measuring Immigrants' Legal Trajectory: Lessons From a French Survey. International Migration Review, https://doi.org/10.1177/01979183241295995

European Union (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts. Official Journal of the European Union, L, 2024(1689), 1–127. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). Official Journal of the European Union, L119, 1–88. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Floridi, L., Taddeo, M., & Turilli, M. (2018). Group privacy: A defence and an interpretation. In L. Taylor, L. Floridi, & B. van der Sloot (Eds.), Group privacy: New challenges of data technologies (pp. 83–100). Cham, Switzerland: Springer.

Hendow, M., Wagner, M., Ahrens, J., Cherti, M., Kierans, D., Kraler, A., Leerkes, A., Leon, L., Rodríguez Sánchez, A., Siruno, L., Tjaden, J., & Vargas-Silva, C. (2024). How fit is the available data on irregular migration for policymaking?. In MIrreM Policy Brief No. 3. Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.13757685

Kraler, A., Reichel, D., & Entzinger, H. (2015). Migration statistics in Europe: A core component of governance and population research. Integrating immigrants in Europe: Research-policy dialogues, 39-58. https://doi.org/10.1007/978-3-319-16256-0_3

Jasso, G., Massey, D. S., Rosenzweig, M. R., & Smith, J. P. (2008). From illegal to legal: Estimating previous illegal experience among new legal immigrants to the United States. International Migration Review, 42(4), 803–843. https://doi.org/10.1111/j.1747-7379.2008.00148.x

Mohan, S. S., Mountz, A., Romero, M., & Visan, A. (2023). How to research 'irregular' migration: approaches and perspectives from the field. In Research Handbook on Irregular Migration (pp. 36-48). Edward Elgar Publishing. https://doi.org/10.4337/9781800377509

Reed-Berendt, R., Dove, E. S., Pareek, M., & Group, U.-R. S. C. (2022). The Ethical Implications of Big Data Research in Public Health: "Big Data Ethics by Design" in the UK-REACH Study. Ethics & Human Research, 44(1), 2–17. https://doi.org/10.1002/eahr.500111

Scheel, S., & Tazzioli, M. (2022). Who is a migrant? Abandoning the nation-state point of view in the study of migration. Migration Politics, 1(1), 002. https://doi.org/10.21468/MigPol.1.1.002

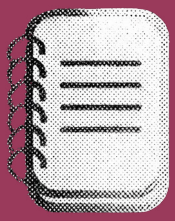
Taylor, L., & Meissner, F. (2024). Migration statistics in times of large-scale mobility data: ethical concerns and concerns with ethics. In W. L. Allen, & C. Vargas-Silva (Eds.), Handbook of research methods in migration (2nd edition, pp. 280-295). Edward Elgar Publishing. https://doi.org/10.4337/9781800378032.00031

Chapter 4

What are good quality data on a phenomenon that is hard to measure?

Denis Kierans and Lalaine Siruno

What are good quality data on a phenomenon that is hard to measure?



Key points

- Irregular migration is difficult to measure, and the data that exist are often limited, inconsistent, or outdated. This chapter introduces a practical framework to help users assess the quality and credibility of such data, rather than taking estimates at face value.
- It distinguishes between key data types—stocks vs flows, estimates vs indicators—and highlights how conceptual ambiguity, observational gaps, and poor documentation can undermine how irregular migration data are interpreted and used.
- When applied to over 250 estimates across 14 countries, the framework reveals significant variation in quality. While some countries produce relatively robust and transparent figures, many rely on outdated, methodically weak or poorly documented estimates. Still, pockets of good practice exist across North America and Europe, which can be built on.
- The chapter argues that responsible use of irregular migration data depends not only on improving data systems, but also on the ability of users to critically assess what data mean, how they were produced, and whether they are fit for purpose.

Introduction: Measuring the unmeasurable?

Policymakers often point to data to justify their decisions, particularly in contested policy spaces, such as immigration (Boswell, 2009; Kraler & Reichel, 2022). Irregular migration, while a point in case, poses distinct challenges to this practice. Although irregular migrants are the subject of intense political and media scrutiny in many countries, information about them is notoriously scant and unreliable (Vollmer, 2011).

This chapter examines what constitutes "good quality" data in this complex landscape. It introduces the MIrreM framework for assessing irregular migration estimates and indicators and discusses how uncertainty (Box 4.1) and other criteria shape data quality. Rather than assuming these figures are fit for purpose, we suggest users interrogate them, asking whether the data are credible, transparent, and suited to question at hand. We conclude with a checklist based on the five MIrreM criteria to support this process.

Box 4.1: Uncertainty in irregular migration data

Denis Kierans

There are three main sources of uncertainty in irregular migration data:

- **Conceptual uncertainty** arises from disagreements over who should count as an irregular migrant. Legal definitions vary across jurisdictions, media usage is inconsistent, and there is no universal agreement on terminology in academic circles, either. The MIrreM project distinguishes three categories: (1) migrants in an irregular situation, (2) migrants with a provisional status or a claim to one, and (3) EU citizens without residence rights. This taxonomy helps make explicit the assumptions behind different datasets and more detail on this can be found in Chapter 2.
- **Ethical uncertainty** emerges from the tension between producing data, such as estimates and indicators, on irregular migration and protecting individual rights. Certain types of data collection run the risk of compromising privacy and creating risks for vulnerable people. Furthermore, due to the political sensitivity of this area, these numbers – irrespective of their quality – can contribute to negative public sentiment and actions towards migrants in general and be used to justify reactionary and punitive policymaking. More information about ethics and data protection can be found in Chapter 3.
- **Practical uncertainty** is related to observability. Many irregular migrants avoid contact with authorities and may not appear in traditional sources of data on migration, such as administrative systems or surveys. Indicators based on events (e.g., border apprehensions, deaths occurring during the migration process in the Mediterranean, the English Channel or along the US – Mexico border) only capture the visible portion of the phenomenon. Estimating what is not directly observed requires the use of methods underpinned by assumptions, some more plausible than others.

This chapter focuses primarily on the conceptual and practical aspects of uncertainty. However, it is important to consider all elements when assessing and using irregular migration data.

What types of irregular migration data are there?

Data on irregular migration can be categorised broadly in two important ways. First, as either stock or flow data:

- **Stocks:** Quantitative estimates of the number of irregular migrants present in a given country or area at a particular point in time. Stock data help to answer questions such as "how many irregular migrants are living in Country X today?"
- **Flows:** Data that describe movements into, within, or out of irregular status over a specified period of time. Flow data help to answer questions like "how many irregular migrants were apprehended at the border in the past month?"

Second, irregular migration data can be distinguished as an estimate or an indicator:

- **Estimates:** quantifying a population group that cannot be directly observed or counted. Estimates help to answer questions such as "how many irregular migrants we living in the United States in 2020?"
- **Indicators:** Counts of observed events, such as border apprehensions or asylum decisions. Indicators help to answer questions like "how many migrants were intercepted by border control last quarter?" Indicators can track trends and inform planning but do not capture full populations. They often feed into estimates.

Box 4.2: Irregular migration to the UK: A Home Office statistical overview

Jon Simmons and Lucy Swinnerton

The Home Office publishes a wide range of statistics that reflect different aspects of irregular migration to the UK, drawn from the department's operational systems.

Daily counts of small boat arrivals in the English Channel provide a near real-time operational snapshot of people detected arriving in the UK this way. These figures, published on GOV.UK,[1] with a public dataset of daily arrivals, are manually collated by Border Force officers and offer a timely but provisional view of one of the most visible forms of irregular entry.

The Home Office's comprehensive publication of statistics on the operation of the UK immigration system are the quarterly 'Immigration System Statistics',[2] which bring together administrative data from operational and case working systems to provide a wide range of information on irregular migration. This includes demographic details of those arriving to the UK by small boat (age, sex and nationality), as well as information on asylum claims made by small boat arrivals, decisions on those claims, and referrals to the National Referral Mechanism (NRM) for potential victims of modern slavery.

Publicly available data on irregular arrivals are not limited to those entering the UK on small boats. Data on inadequately documented air arrivals (individuals arriving without valid or with fraudulent documents) and detections in-country and at ports are also recorded on Home Office systems and included in published statistics. The department also publishes statistics on immigration enforcement activity, including return of small boat arrivals with no right to remain in the UK. This information helps to contextualise the experiences and outcomes of individuals arriving to the UK via irregular routes and compliments other forms of research.

Data on irregular arrivals are derived from live operational systems and so numbers may change as more information is added. The statistics sometimes present a partial picture of most recent events due to time required for data validation or additional safeguarding assessments, or the time necessary to come to a decision on often complex individual cases. Although the statistics are drawn from live systems and may be subject to revision, they do nonetheless provide a consistent and broad basis for understanding patterns over time.

The Home Office statistics are accredited Official Statistics published according to the Code of Practice for Statistics overseen by the UK Statistics Authority. The published 'User Guide to Immigration System Statistics'[3] provides documentation to support understanding of the data, including definitions, data sources, and known limitations.

While the statistics published by the Home Office provide a wide range of valuable insights, they cannot capture the full extent of irregular migration. Some individuals evade detection entirely while others may see their status change. Others of course regularise their status, for example through a successful application for refugee status. The Home Office has published some initial information in relation to those who arrive on a visa but subsequently claim asylum (including individuals who overstay their visa and claim after their visa status is no longer valid),[4] but this is novel and complex data and so unlike the regular quarterly outputs at present it is regarded as 'statistics in development'.

It is also possible that the same person may be detected at multiple times in different operational systems, which cannot always or simply be linked. These statistics therefore offer a partial view of irregular migration and users should interpret data with care, carefully considering the limitations and assumptions that underpin them.

1 See https://www.gov.uk/government/publications/migrants-detected-crossing-the-english-channel-in-small-boats.

2 See https://www.gov.uk/government/collections/immigration-statistics-quarterly-release.

3 See https://www.gov.uk/government/publications/user-guide-to-home-office-immigration-statistics--9/user-guide-to-immigration-statistics.

4 See https://www.gov.uk/government/publications/source-of-asylum-claims-in-2024/source-of-asylum-claims-in-2024.

How do you assess the quality of irregular migration data?

To assess the quality of these data, MIrreM used a structured set of criteria which are informed by the FAIR data principles: Findability, Accessibility, Interoperability, and Reusability (Wilkinson et al., 2016). The following framework outlines the five criteria against which we assessed the quality of the irregular migration estimates.

MIrreM criteria	High (3 points)	Medium (2 points)	Low (1 point)
Accessibility	All raw data used to construct the estimate is publicly available and electronically accessible with no permissions required.	At least some of the raw data used to construct the estimate is only available on request from relevant authorities. If some of the data is not available at all, then give 1 point.	At least some of the raw data used to construct the estimate is not available for most potential users.
Documentation	Full documentation about data and methods are available and accessible. The level of information allows for replication of the estimates.	Limited information on data, estimation methods, and quality are available and accessible. Insufficient details to replicate the estimates.	Information on data and estimation methods is neither available nor accessible.
Reliability	Analysis includes demonstrated reliability indicators, with limitations clearly specified (e.g. ranges, alternative calculations, characterisation as minimum or maximum estimate).	Some discussion of reliability, but no indicators in quantitative terms.	Missing a discussion of reliability.
Methodology	Methodology is adequate and comprehensive including, but not limited to, rigorously implemented multiplier or residual studies.	Methodology is adequate, even if not comprehensive, including but not limited to: (1) Simple multiplier calculations; (2) Simple residual estimates; (3) Adjustment of older estimates with partly insufficient data; (4) Aggregate estimates for different groups, partly relying on plausibility calculations.	Inadequate method and application of the method; resulting estimate lacks foundation
Data	The analysis relies on an adequate dataset not likely to have a considerable bias, including no bias for any group estimates. There are no strong assumptions regarding the data.	The analysis relies on a biased dataset. There are plausible adjustments and assumptions. This includes cases in which the dataset does not provide the information necessary or it is necessary to make strong assumptions.	The analysis relies on a biased dataset, without proper adjustments. The assumptions regarding data are not plausible.

Table 4.1: Criteria for the MIrreM quality evaluation of estimates (Source: Kierans and Vargas-Silva, 2024)

Alongside scoring each estimate on five individual criteria, we also developed an aggregate quality score. It combines the different dimensions into a single number to help users quickly understand the overall robustness of an estimate.

Although we hope it is helpful, we recognise it is not objective. It reflects deliberate decisions about what matters most. Some criteria – such as the soundness of methods or how uncertainty is handled – are more central to quality than, for example, whether data are fully open-access. We therefore weighted the criteria: methodology and reliability counted more than data access or documentation. We also introduced thresholds to stop poor methods or unreliable outputs from boosting the total score, reflecting feedback that no amount of documentation or access can compensate for fundamental flaws.

A slightly different quality framework was used for the assessment of indicators. In particular, the methodology and data criteria were excluded. This is because, unlike estimates, indicator data are observed or registered event collected and reported by national statistical institutes or enforcement authorities like FRONTEX. Instead, we grouped validity and reliability to assess whether an indicator is in fact measuring the type of irregular flow it is supposed to measure.

MIrreM Criteria	High (3 points)	Medium (2 points)	Low (1 point)
Accessibility	Data is publicly available and electronically accessible with no permissions required	Data is available on request from relevant authorities	Data is available, but access and use are exclusive to authorities
Documentation	Sufficient and transparent information on data and methods are available and accessible; a comprehensive quality report is also available	Limited information on data, methods, and quality are available and accessible	Information on data, methods, and quality are neither available nor accessible
Validity and reliability	Data is representative of the phenomenon it is supposed to measure and adequately reflects the type of irregular migration being measured; data is relatively complete (not highly selective) and does not indicate internal contradictions	Data is selective and points to some internal contradictions	Data is neither valid nor reliable

Table 4.2: Criteria for the MIrreM quality evaluation of indicators (Source: Siruno et al., 2024)

We understand others may have different views on how to assess these data, which we welcome. Our aim is to encourage critical reflection and debate around irregular migration data. More information about the MIrreM data collection process and tools are available in Vargas-Silva et al. (2025).

What do high and low quality data look like?

We found wide variation in the quantity and quality of national estimates (n=260). For instance, the United States produces multiple annual estimates of high reliability. The UK and Germany, which have the largest estimated irregular migration populations in Europe, rely on relatively outdated figures. Austria excels in access to the underlying data but does worse against other criteria. Canada and Portugal have few estimates, all scored poorly. No estimates were located for Bosnia and Herzegovina, Morocco, Serbia or Türkiye.

Country	Aggregate	Access	Documentation	Reliability	Method	Data
Germany	**11.1**	2.0	2.9	3.0	3.0	2.5
France	**10.7**	2.8	2.8	2.8	2.8	1.9
United States	**10.3**	1.8	3.0	2.6	2.7	3.0
United Kingdom	**10.1**	1.9	3.0	2.7	2.7	2.0
Italy	**10.1**	1.7	2.4	2.4	3.0	3.0
Spain	**10.0**	2.6	2.0	2.9	2.6	2.5
Belgium	**9.5**	2.2	2.7	2.0	2.7	2.7
All countries	**9.1**	**2.0**	**2.4**	**2.4**	**2.5**	**2.2**
Greece	**9.1**	2.1	2.1	2.7	2.0	2.9
Austria	**8.3**	3.0	2.0	2.3	2.1	1.3
Netherlands	**7.9**	1.2	2.2	2.2	2.2	1.3
Ireland	**7.8**	1.6	2.2	2.4	1.9	1.6
Poland	**7.6**	2.4	1.9	2.1	2.4	1.0
Portugal	**7.2**	1.8	1.8	1.8	1.8	1.9
Finland	**5.9**	1.9	1.6	1.9	1.6	1.4
Canada	**4.0**	1.0	1.0	1.0	1.0	1.0

Table 4.3: Average MIrreM quality point scores for irregular migration stock estimates, criteria and aggregate, by country. Note: Aggregate scores range from 4 (low) points to 12 (high) points; the other criteria range from 1 to 3 points. (Source: Kierans et al., 2024)

The MIrreM project analysed six of the most frequently discussed indicators related to irregular migration flows, all of which are published by Eurostat.

Type of flow	Indicator – inflow	Indicator – outflow
Geographic	(1) TCNs refused entry at the external borders	(3) TCNs ordered to leave
	(2) TCNs found to be illegally present inland because of illegal entry	(4) TCNs returned to a third country following an order to leave
Status-related	(5) Negative first instance asylum decisions	
	(6) Decisions withdrawing status granted at first instance decision	

Table 4.4: Selected irregular flow indicators

We found these indicators to be of relatively good quality. However, it remains difficult to establish whether they are fully accurate and reliable. For example, double counting and missing information can affect measurement precision, especially when disaggregated by age and sex. Furthermore, because the data are typically produced by bureaucracies with limited oversight, there are few opportunities to validate the numbers by cross-checking them with other information.

Criteria	Quality	Explanation
Accessibility	High	Data are publicly available and electronically accessible from Eurostat with no permissions required. General alignment with the FAIR Data Principles.
Documentation	High	Sufficient and transparent information on data and methods available and accessible; comprehensive quality report for almost all countries covered.
Validity and reliability	Medium	The data is generally representative of the specific flows being measured, but there is some variation re: external validity and reliability (percentage of missing values and inter-item correlation tests).

Table 4.5: Summary of quality assessment of selected Eurostat indicators (Source: Siruno et al., 2024)

We hope that these examples highlight the value of systems of critical appraisal by helping users understand and distinguish between different sources of irregular migration data.

Conclusion

Measuring and estimating irregular migration will always be difficult. But better data and their use is possible. A step in the right direction is to incorporate critical appraisals of the data as a matter of course, especially for those shaping policy and the public debate. To this end, the MIrreM quality criteria can be used or adapted as a preliminary, rapid-fire assessment tool for irregular migration data.

Not all estimates or indicators will meet all criteria fully. However, if you are unable to answer the relevant questions on the checklist, we recommend learning more about the data before you use it in your work or draw conclusions from the data.

- **Accessibility**: Are the raw data publicly available?
- **Documentation**: Are the methods and sources clearly explained?
- **Reliability**: Are uncertainty and limitations explicitly addressed?
- **Methodology**: Is the *estimate's* approach comprehensive, appropriate and replicable?
- **Data quality**: Are the data underpinning the *estimate* relatively unbiased and fit for purpose?
- **Validity and reliability**: Are you clear on what the *indicator* measures and does it do so accurately and consistently?

Figure 4.1: Checklist for rapid assessment irregular migration estimates and indicators

To conclude, we emphasise that assessing the quality of irregular migration data is not merely an academic or technical matter. Given irregular migration data's uneven quality and limited availability – combined with the political and public sensitivity of the issue – it is easy to misinterpret, with potentially serious consequences. Avoiding this requires investment not only in data systems, but also in the capacities of individuals and institutions to interpret and use irregular migration data responsibly.

References

Boswell, C. (2009). The Political Uses of Expert Knowledge: Immigration Policy and Social Research. Cambridge University Press.

Kierans, D. & Vargas-Silva, C. (2024). The Irregular Migrant Population of Europe. MIrreM Working Paper No. 11. Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.13857073

Kraler, A. & Reichel, D. (2022). Migration Statistics. In: Scholten, P. (eds) Introduction to Migration Studies. IMISCOE Research Series. Springer, Cham. https://doi.org/10.1007/978-3-030-92377-8_27

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

Vargas-Silva, C., Leerkes A., Kierans, D., Siruno, L. & Kraler, A. (2025). Tools for collecting information on irregular migration estimates and indicators [version 1; peer review: 2 approved]. Open Res Europe 2025, 5:176. https://doi.org/10.12688/openreseurope.20695.1

Vollmer, B. A. (2011). Policy Discourses on Irregular Migration in the EU-'Number Games' and 'Political Games'. European Journal of Migration and Law, 13(3), 317-339. https://doi.org/10.1163/157181611X587874

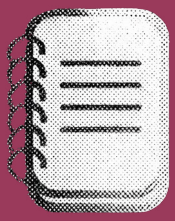
Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3. https://doi.org/10.1038/sdata.2016.18

Chapter 5

Innovations in methodological approaches to estimate irregular migrant stocks and flows

Alejandra Rodríguez-Sánchez and Jasper Tjaden

Innovations in methodological approaches to estimate irregular migrant stocks and flows



Key points

- This chapter builds on Rodríguez-Sánchez and Tjaden (2023), who reviewed the main methods for estimating irregular migrant stocks and flows, spanning both established and more experimental approaches.
- Measuring irregular migration remains a fundamental challenge: the population is difficult to observe, and even widely used methods such as residual estimation or capture–recapture provide only partial pictures.
- Traditional techniques continue to form the backbone of the field, but improvements have often come from incremental innovations, such as using mortality data to refine life-course approaches, or expanding residual methods with large government databases and machine learning.
- More novel directions, such as exploiting consular registers, driver's licence data, or online search behaviour, show promise in filling gaps, though these remain context-specific and experimental.

Introduction

The scientific study of irregular migration, its description, and its estimation are closely connected with the search for appropriate measurements and empirical observations related to this form of migration. Aside matters related to the definition of irregular migration, treated in other chapters in the book, and especially in Kraler and Ahrens (2023), there are multiple challenges associated with measuring irregular migration stocks and flows and attaining estimates of the size of these quantities.

Some of these challenges were highlighted by the seminal CLANDESTINO project – an EU–funded project (2007-2009) which reviewed data and methodologies on irregular migration over a decade ago (Jandl, 2011). Key obstacles include irregular migrants' reluctance to disclose their status in surveys or censuses, the absence of adequate sampling frameworks, and their elevated mobility patterns—all necessitating alternative research approaches. In MIrreM, we seek to update this review with advances in literature in terms of data

sources and methodologies. For a comprehensive and detailed overview of each method, see the review paper by Rodríguez-Sánchez and Tjaden (2023). In that review, a detailed explanation of how the different methods work, typical databases used, and their strengths/weaknesses are documented. Depending on which definition of irregular migration we employ, some methods might be better suited than others to capture the different operationalizations, especially as these will be based on different data sources

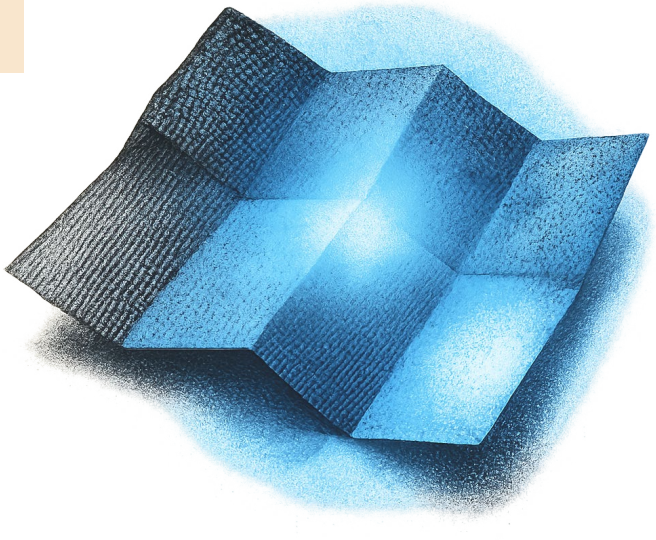
Box 5.1: Traditional and innovative approaches

Alejandra Rodríguez-Sánchez and Jasper Tjaden

By "Traditional" we refer to approaches covered by previous methodological overviews (Jandl, 2011; Pinkerton et al., 2004) on which our overview builds. These are well established methods that are used to estimate irregular migration across the world. We included these traditional methods in our review out of a desire to be comprehensive, but also because some of the innovations build from well-established methods, like the residual approach. We defined "Innovative" approaches as those methods that either use novel data sources (e.g., digital behavioural data) or apply a new estimation method to standard data sources. These approaches improve upon some of the limitations of established methods. The innovative approaches were identified through literature review and discussions with experts.

Estimating irregular migrant stocks and, especially, irregular migration flows, despite important advances, remains a challenging endeavour. Methods estimating stocks measure the total number of irregular migrants residing in a country at a specific point in time (e.g., the year 2022), whereas methods estimating flows capture changes in that population over a defined period (e.g., 2015–2020), attaining measures of inflows or outflows. MIrreM's innovative pilot studies which are summarized in the next chapter, and which were based on this overview, are aimed at tackling some of these challenges.

This chapter of the Handbook is intended to highlight some of the most innovative aspects of approaches we found through our scoping review for measuring both stocks and flows. A quick overview of the methodologies we found in this search can be found in Table 1. We grouped the methods based on their core data (e.g., government data, non-government data, survey, mixed data, and digital data) and estimation strategies, plus a brief description of their main idea, rather than focusing on minor differences across methods. Our review encompasses both traditional and innovative methods (see Box 1).



Type	Government administrative & census	Description
Traditional	Capture-Recapture / Multiple Systems Estimation	The many varieties of this approach are based on a statistical method that estimates population size by comparing how many individuals appear in multiple data sources. For irregular migration, it matches people between enforcement records (like police apprehensions) and general population databases. The detection rate from this overlap is used to estimate the total irregular migrant population. Variations include Multiple Systems Estimation and repeated capture-recapture, all using the same core principle of linking individuals across datasets.
	Regularisation & formal status adjustments	Amnesties or regularizations grant legal status to irregular migrants, revealing population size as a byproduct. Case-by-case adjustments, called "silent regularization," happen continuously without formal announcements, also providing population estimates.
	Life course events	This method compares observed demographic rates (e.g., births, deaths, hospitalizations) among migrants to expected rates in non-migrant populations. When migrant rates exceed projections, the excess suggests the presence of unregistered irregular migrants who are not captured in official statistics.
	Labour demand models	This approach estimates labour demand from economic output, then subtracts reported foreign workers to find irregular migrant workers. A similar method estimates the shadow economy size and assumes migrants' share of irregular work.
	Flow-Stock models	Here, the main idea is that irregular migrant population can be obtained by understanding the initial size of this population at some point, and adjust by inflows (adding subsequent overstayers, border crossings) and outflows (subtracting deportations, emigration, deaths, status changes).
	Irregular border crossings & apprehensions	Irregular migration flows are measured by counting people apprehended at borders or after entry, through checks within a country.
	Database systems enabling the identification of visa overstayers	This method calculates irregular migrants by subtracting legal departures from legal arrivals plus status changes in tracking systems. For instance, the US ADIS system matches entry records with departure and status change data for air, land, and sea travelers. Visa overstayers are identified as those who entered legally but neither left nor gained new legal status. The approach may use biometric and biographical verification to detect overstayers.
Innovative	Using consular registers of migrant communities in combination with other data-sources	This method combines consulate card registrations with survey data. For example, Mexican consulates in the US register nationals and issue cards by region. These counts estimate total resident Mexican nationals, then survey data on undocumented percentages from the American Community Survey is applied to calculate irregular Mexican migrants.
	Drivers' licenses	In some states in the US, comparing the number of drivers licenses before and after law changes that grant irregular migrants access to driver's licenses (e.g., the AB60 reform in California, US) could be indicative of the size of undocumented present, similarly to regularization given the widespread use of cars.
	Consumption data	In this method, estimates of expected food consumption (like rice) based on development levels and daily consumption is compared to actual consumption levels. The excess consumption, unexplained by the official population, is attributed to irregular migrants and converted to population estimates using consumption pattern assumptions.
	Biometrics & surveillance	This method uses AI facial recognition and biometric data to identify irregular migrants by comparing faces against government databases. For example, migrants who entered the country irregularly or overstayed the terms of their visa will not appear in matching official biometric records. Large-scale application across complete databases enables population estimation.

Type	Non-government administrative data	Description
Innovative	Institutional registers: College enrolment, prison population, and human trafficking	Institutional registers (such as universities, prisons, trafficking victim services) capture legal status data. Universities, for example, can identify irregular students through admissions when documentation is missing. Prison systems, in turn, record migrants' legal status after they have committed crimes. NGOs can provide information on trafficking victims. Hence, these sources can provide direct counts of irregular migrants within specific populations potentially useful in combination with other approaches.
	Missing, disappeared or deceased migrants	This method tracks individuals who are missing, disappeared, or died while attempting irregular border crossings or migration journeys. Although this is not a count of irregular migrant flows, it might provide further evidence to understand dynamics of irregular migration flow.
	Survey data	
Traditional	Self-identification in surveys	Surveys can help determine respondents' legal status either by asking directly about residence permits or indirectly by inquiring about rights or other characteristics linked to having/or not having a legal status.
	Expert surveys / Delphi surveys	Researchers survey experts who estimate the population size in specific areas. The Delphi method, for example, involves multiple survey rounds where experts see colleagues' responses and discuss until reaching consensus through iterative feedback.
Innovative	Statistical imputation	This approach uses survey data with direct or indirect legal status questions to train predictive models, then applies these models to larger national surveys lacking such legal status-related questions to identify irregular migrants. Various techniques are used including logical edits, logistic regression, or machine learning.
	Retrospective surveys	In this method surveys on regular migrants can be used to retrospectively infer their previous irregular status periods, allowing for an estimation of previous legal statuses, especially if samples are sufficiently large.
	Mixed data sources	
Traditional	Multiplier/ simple extrapolation	Find the ratio of irregular migrants compared to all foreign-born individuals, then project this proportion onto different geographic scales or contexts.
	Residual estimation method	This method calculates irregular migrant stocks by taking the total migrant count or foreign-born population from census data and subtracting the estimated number of legal migrants. The difference represents the irregular migrant population size.
Innovative	Alternative residual: labour force survey and social security register	This method is based on a comparison of employment data from surveys and official registers. When survey data shows more immigrant workers in certain sectors than official registrations, it suggests irregular migrant workers in those sectors.
	Digital data	
Innovative	Online search behaviour	Search engine data (like Google Trends) tracking searching behaviour of people, could employ information on trends for queries such as "will I be deported" to gauge policy impacts. Analysing search volume changes after events like the "Muslim ban" could help in specific groups of irregular migrants and their responses to policy shifts.

Table 5.1: Overview of methodological approaches covered in Rodríguez-Sánchez & Tjaden (2023) review (Note: Author's own elaboration)

Review criteria

Our review of each methodology was based on a series of criteria we deemed fundamental to understanding the scope of each method, meaning which population the method is able to produce estimates for, and the quality of its estimates. The selection of which features of the methodologies to highlight was based on existing common standards for the evaluation of scientific evidence, such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), the Grading of Recommendations, Assessment, Development and Evaluation (GRADE), among others (Guyatt et al., 2008; Page et al., 2021).

Each of the methods was evaluated in Rodríguez-Sánchez and Tjaden (2023) according to the following criteria:

- **Main Idea:** Explains the core concept of the method in simple terms for non-experts.
- **Data Source:** Identifies main data types (e.g., administrative, surveys, census) and gives examples.
- **Coverage / Definition:** Describes which sub-groups of irregular migrants are included or excluded, highlighting potential biases.
- **Estimation Assumptions:** Outlines key assumptions needed for the method to estimate the total irregular migrant population accurately.
- **Reliability:** Assesses whether the method gives consistent results over time.
- **Scalability:** Evaluates whether the method can be applied in different countries.
- **Ethical Issues:** Flags ethical concerns in data use, collection, and potential risks to migrants.
- **Examples:** Provides references to studies that apply the approach.

What we found

Among the traditional approaches, covering both indirect and direct approaches as classified by Jandl (2011), we review multiplier or simple extrapolation, the capture-recapture or multiple system estimation approach, the residual estimation method, self-identification in surveys, and expert or Delphi surveys. Although these methods possess important drawbacks we highlight in the literature, these methods are well-known and considered standard. We found important innovations regarding the multiplier, in Drbohlav and Lachmanová (2023), which document the results of implementing the multiplier in practice; and also, innovative work in the residual method, with the use of machine learning, larger government databases on social programs, and the evaluation of robustness of the method to core methodological assumptions (van Hook et al., 2021).

Moreover, among traditional approaches, we included the use of specific events, such as large-scale regularization and formal status adjustment programs (Sabater & Domingo, 2012), and life course events. Changes in legislation have offered the opportunity to understand the number of individuals lacking legal status in the past (Kraler, 2019). In turn, the life course events, in which administrative or register data sources on births, deaths, or hospitalizations, can offer important clues as to the sizes of populations as long as these can be extrapolated to the larger population. In particular, we highlight the potential of data on mortality (Surkyn et al., 2023), an approach which holds promise to be implemented in various countries relying on similar data.

Another class of traditional approaches we cover follows statistical modelling practices. For example, labour demand and the flow-stock modelling. In the labour demand modelling (Hess, 2006), only irregular migrant workers are estimated on the basis of reported economic output based on administrative data. In the flow-stock model (Fazel-Zarandi, Feinstein & Kaplan,

2018; Rodilitz & Kaplan, 2021), in turn, used in the US, and which have been criticized for providing implausible estimates that go orders of magnitude beyond existing estimations (Capps et al., 2018), information on cumulative inflows (visa overstayers, irregular border crossings/apprehensions) minus cumulative outflows (deportations, voluntary emigration, mortality, status changes) is used to derive an estimate.

Finally, among traditional approaches we also included the use of official, administrative, and commercial databases that allow for the estimation of irregular flows or stocks. Although this could be considered partly innovative, as new data sources have become available, irregular border crossings (Savatic et al., 2021; FRONTEX, 2022), data bases on asylum claims and refugee status (Ghui & Blangiardo, 2019), migrant deaths and apprehensions, and database systems enabling identification of visa overstayers. In the US, for example, overstayer events are estimated by considering all arrivals through air, sea, and land, matched to records of exits (Department of Homeland Security, 2022; Warren, 2017).

Advances and challenges

Among the innovative approaches, we document the development of important approaches. For example, in terms of databases, institutional registers on college enrolment (Hsin & Reed, 2020), as well as driver license register data (Lueders & Mumper, 2022), and consular registers of migrant communities (Bhandari et al., 2021). On their own, these databases cannot, on their own, be used to estimate the total number of irregular migrants, but when put in combination with other data sources they have the potential to provide important clues about irregular migrant population size.

On the more methodological side, the use of statistical imputation in large databases, often in connection to the residual method (Gálvez-Iniesta, 2020), constitutes an important innovation worth mentioning (Borjas & Cassidy, 2019; Ro & van Hook, 2022). Also, the use of innovative data sources on consumption and online search behaviour can be highlighted among the most important innovations (Nixon, 2022; Böhme, Gröger, & Stöhr, 2020). New databases that also inform irregular migrant flows, for example, have been created as a result of important citizen-driven projects (UNITED for Intercultural Action. (n.d.), or the creation of more encompassing surveillance programs. An example of the first is the Missing Migrants Project collecting data on the population of migrants dead and presumed missing while en route (García-Borja & Black, 2022). Knutson (2021) and Molnar (2019) discuss the uses of socio-technical systems based on artificial intelligence that enable facial recognition in the enforcement of migration.

Significant progress has been achieved in refining estimation methods to address the shortcomings of traditional techniques like the residual method. These improvements have been driven by the adoption of novel methodologies and the growing availability of diverse data sources (Vespe et al., 2017). However, estimation approaches remain largely fragmented, often shaped by the specific type of data source employed.

References

handari, R., Feigenberg, B., Lubotsky, D., & Medina-Cortina, E. (2021). Projecting Trends in Undocumented and Legal Immigrant Populations in the United States (No. NB21-16). NBER. https://www.nber.org/programs-projects/projects-and-centers/retirement-and-disability-research-center/center-papers/nb21-16

Borjas, G. J., & Cassidy, H. (2019). The wage penalty to undocumented immigration. Labour Economics, 61, 101757. https://doi.org/10.1016/j.labeco.2019.101757

Böhme, M. H., Gröger, A., & Stöhr, T. (2020). Searching for a better life: Predicting international migration with online search keywords. Journal of Development Economics, 142, 102347. https://doi.org/10.1016/j.jdeveco.2019.04.002

Capps, R., Gelatt, J., Van Hook, J., & Fix, M. (2018). Commentary on "The number of undocumented immigrants in the United States: Estimates based on demographic modeling with data from 1990-2016". PloS one, 13(9), e0204199. https://doi.org/10.1371/journal.pone.0204199

Department of Homeland Security. (2022). FY 2020 Entry/Exit Overstay Report (Entry/Exit Overstay Report). Department of Homeland Security. https://www.dhs.gov/sites/default/files/2021-12/CBP%20-%20FY%202020%20Entry%20Exit%20Overstay%20Report_0.pdf

FRONTEX. (2022, October 7). Risk Analysis for 2022/2023. Publications. https://frontex.europa.eu/publications/risk-analysis-for-2022-2023-RfJIVQ

Gálvez Iniesta, I. (2020). The size, socio-economic composition and fiscal implications of the irregular immigration in Spain [WorkingPaper]. https://e-archivo.uc3m.es/handle/10016/30643

Ghio, D., & Blangiardo, G. C. (2019). Exploring the link between irregular migration and asylum: the case of Italy. Genus, 75(1), 14. https://doi.org/10.1186/s41118-019-0060-3

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ, 336(7650), 924–926. https://doi.org/10.1136/bmj.39489.470347.AD

Hsin, A., & Reed, H. E. (2020). The Academic Performance of Undocumented Students in Higher Education in the United States. International Migration Review, 54(1), 289–315. https://doi.org/10.1177/0197918318825478

Jandl, M. (2011). Methods, Approaches and Data Sources for Estimating Stocks of Irregular Migrants: Estimating stocks of irregular immigrants. International Migration, 49(5), 53–77. https://doi.org/10.1111/j.1468-2435.2011.00701.x

Kraler, A. (2019). Regularization of Irregular Migrants and Social Policies: Comparative Perspectives. Journal of Immigrant & Refugee Studies, 17(1), 94–113. https://doi.org/10.1080/15562948.2018.1522561

Kraler, A., & Ahrens, J. (2023). Conceptualising migrant irregularity for measurement purposes. In MIrreM Working Paper No. 2 (Version 2). Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.7788961

Lueders, H., & Mumper, M. (2022). Driving while unauthorized: Auto insurance remains unchanged when providing driver licenses to unauthorized immigrants in California. Journal of Risk and Insurance, 89(3), 669–696. https://doi.org/10.1111/jori.12382

Nixon, S. (2022). Of Rice and Men: Rice Consumption-Based Estimates of Undocumented Persons in Malaysia. International Migration Review, 019791832211264. https://doi.org/10.1177/01979183221126466

Pinkerton, C., McLaughlan, G., & Salt, J. (2004). Sizing the illegally resident population in the UK (No. 58/04). Migration Research Unit University College London. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=949bc2f103e29ec74a63070eaf24fd4f9fccdb4b

Sabater, A., & Domingo, A. (2012). A New Immigration Regularization Policy: The Settlement Program in Spain. International Migration Review, 46(1), 191–220. https://doi.org/10.1111/j.1747-7379.2012.00884.x

Savatic, F., Thiollet, H., Jualin, T., Mesnard, A., & Senne, J.-N. (2021). Borders Start with Numbers. Measuring Migration in Times of Crisis (MigrAtion Governance and AsYlum Crise-MAGYC D 8.2). https://www.magyc.uliege.be/upload/docs/application/pdf/2021-12/d8.2v1october2021.pdf

Surkyn, J., Ahmad Yar, A. W., Ceylan, D., Vanthomme, K., & Bircan, T. (2022). New publication 'An analysis of mortality rates for estimating undocumented migrants in Belgium' — Hummingbird (HumMingBird Project 870661 – H2020). Retrieved 28 September 2023, from https://hummingbird-h2020.eu/news/news-items/D7.2

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, September 7). Datasets on Irregular Migration and Irregular Migrants in the EU. JRC Publications Repository. 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

Fazel-Zarandi, M. M., Feinstein, J. S., & Kaplan, E. H. (2018). The number of undocumented immigrants in the United States: Estimates based on demographic modeling with data from 1990 to 2016. PLOS ONE, 13(9), e0201193. https://doi.org/10.1371/journal.pone.0201193

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. International Journal of Surgery, 88, 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Ro, A., & Van Hook, J. (2022). Comparing the Effectiveness of Assignment Strategies for Estimating Likely Undocumented Status in Secondary Data Sources for Latino and Asian Immigrants. Population Research and Policy Review, 41(2), 449–464. https://doi.org/10.1007/s11113-021-09658-3

Rodilitz, S., & Kaplan, E. H. (2021). Snapshot Models of Undocumented Immigration. Risk Analysis, 41(9), 1643–1661. https://doi.org/10.1111/risa.13658

Rodriguez Sanchez, A., & Tjaden, J. (2023). Estimating Irregular Migration – A Review of Traditional and Innovative Methods. In MIrreM Working Paper No.4 (version 2). Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.8380854

Hess, S. (2006). The demand for seasonal farm labor from Central-and Eastern European countries in German agriculture (CIGR Journal). Institute of Agricultural Economics. https://cigrjournal.org/index.php/Ejounral/article/download/647/641/0

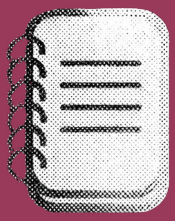
UNITED for Intercultural Action. (n.d.). About the “UNITED List of Refugee Deaths”. UNITED Against Refugee Deaths. Retrieved June 9, 2025, from UNITED for Intercultural Action’s website

Chapter 6

Data traces and the inevitable visibility of irregular migration

Alejandra Rodríguez Sánchez

Data traces and the inevitable visibility of irregular migration



Key points

- Irregular migrants, like all people, leave behind data traces – through surveys to digital activity and use of public services. These traces can indirectly signal their presence, even when their legal status remains unknown or unrecorded.
- This chapter explores how metadata and alternative data sources, such as social media, mobility records, and mortality registers can complement traditional statistics and offer new ways to estimate irregular migration.
- Pilot studies from the MIrreM project show that while these approaches are limited and ethically complex, they can be adapted and combined to improve understanding of irregular migrants, including their presence, behaviours and needs.

Introduction

One of the greatest challenges in researching irregular migration is the lack of reliable and up to date information on migrants and their legal status. For example, existing estimating methods to learn about irregular migration stocks and flows and the estimates they produce lack a reference value that can serve as a "gold standard" precisely because data on irregular migrants is incomplete. In addition to the inherent difficulty of studying mobile populations, researching individuals with irregular status is further complicated by the risks they face if their identity or legal status is revealed. The threat of deportation and the dangers of being forced to return to a place they may have fled for economic, political, or personal reasons create strong, legitimate disincentives to sharing information on their legal status.

Nevertheless, despite these concerns, data on migrants with irregular status is passively extracted, and collected by researchers, civil organizations, private companies, and immigration authorities, on a daily basis. This chapter offers a brief reflection on the differences and similarities in which this kind of data is being extracted, the potential uses researchers can make of this information, and especially the approaches we employed in MIrreM.

Metadata, data about data that offers context rather than direct description, can, in the context of migration, though not originally intended for migration research, serve as valuable indirect indicators of irregular migration patterns, behaviours, and trends. In this way, metadata can act as a complement to traditional sources like surveys, censuses and population registers to enhance our understanding of this complex issue. Like everyone else, irregular migrants leave traces of their presence within a country, especially in today's highly digital and technologically mediated environments. Whether making purchases, applying for a driver's license, visiting a doctor, responding to a government survey, or logging onto social media, irregular migrants generate records that document their presence—often without their knowledge or intent. To better understand the population size of irregular migrants, as well as their needs and living conditions, researchers can tap into the data generated by such interactions, in combination with official data. Essentially, this viewpoint aligns in part with indirect demographic methods but extends beyond them by embedding the indirectness within the data itself, rather than relying on underlying assumptions or parameters.

Box 6.1: Metadata in the context of migration

Alejandra Rodríguez Sánchez

Metadata refers to data about data—information that provides context, structure, or insight into how data is produced, rather than describing the subject directly. In the context of migration, metadata includes digital traces such as social media logins, geolocations, IP addresses, mobile network records, and internet searches. These data can signal movement patterns, border crossings, or changes in residence. Though not originally intended for migration research, such metadata—often derived from social media platforms, telecommunications networks, or administrative systems—can serve as valuable indirect indicators of irregular migration flows, behaviours, and trends. Even traditional sources like censuses, population registers, or household surveys can be viewed through this lens, offering a way to connect diverse data approaches and better understand the complex challenge of estimating irregular migration.

Mixed approaches combining direct and indirect data on migrants

The presence of irregular migrants is generally unknown to authorities. Contact between irregular migrants and immigration authorities may occur through various means, such as encounters at entry points using forged documents, random border checks, sea rescues, police operations, or interactions with other government agencies. However, the completeness of data collected through these operations has been criticized for relying heavily on the number of border personnel, thereby conflating enforcement activity with the actual presence of migrants (Savatic et al., 2024).

Population registers are another important source of official data on migrants. A key example is Spain's registration system, known as empadronamiento, which all residents are required to complete to access social services (Eniquez, 2019). However, due to recent legal changes, many irregular migrants may choose not to register, limiting the coverage of this data source, thus highlighting the voluntary aspect of participation (Velasco, 2021). Additionally, the use of data from registration systems is often restricted by data protection laws, which can prohibit the use of personally identifiable information or prevent the lawful matching of records across systems.

One of the most common ways researchers gather information on irregular migrants is through

sampling and household or individual surveys—an approach that has been used in past studies (Rodríguez-Sánchez & Tjaden, 2023; Jandl, 2011). These efforts often rely on institution-led data collection, such as censuses or household surveys. However, the legal status of migrant respondents—who belong to a hard-to-reach demographic—is frequently unknown, making the use of probabilistic sampling difficult. Surveys rarely include direct questions about legal status, as doing so may lower response rates or result in incomplete data (Bachmeier et al., 2014; Young & Madrigal, 2017). Notwithstanding these difficulties, other methods, such as the indirect method, which is used in a standard fashion to estimate the number of unauthorized migrants in the United States, relies on data from the American Community Survey (ACS), a large-scale household survey that includes all migrants. However, ACS underrepresent immigrants, especially newly arrived ones. Although efforts are made to adjust for this underrepresentation in the estimation of irregular migrants employing the residual method (van Hook et al., 2015).

Metadata, indirect approaches, and irregularity

Metadata can help in detecting and tracking population groups without requiring their direct participation or consent. For populations that actively avoid detection or are unlikely to participate in traditional surveys, such as irregular migrants or other marginalized communities, metadata may provide the only viable method for demographic observation and measurement.

Viewed from a certain perspective, this kind of information is already in use, though it is rarely labelled in these terms. For example, when the legal status of immigrants is not known, the use of statistical imputation has been employed to learn about the size and characteristics of irregular migrants. Statistical and machine learning methods can be used to predict what would be the most likely status of respondents based on the responses migrants give to other questions deemed to be less sensitive (Van Hook et al., 2015; Borjas & Cassidy, 2019), especially when such algorithms are trained in complete data (i.e., data capturing more direct evidence of irregularity). Therefore, even when not intending to answer questions about legal status, information provided by the correlation between legal status and other attributes of respondents can lead to plausible estimations.

The search for potential correlates or statistical signals of irregularity within existing data is an interesting approach. A compelling example of why it is helpful to view data on irregular migrants as metadata emerges when analysing the "digital traces" left by migration behaviour. Data from social media, internet searches, call detail records, economic transactions, and similar sources—often referred to in the recent past as "big data" sources to emphasize their greater volume, variety, variability and velocity of generation (Sîrbu et al., 2021), especially in migration (Tjaden, 2021)—are not originally created to study migration. Rather, this type of digital data arises from commercial or operational activities primarily intended for business or statistical purposes.

For example, META, a company owning some of the largest social media platforms in the world, offers the possibility to run targeted ads campaigns for marketing, political or civic purposes on Facebook, Instagram, Messenger or the Audience Network. To get these ads to the people most likely to be interested in the products or services advertised, META classifies its user base in various demographics, often without the awareness of the users themselves.

Attributes such as gender, age, location, interests, and places where the user has been or lived in the past, etc. are used to create targeted groups. It is this last attribute that researchers have employed to derive additional information on the stock of migrants in a given country (Zagheni, Weber & Gummadi, 2017). Recently, based on individual-level data, Chi et al. (2025) used Facebook profile locations to identify users who changed countries. If the change lasted most of the year, the authors classified users as migrants to the specific destination, while going beyond existing work on stocks by pairing the definition of migrant using duration of stay as a further criterion. In this way, researchers can circumvent asking respondents to disclose information about them and instead gather this information from behavioural clues.

Pilot studies in MIrreM

In a series of pilot studies developed in the context of MIrreM, we have made use of a similar intuition regarding metadata to develop innovative methodologies to estimate irregular migration. We used data from the following sources: Facebook users (Rodríguez-Sánchez & Tjaden, 2025a); digital surveys on US based Mexicans and Venezuelan users of social media platforms (Tjaden & Rodríguez-Sánchez, 2025); air passenger data collected on all flights across the world towards and outside of Europe (Bernasconi & Recchi, 2025); information on employment conditions of immigrants in the UK (Salihoğlu & Vargas-Silva, 2025); the effects of changes in laws governing access to healthcare through the National Health Institute (NHS) in the UK (Rodríguez-Sánchez & Tjaden, 2025b); and a matching of official registers capturing mortality and population in Belgium (Surkyn & Bircan, 2025), a method that shares many more commonalities with traditional approaches as discussed in Chapter 5).

The results of these pilot studies advance the literature in this domain and connect to a growing literature employing innovative data sources and methodologies to tackle the challenges of estimating irregular migration (Rodríguez-Sánchez & Tjaden, 2023).

Tjaden and Rodríguez-Sánchez (2025) demonstrated that Facebook ads can effectively reach irregular migrants and that list experiments provide more reliable estimates of legal status than direct questions in the United States context, particularly among well-establish immigrant groups such as Mexicans residing in the US.

Salihoğlu and Vargas-Silva (2025) demonstrated that some of the conceptual and measurement challenges in studying irregular migrants can be addressed through analysis of the informal economy and migrants' characteristics, using a clear conceptual framework and probabilistic tools such as national labour force surveys from Turkey and the UK to estimate their presence within it.

Rodríguez-Sánchez and Tjaden (2025b), in turn, showed that healthcare reforms in the UK resulted in a slight decline in new GP registrations at practices serving large migrant populations, which, when combined with arrival data at local levels, could be used to estimate undocumented migration flows using a multiplier approach.

Bernasconi and Recchi (2025) analysed net air travel flows in the Schengen area, using passenger data and adjusted official net migration figures, providing estimates of irregular inflows by region of origin for 2019, closely matching the notion of visa overstayers.

Drawing on Belgian population register data, Surkyn and Bircan's study (2025) shows that mortality rates can serve as a robust indicator for estimating the size and changes of irregular migrants over time, providing detailed insights by gender, age group, and even region of origin.

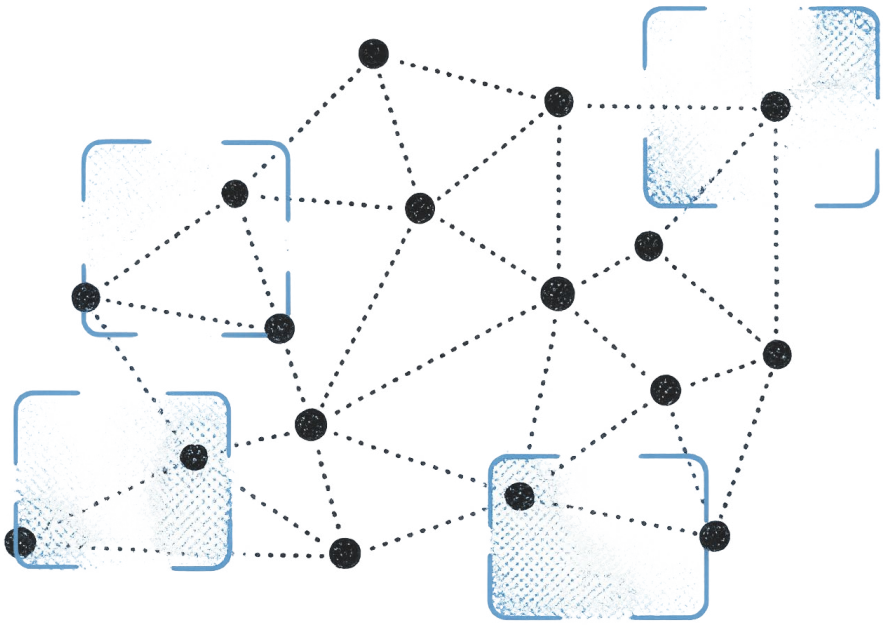
Finally, Rodríguez-Sánchez and Tjaden (2025a), in turn, showed that Facebook stocks of migrants, when examined by means of predictive modelling and machine learning, can provide hints at the hidden numbers not measured by official migrant stocks, offering a global comparison of irregular migrant stocks.

Conclusion

Computational approaches, of which digital data is an important component, have already enriched migration research (Drouhot et al., 2023). Thinking about data about migrants without a legal status as meta-data is helpful in understanding the information and in the development of further methods to estimate the number of irregular migrants. This type of data also comes with important limitations. The definition of what constitutes "migration", "place of birth", "migration status" etc. may vary across the different sources of metadata that exist and, importantly, may not be comparable to official or research standard definitions. These alternative data sources have not been created for research purposes.

Despite the insights generated by metadata, one major limitation is the inability to learn something more about the demographics and living conditions of the population of irregular migrants thus estimated. Moreover, there are important risks associated with employing what previous research as defined as "footprints" when not put into the larger context of statistical information on other population statistics (Gelatt, Fix, & van Hook, 2018), such as information on birth, death, school enrolment, housing, and other records.

Assessing when metadata is generated, whether it results from voluntary or involuntary actions, and understanding the potential coverage of such alternative data sources are key steps in determining how much insight can be gained from using innovative data in migration research. While each of these data traces can only offer a partial view of migrants with an irregular status, together these different sources of information underscore the inescapable visibility of irregular migrants and the potential to better understand their presence and the challenges such communities are facing.



References

Bachmeier, J. D., Van Hook, J., & Bean, F. D. (2014). Can we measure immigrants' legal status? Lessons from two US surveys. International Migration Review, 48(2), 538-566. https://doi.org/10.1111/imre.12059

Borjas, G. J., & Cassidy, H. (2019). The wage penalty to undocumented immigration. Labour Economics, 61, 101757. https://doi.org/10.1016/j.labeco.2019.101757

Drouhot, L. G., Deutschmann, E., Zuccotti, C. V., & Zagheni, E. (2023). Computational approaches to migration and integration research: promises and challenges. Journal of Ethnic and Migration Studies, 49(2), 389-407. https://doi.org/10.1080/1369183X.2022.2100542

Enríquez, C. G. (2019). Inmigración en España: una nueva fase de llegadas. Análisis del Real Instituto Elcano (ARI),(28), 1.

Savatic, F., Thiollet, H., Mesnard, A., Senne, J. N., & Jaulin, T. (2024). Borders Start With Numbers: How Migration Data Create "Fake Illegals". International Migration Review, 59(3), 1432-1463. https://doi.org/10.1177/01979183231222169

Rodriguez Sanchez, A., & Tjaden, J. (2023). Estimating Irregular Migration – A Review of Traditional and Innovative Methods. In MIrreM Working Paper No.4 (version 2). Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.8380854

Sîrbu, A., Andrienko, G., Andrienko, N., Boldrini, C., Conti, M., Giannotti, F., ... & Sharma, R. (2021). Human migration: The big data perspective. International Journal of Data Science and Analytics, 11, 341-360. https://doi.org/10.1007/s41060-020-00213-5

Tjaden, J. (2021). Measuring migration 2.0: a review of digital data sources. Comparative Migration Studies, 9(1), 59. https://doi.org/10.1186/s40878-021-00273-x

Van Hook, J., Bachmeier, J. D., Coffman, D. L., et al. (2015). Can we spin straw into gold? An evaluation of immigrant legal status imputation approaches. Demography, 52, 329–354. https://doi.org/10.1007/s13524-014-0358-x

Velasco, F. R. (2021). El arduo camino hacia la universalidad de la asistencia sanitaria de los inmigrantes irregulares en España. e-Revista Internacional de la Protección Social, 6(1), 343-369. http://dx.doi.org/10.12795/e-RIPS.2021.i01.16

Young, M. E. D. T., & Madrigal, D. S. (2017). Documenting legal status: A systematic review of measurement of undocumented status in health research. Public Health Reviews, 38, 1–25. https://doi.org/10.1186/s40985-017-0073-4

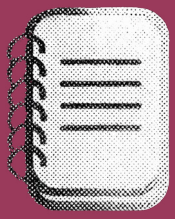
Zagheni, E., Weber, I., & Gummadi, K. (2017). Leveraging Facebook's advertising platform to monitor stocks of migrants. Population and Development Review, 721-734. https://doi.org/10.1111/padr.12102

Chapter 7

Register data sources on migrant stocks

Laura Peitz

Register data sources on migrant stocks



Key points

- Analysing irregular migrant stocks using register or other administrative data can prove challenging given the usually undocumented nature of the phenomenon; yet, some recent efforts highlight the potential of register data.
- The German Central Register of Foreigners (AZR) provides longitudinal data on non-Germans staying or having previously stayed in the country, including subsets of irregularly staying migrants, allowing for in-depth analyses of various research questions.
- Throughout, the chapter provides real-life examples of how register data has been used in irregular migration research across different countries.

Introduction

Counting the number of irregular migrants or more closely investigating this target group using administrative data presents profound challenges. In most national contexts, registration in population registers or similar systems is tied to a legal residence status. Consequently, migrants without the legal right to stay are typically excluded from such databases altogether. Moreover, even if they could be registered, irregularly staying migrants might deliberately avoid contact with public authorities to minimize the risk of detection and possible deportation. This further limits their visibility in administrative data systems. Single administrative data sources may contain data on irregular migrants interacting with particular public institutions – for example for schooling or urgent healthcare, or upon regularization or police force encounters – but this data is oftentimes incomplete, fragmented, inconsistent, and usually not linked to broader administrative registers.

While these obstacles pose significant barriers to the statistical inclusion of irregular migrant populations, some efforts have recently emerged to provide details on irregular migrant stocks based on administrative data sources (see also UNECE, 2025). In **Italy**, for instance, irregularly staying migrants can be identified by comparing data from various administrative sources and applying the Signs of Life method (see Box 7.1).

A similar approach has recently been applied in **Chile**, where the number of irregular migrants is estimated by integrating data from post-census administrative records on education, tourist stays, and police reports, and comparing this data against the baseline of residence permit applicants (see Box 7.2). In **Spain**, everybody is encouraged to register in the municipal population registers of their municipality (*Padrón Municipal*). The registration is a prerequisite for accessing basic rights and public services, such as health care and schooling. It is independent of legal status and – crucially – is not used for immigration control. As a consequence, the *padrones* even include irregularly staying migrants. When comparing or linking the *padrón* data to other administrative datasets, it is possible to assess questions around migrant irregularity, such as deriving the number of irregularly staying migrants from a comparison of the padrones with the database of legal stay permits (González-Enríquez, 2016). In the **UK** and **Poland**, recent efforts have been undertaken to produce a time series of the number of irregularly staying migrants based on the ethnic economies approach and non-linear count regression models. The assumption is that regularly settled ethnic groups provide support for individuals of similar ethnicity from their countries of origin to circumvent national restrictions on migration rules regarding work. Based on this, the numbers of detentions extracted from official police and border enforcement data are scaled up to the regularly residing foreign population using non-linear count regression models to estimate the number of irregularly staying migrants per country of origin (Beręsewicz, 2024). [1]

Box 7.1: Applying the 'Signs of Life' method: The case of Italy

Marco Marsili

In Italy, the National Institute of Statistics (Istat) identifies the number of irregularly staying migrants by applying the Signs of Life (SoL) approach. To this end, in a first step, data on migration (changes of residence) are drawn from the centralized population register (ANPR, managed by the Ministry of the Interior). These data are subjected to standard control and correction procedures. In general, the quality of the data is quite high; in case of partial non-response, the information is filled using donor hot-deck methods of imputation or by retrievals from the previous year's census, where available.

In a second step, the information of the ANPR is integrated into a demographic data system (Midea-Anvis, MIcro-DEmographic Account - Virtual Statistical register of the population) which, in addition to migrations, also incorporates data of other population changes (births, deaths, acquisitions of citizenship). Midea-Anvis is a counting system based on micro-data, in which all data are integrated with each other and with respect to the population of the last census, in order to verify the stock-flow coherence of the information acquired.

The last step is comparing Midea-Anvis with a large set of administrative archives (AIDA, Integrated Archive of Administrative Data), including, among others, the tax, social security, energy consumption, and education registers as well as the Cadastre of buildings and constructions. Each administrative archive in AIDA provides life signals on habitually resident persons who have spent a significant amount of time in Italy over the last three years. The comparison between AIDA and Midea-Anvis produces three distinct datasets:

1. individuals present in Midea-Anvis and confirmed as residents through the life signals system in AIDA (the so-called "usual resident population");
2. individuals not present in Midea-Anvis but with strong life signals in AIDA (under-coverage);
3. individuals present in Midea-Anvis but without life signals in AIDA (over-coverage).

1 This research was led by Brendan Georgeson (Office for National Statistics, UK) and Maciej Beręsewicz (Poznań University of Economics and Business, Poland).

The second of these datasets is relevant information about irregular migrants. It comprises all those individuals who, despite not having specific authorization to reside in the national territory (for example, because they have an expired residence permit) show signs of administrative life in Italy (e.g. because of working, studying or avwaiting residence permit renewal). Hence, as regards irregular or undocumented migrants, the current structure of the data production system allows to correctly focus on a specific group of irregularly staying migrants. the other side, the evaluation of fully undocumented migrants is most challenging as, by definition, they do not show any sign of life. Nonetheless, Istat also produces national estimates of fully undocumented migrants on a yearly basis. The sources used to produce these estimates have varied over the years, depending on data availability, including sample surveys and data from administrative sources. In recent years, the methodology has been improved by also integrating data of the Ministry of the Interior relating to police stops on the territory or at the border, as well as data relating to actual repatriations to countries of origin.

In contrast, **Germany** presents a unique case in this regard, as it has been using a Central Register of Foreigners (AZR) for over 70 years, which includes comprehensive data on the majority of non-nationals staying in the country, even parts of those without legal residency status. The reason is that in Germany, many migrants who are formally obliged to leave the country are issued a *Duldung* (tolerated status) while their removal is temporarily suspended due to either actual obstacles (e.g. missing travel documents or illness) or legal reasons (e.g. family unity) preventing deportation. These migrants are well captured in the Central Register of Foreigners, meaning that the register is suited for detailed analyses aimed at specific subgroups of irregular migrants. The rest of this chapter will present the Central Register of Foreigners and the potentials and pitfalls of using its administrative data in irregular migration research.

Box 7.2: Chile's experiences in integrating data for estimating the foreign population with irregular migration status[2]

Julibeth Rodríguez and Felipe Mallea

Since 2014, Chile has witnessed an increasing migrant flow, which has meant that the country must assume the challenge of officially measuring the phenomenon. To this end, the National Statistics Institute (INE) and the National Migration Service (Sermig) have developed a methodology for estimating the number of foreign nationals residing in Chile between censuses by linking border-control data with residency applications after the 2017 Census.

This study employs a methodology that integrates baseline data from the census with data from post-census administrative records by linking microdata of various government institutions. A critical component is using administrative records to identify populations with a potentially irregular migration status.

2 This work was conducted by the Studies Department at the National Migration Service of Chile and the Demography Subdepartment at the National Institute of Statistics of Chile (team members: Gabriel Santander, Consuelo Salas, Marisol Opazo, Pablo Roessler, Felipe Hugo, Luis Rodríguez, Miguel Ojeda, Francisco González). More details are available at: https://serviciomigraciones.cl/estudios-migratorios/estimaciones-de-extranjeros/ and https://www.ine.gob.cl/estadisticas/sociales/demografia-y-vitales/demografia-y-migracion.

The methodology acknowledges the complexities in precisely quantifying irregular migration, which are due to the multifaceted character and the dynamic fluidity of migration status. Consequently, the scope of the estimation is deliberately confined to a specific subset of the population with irregular status, a delimitation necessitated by the availability of relevant data sources and by the objective of clearly distinguishing the populations with regular and irregular migration status.

The estimation of the population with irregular status specifically includes individuals who do not possess any type of residency permit application, who have been in the country for at least six months, and who have no recorded departure for the period ending on December 31, 2023. The sources for the estimation are as follows:

1. The biometric control system (between June and December 2023)
2. Expired tourist visa extensions or police reports (including both formal denunciations for unauthorized entry and self-reported clandestine entries)
3. Primary and secondary student enrollment in Chilean educational institutions of those who are assigned a provisional identifier because they lack a national identification number (RUN)

By including a wide range of administrative records, we can account for the two main areas that form the basis of irregular migration of foreign nationals in Chile: (1) those who enter the country clandestinely and who cannot apply for a residency permit, and (2) those who enter the country legally and who cannot apply for a residency permit. With these two areas and their combination with records from border control, it can be determined whether the person was in the country for the period ending on December 31, 2023. In processing the data, 33,251 people who left the country were excluded. In contrast, we included those whose presence and residence in Chile were shown by the records of their administrative acts to be subsequent to their exit from the country.

The final dataset for the population with irregular migration status comprised 336,984 individuals, whose information was categorized by primary source: 261,449 from police reports, 10,217 from expired tourist visas, and 65,318 from official enrollments without a national identification number (RUN).

In conclusion, this study presents a methodology that contributes to international migration statistics by integrating census data with diverse administrative data. While recognizing the inherent uncertainties in estimating the population with irregular migration status, the results offer valuable insights for targeted public policy design and demonstrate potential for adaptation in other countries, which would thereby improve the quality and comparability of regional migration data.

Structure and contents of the AZR

The *Ausländerzentralregister* (AZR, Central Register of Foreigners) is Germany's primary administrative register for non-German nationals living in Germany. Established in 1953 and governed by law (*Gesetz über das Ausländerzentralregister*), the AZR plays a central role in federal and local migration governance, strategic planning as well as in daily migration-related administrative activities. The data is entered into the register primarily by local immigration offices (*Ausländerbehörden*) as well as other public institutions such as the Federal Office for Migration and Refugees (*Bundesamt für Migration und Flüchtlinge*, BAMF) or federal and state police forces. Various public authorities use

the data stored in the AZR to support their case-by-case decision-making, operational planning, and political decision-making. Over the past years, the register has become an important, though complex, information source on various migration-related questions for public administration, the government, the media, the broader public, and for researchers interested in understanding patterns and dynamics of migration and residence (Brückner, 2019; Peitz, 2025; Tanis, 2022; Weber, 2022).

The AZR covers all non-German nationals who reside in or have resided in Germany for more than three months. In addition, it includes data of individuals who have filed an asylum claim and of those who have been issued residence law decisions, such as expulsion or deportation orders. The data recorded is stored in the AZR for the duration of an individual's stay in Germany, and usually for ten years after their departure (five years after death). All data entries are deleted from the AZR upon naturalization, without the possibility of further tracking these individuals given the lack of a central population register in Germany.

The AZR contains various data attributes per individual. Which types of attributes are stored depends on the specific group of migrants. Only rudimentary information is stored in the case of EU citizens, while the most comprehensive data is collected on individuals entering the asylum system.

The variables contained include:

- Personal data: unique AZR identifier, full name, date of birth, gender, nationality, marital status
- Border crossings: entries, voluntary departure, forced return
- Residence status: temporary and permanent residence titles, *Duldung*, obligations to leave the country
- Asylum procedure: application filed, asylum status, rejection

In addition to this "core" data, the AZR has in recent years been expanded by multiple additional variables, including language skills, education and profession, postal address, and integration course information, but the quality of these variables varies (see below).

With exception of time-invariant personal data, data entries in the AZR are usually location- and time-stamped: They contain the date of the respective data entry as well as the municipal level of the executive authority (which usually corresponds to individuals' place of residence), along with the respective federal state. Importantly, whenever new information is entered for many of the 'core' variables, the previous data entry is not overwritten. Instead, all previous information on these variables is kept as long as an individual's data is stored in the AZR (see Gleiser & Hinz, 2024, p. 8). This way, the AZR data allows for longitudinal and flow analyses.

Irregular migration stocks and flows based on AZR data

The AZR can provide indicators on irregular migration stocks and flows. However, one needs to carefully delineate the groups of irregular migrants who are, and who are not, included in AZR data. Being an administrative register utilized and filled by public authorities, the AZR, virtually by definition, contains only data on migrants with contact to the authorities. Based on the MIrreM taxonomy (Kraler, 2023), the following groups of irregularly staying migrants (migrants with an obligation to leave the country) can be identified using the AZR: individuals who are issued a return decision, whose status is expired or revoked, and whose removal is formally suspended.[3] The following flows into and out of irregularity can be traced based on AZR data: inmigration, being born into an obligation to

3 Due to the specific filter functions in the AZR, identifying these groups is possible in the most current cross-sectional dataset. It is, however, not necessarily possible for all these groups retrospectively in the longitudinal dataset.

leave, loss of status, death, voluntary and enforced departure, and regularisation (including the permit issued when entering new asylum procedures).

Undocumented migrants who have never come into contact with authorities remain outside the scope of the AZR, making the register incomplete for fully irregular populations. Similarly, data for irregular migrants who are identifiable using the AZR is only up-to-date as long as individuals show administrative signs of life in the AZR or have undoubtedly terminated their irregular status (by a registered departure or by regularization). In the case of individuals who no longer show signs of life in the AZR and who do not have a documented termination of irregularity, it is not always clear from the data whether they have left the country without official knowledge or gone into hiding.

However, given Germany's *Duldung* system and the wide spread of this provisional 'status', which regularly needs to be renewed (Schütze, 2023), many migrants staying in Germany irregularly are in contact with the authorities and have recurrent positive data entries in the AZR. In addition, a large group of irregularly staying migrants in Germany are rejected asylum seekers, who always have been in contact with authorities at some point in time, and, when possessing a *Duldung*, subsequently are. As a consequence, the AZR contains the entire residence history of virtually the full sample of individuals obliged to leave the country and showing signs of life in the AZR (e.g. by being in possession of a *Duldung*). And, given AZR storage policies, it contains not only data of those present in Germany at the time of data extraction, but also of those who have been present within the previous ten years.

This highlights the unique advantages of using AZR data in analyses on irregularly staying migrants: Being able to analyse close to a full sample of the specified sub-group, usually relatively promptly and without temporal delay, and in greater detail than with other administrative data sources. This includes the availability of time- and location-stamped data allowing for longitudinal analyses with a dynamic perspective, such as analyses related to the duration of stay or timing of key transitions in great temporal granularity (for an example, see Box 7.3). The AZR can also be used to investigate period effects (such as new laws, political or societal events) with a quasi-experimental framework (e.g. Peitz & Carwehl, 2025). Also, the AZR permits interregional comparison or the assessment of interregional mobility patterns, and serves as a solid sampling frame to draw representative samples for survey research.

Box 7.3: Using AZR data to analyse pathways out of irregularity: An application example

Laura Peitz

A recent analysis of pathways out of an obligation to leave presents a compelling case for using longitudinal administrative data to understand temporal dynamics in irregular migration (Peitz, 2025). Drawing on data from the German Central Register of Foreigners (AZR) between 2013 and 2022, the study examines the trajectories of over 400,000 individuals whose asylum claims were finally rejected. This rich dataset allows for a nuanced temporal analysis of how individuals exit irregularity – through voluntary departure, deportation, or regularization.

By applying event history analysis within a competing risk framework, the study shows that temporal dynamics significantly shape outcomes. Voluntary departures are most likely in the first two years following a final rejection, while regularizations increase in likelihood with the length of stay. In contrast, deportations remain relatively rare and largely concentrated in the early years following a final rejection.

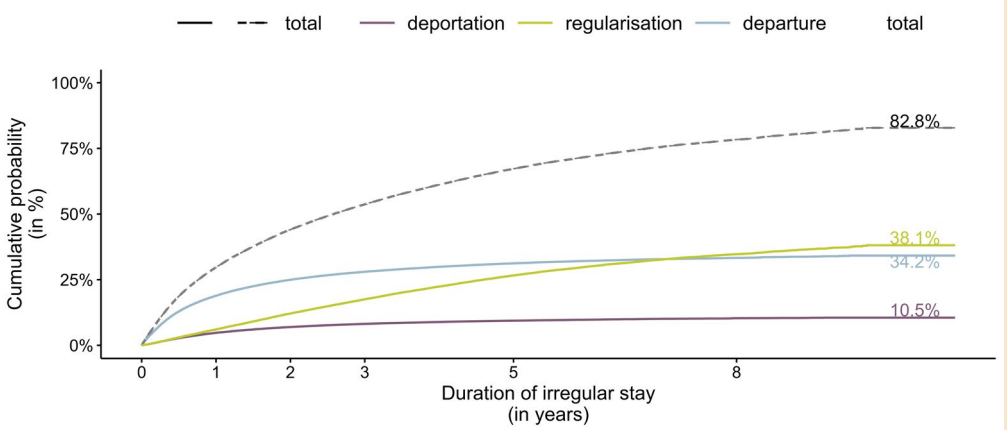


Figure 7.1: Cumulative incidence of competing pathways out of irregular status (Source: Peitz 2025, p. 19)

The analysis also reveals how policy instruments interact with time. For instance, the impact of different types of Duldung is distinctly time-sensitive: the restrictive 'Duldung light' for individuals with unclear identity tends to prolong irregularity, while employment and training-related types of Duldung facilitate regularization – but only after years of legal limbo. The study also shows that while designating countries of origin as "safe" aims to accelerate return, the actual timing of exits varies more by structural factors than by policy labels alone.

In sum, administrative longitudinal data unlocks critical insights into when and how irregular migrants transition out of legal limbo. Such data enables the evaluation of migration policies over time, offers evidence for reforming regularization schemes, and underscores the importance of integrating temporal dimensions into migration research and governance.

References:

Peitz, L. (2025). Return or regularization? A temporal analysis of rejected asylum seekers in Germany (RSC Working Paper 2025/06). https://hdl.handle.net/1814/78047

At the same time, working with AZR data for scientific purposes comes with some pitfalls that need to be carefully addressed. The AZR is not primarily a scientific database. Rather, its structure reproduces the legal complexity of Germany's migration governance system, and data quality can reflect the local administrative capacity of individual authorities and the effectiveness of data exchange between agencies. For instance, some variables, such as residence permits, can carry a multitude of values that need to be combined into meaningful categories during the data management phase.

Despite the recent introduction of automated data entries, implausibility checks during data entry, and automated interfaces between the applications used by different agencies, some issues potentially impairing data quality remain. These can be due to delayed reporting, incomplete status transitions, differences in interpretation or practices across federal and local agencies, or incorrect data entry practices. As a consequence, the raw data can show implausible entries that need to be carefully addressed in data cleaning and processing. Therefore, both for data management and for the interpretation of data and results, detailed knowledge on the legal and administrative procedures is needed.

Data quality is usually fairly high on 'core' variables, as they are vital for administrative processes and their systematic collection is legally required. In contrast, (some of the) additional variables are not essential for daily administrative tasks, their data collection is legally required only for specific groups (e.g. asylum seekers), or they are handled using other applications than the AZR. Data on these variables is therefore not necessarily entered and kept in the quality that would be necessary for scientific purposes. When wanting to exploit information of such variables, researchers should critically balance the data's informational value against potential bias.

Finally, when publishing and communicating results that are based on register data such as the AZR, it is crucial to clearly state the specific subgroup of irregularly staying migrants that are covered by the data and to be transparent about the informed but inevitably arbitrary decisions made during the data cleaning process. Also, in light of potential implausibilities and case incompleteness, it may be more appropriate to highlight patterns, interrelationships, and dynamics rather than presenting absolute figures that risk conveying a false sense of certainty.

Conclusion

The AZR systematically and comprehensively captures several subgroups of irregular migrants – a category of persons who are usually not well captured by administrative data. Despite the challenges that arise when using such administrative data for scientific analysis, the AZR remains one of the most wide-ranging administrative databases on non-nationals in Europe, and a valuable source for informative irregular migration analyses. In the future, record linkage with other administrative data sources could even enhance the potential of AZR data for research around irregular migration. Prerequisite for such record linkage is a legal basis as well as a careful design in accordance with ethical considerations and data protection regulations.

The example of the AZR also shows the added value of making administrative migration databases exploitable for more advanced statistical research, such as longitudinal analyses. For this reason, a random sample of the AZR data is now available to researchers for scientific purposes, via the Research Data Centre of the German Federal Office for Migration and Refugees (Gleiser & Hinz, 2024; Gleiser et al., 2024).

References

Beręsewicz, M. (2024). Estimation of the number of irregular foreigners in Poland using non-linear count regression models. Economic Commission for Europe.

Brückner, G. (2019). Das Ausländerzentralregister als Längsschnittdatenquelle. WISTA, 2019(1), 35–67. https://www.destatis.de/DE/Methoden/WISTA-Wirtschaft-und-Statistik/2019/01/auslaenderzentralregister-laengsschnittdatenquelle-012019.pdf?__blob=publicationFile&v=1

Gleiser, P., & Hinz, T. (2024). Datenreport zum AZR-Forschungsdatensatz 2023 (Veröffentlichungen des BAMF-FDZ 01/2024). https://doi.org/10.48570/bamf.fdz.voe.01/2024.d.2024.azrdatenreport.1.0

Gleiser, P., Salas Escobar, C., & Hinz, T. (2024). Codebuch zum AZR-Forschungsdatensatz 2023 (Veröffentlichungen des BAMF-FDZ 02/2024). https://doi.org/10.48570/bamf.fdz.voe.01/2024.d.2024.azrcodebuch.1.0

González-Enríquez, C. (2016). Spain: Irregularity as a Rule. In A. Triandafyllidou (Ed.), Irregular Migration in Europe: Myths and Realities (pp. 247–266). Routledge. https://doi.org/10.4324/9781315589848

Kraler, A. (2023). Taxonomy of Migrant Irregularity (MIrreM Deliverable D1.2). University for Continuing Education Krems. https://doi.org/10.5281/zenodo.7875543

Peitz, L. (2025). Return or regularization? A temporal analysis of rejected asylum seekers in Germany (RSC Working Paper 2025/06). https://hdl.handle.net/1814/78047

Peitz, L., & Carwehl, A.-K. (2025). Zwischen Duldung und Bleiberecht: Wie und für wen wirkt das Chancen-Aufenthaltsrecht? (BAMF-Kurzanalyse 03/2025). https://doi.org/10.48570/bamf.fz.ka.03/2025.d.2025.chaufr.1.0

Schütze, T. (2023). The (Non-)Status of 'Duldung': Non-Deportability in Germany and the Politics of Limitless Temporariness. Journal of Refugee Studies, 36(3), 409–429. https://doi.org/10.1093/jrs/feac056

Tanis, K. (2022). Die Wohnhistorie Geflüchteter in Deutschland (BAMF Kurzanalyse 01/2022). https://nbn-resolving.org/urn:nbn:de:0168-ssoar-79874-5

UNECE. (2025, March 12). Measuring Hard-to-Reach Groups in Administrative Sources. United Nations Economic Commission for Europe. https://unece.org/sites/default/files/2025-04/HardToReachGroups%20Consult.pdf

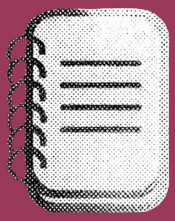
Weber, J. (2022). Binnenmobilität von Geflüchteten mit Schutzstatus in Deutschland: Eine explorative Analyse auf Basis des Ausländerzentralregisters (Forschungsbericht No. 39). https://www.bamf.de/SharedDocs/Anlagen/DE/Forschung/Forschungsberichte/fb39-binnenmobilitaet.pdf?__blob=publicationFile&v=10

Chapter 8

Getting into the flow - what do we know now, 15 years since CLANDESTINO?

Lalaine Siruno and Arjen Leerkes

Getting into the flow - what do we know now, 15 years since CLANDESTINO?



Key points

- Many countries have stock estimates of irregular migrants, but flow estimates are still scarce.
- Irregular migration flows are more often measured through statistical indicators.
- Compared to CLANDESTINO, the MIrreM Project has found that there are now more irregular flow indicators, particularly for geographic flows, and to some extent, also asylum-related status flows.
- The available EU-level indicators for irregular flows are generally of good quality, particularly with respect to accessibility and documentation, but there are some limitations in terms of validity and reliability.

Introduction: Counting the uncountable, measuring the unmeasurable

Immigration, particularly irregular migration, has become increasingly politicised, largely due to its entanglement with debates over national security, economic pressures, and cultural identity. In Europe, irregular migration has emerged as a "major issue" since the 1990s (De Genova, 2002; Sassen, 1999). In the EU in particular, it has dominated the policy landscape since 2015, following the large-scale arrival of asylum seekers from Syria, Afghanistan, and other countries (Siruno et al., 2024). Irregular migration has been conflated with asylum migration (Cantat et al., 2023), and public discourse often spotlights the perceived large "flows" of migrants, framing migration as a problem to be managed rather than a complex social phenomenon.

This reality underscores the importance of collecting good-quality data on the size of the irregular migrant population, not only stocks (i.e., the total number of irregular migrants residing in a particular location at a specific point in time, offering a snapshot of the migrant population but also flows), but also flows (i.e., the movement of irregular migrants over a defined period, capturing arrivals, departures, and net migration, providing a dynamic perspective on migration patterns and

trends). The first large-scale EU-funded project to do so was CLANDESTINO (Undocumented Migration: Counting the Uncountable: Data and Trends Across Europe)[1], which ran from 2007 to 2009. The final report presented the following conclusions (Jandl et al., 2008, p. 17):

The review of efforts to estimate the size of irregular migration on a European level has shown that the numbers indicated are based on very rough estimates. Often, we do not know which groups of irregular migrants are in [sic] included in a stock estimate, nor we do not know whether a flow estimate is meant to measure net inflows or gross inflows (without substraction [sic] of outflows).

Jandl (2008, p. 20) further pointed out that compared to stocks, flows are generally not well-measured:

...Given the highly volatile nature of migration flows, the scarcity of reliable indicators on illegal migration flows, and the dearth of appropriate methods for estimating such flows, most efforts have concentrated on estimating stocks of undocumented migrants rather than flows.

Now over a decade since CLANDESTINO, and with managing irregular migration flows a mainstay policy priority in the EU and other countries, this chapter outlines the main findings from the MIrreM Project's Work Package on Flows (WP4). More specifically, it provides a summary of the current approaches to measuring irregular migration flows, and addresses the question: what do we know now about irregular migration flows, 15 years since CLANDESTINO?[2]

Expanding the temporal and geographic scope, improving the quality assessment criteria

MIrreM is a follow-up to CLANDESTINO, and the following Table shows a basic comparison between the two projects in terms of timelines and geographic coverage:

	CLANDESTINO	MIrreM
Implementation period	2007-2009	2022-2025
Timeframe covered	2000-2007	2008-2023
Countries covered	12	20
	EU countries: Austria, Czech Republic, Germany, Greece, France, Hungary, Italy, Netherlands, Poland, Slovakia, Spain, and UK	**EU countries:** Austria, Belgium, France, Finland, Germany, Greece, Ireland, Italy, Netherlands, Poland, Portugal, and Spain **Other countries:** Bosnia and Herzegovina, Canada, Morocco, Serbia, Tunisia, Türkiye, UK, and USA

Table 8.1: Basic comparison between the CLANDESTINO and MIrreM Projects

1 https://irregular-migration.net/

2 This chapter draws mainly from the following WP4 deliverable, which was published in 2024, hence, 15 years since the conclusion of the CLANDESTINO Project in 2009: Siruno, L., Leerkes, A., Hendow, M., & Brunovská, E. (2024). MIrreM Working Paper on Irregular Migration Flows. University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.10702228

A notable difference is the inclusion of non-EU countries in the MIrreM project. And as discussed in detail in Chapter 4 of this Handbook, MIrreM has developed and used a more structured set of criteria to assess the quality of irregular migration data.

In addition, for irregular migration flows in particular, the MIrreM project highlighted the distinction between estimates and indicators. Estimates refer to statistical calculations or approximations that quantify both observed and non-observed or unknown irregular migration flows. Indicators, on the other hand, refer to metrics or variables that relate only to observed or measured irregular migration flows. In other words, indicators of irregular migration flows show the number of actual observations or cases, such as detections of illegal border crossings, whereas estimates use indicators to come to conclusions about a broader trend, including non-observed components, such as the total number of adults, detected and undetected, who crossed into a country without the legal right to do so. Two related but different sets of criteria were developed to assess the quality of irregular flow estimates and indicators.[3]

What we know now about irregular migration flows, 15 years since CLANDESTINO

Post-CLANDESTINO, scholars observe that available migration data often remain "inaccurate, inconsistent and incomplete" as they are based on differing definitions (Bijak et al., 2019, p. 471). In addition to differing definitions and measures, there are persistent and interlinked gaps based on the drivers or reasons behind migration, geographic coverage, demographic characteristics, and time lag in the availability of data (Ahmad-Yar & Bircan, 2021). International migration flows are particularly difficult to measure, and this is the case even with advancements in technology and data science (McAuliffe & Ruhs, 2017). Several international organisations, including UN DESA and the OECD, have been collecting and publishing international migration flows data, but different definitions and data collection methods present challenges in harmonisation and comparability (Yildiz & Abel, 2021). As there is an inherent challenge in collecting data on clandestine or irregular processes, the difficulties are even more pronounced when capturing data on irregular migration flows (McAuliffe & Sawyer, 2021, p. 48, emphasis added). So, while many countries have available stock estimates, there is persistence in the scarcity of available flow estimates as observed in the CLANDESTINO Project.

Because of this, irregular migration flows are more often measured through statistical indicators, particularly geographic and status-related flow indicators. Table 8.2 below provides a summary of the findings from the CLANDESTINO and MIrreM projects related to different types of irregular migration flows. In view of findings from CLANDESTINO, the conclusion reached then, namely that the methodologies for analysing irregular border crossings, visa overstays, and overall irregular migration flows lag behind the study of irregular resident stocks (Vogel et al., 2008), still rings true. However, the MIrreM Project has found that there are now more irregular flow indicators, particularly for geographic flows, and to some extent, also asylum-related status flows.

3 However, and as this piece underscores, compared to stocks, there is a notable lack of available estimates on irregular migration flows.

Type of flow	**CLANDESTINO** (CLANDESTINO Project, 2009)	**MIrreM** (Siruno et al., 2024)
Demographic	Quantitative importance is low or largely unknown but 'causing considerable human rights concerns.'	There is still a scarcity of data concerning demographic flows. While birth and death registrations are fundamental aspects of civil registration systems across the world, there is a noticeable lack of systematic data on births and deaths and irregularity.
Geographic	Most visible flows and border guard apprehensions are commonly used as indicators. However, published data from countries 'have not achieved full comparability' and there is 'even less information' on outflows (e.g., departures from the EU, movements to another EU country)	Geographic flows are still the most visible flows and there are now more geographic flow indicators. From 2021 for example, in addition to third-country nationals (TCNs) refused entry at the external borders and returned following an order to leave, the enforcement of migration (EIL) statistics collected by Eurostat also include the following (Eurostat, n.d.): • TCNs found to be illegally present by 'place of apprehension' and by the 'grounds of apprehension' • Quarterly rather than annual statistics on returns, and mandatory breakdowns by type of return, type of assistance received and destination country • Unaccompanied minors ordered to leave and on unaccompanied minors who returned following an order to leave.
Status-related	EU enlargement and regularisation programmes in Italy, Spain and Greece resulted in status-related outflows being far higher than inflows. Visa overstaying is the most relevant inflow indicator.	There are now also more data on asylum-related status flows. The collection of asylum statistics on the EU-level has also evolved over time. In 2021 for example, the implementation of the new Asylum Guidelines also involved the collection of data on the reason for the decision withdrawing status (revocation, ending, refusal to renew, unknown) in addition to first instance or final negative decisions and decisions withdrawing status granted at first instance or as final decision.

Table 8.2: Summary of flow trends from the CLANDESTINO and MIrreM Projects

Box 8.1: Frontex data on "illegal border crossings" and the political construction of "illegal" immigration

Filip Savatic

Since 2009, Frontex, the Border and Coast Guard Agency of the European Union (EU), has published a dataset on "illegal border crossings" (IBCs) into the EU and Schengen Area which is publicly accessible through the institution's website.[4] This dataset was initially labelled "irregular border crossings" until 2022, with the change reflecting a striking shift. Over time, particularly after the so-called "migration crisis" of 2015, this dataset has been increasingly referenced by mainstream media, researchers, international organizations, and other actors as a measure of "illegal" migration to Europe.

However, the use of these data as an indicator of irregular migration is problematic for several reasons. First, they capture only detected entries, and may, depending on type of border and context, represent an undercount of actual crossings. Second, they represent crossings and not people and thus may record repeat crossings made by the same individual multiple times, leading to an overcount of movements. Most importantly, the database does not consider valid protection claims of those detected while irregularly crossing a border. As article 31 of the Geneva Refugee Convention states, irregular entry is permitted when individuals are fleeing persecution (United Nations, 1951/1967). Given the absence of legal pathways for refugees to reach Europe, most asylum seekers reach the continent without any prior authorization, with many subsequently obtaining refugee status.

Deploying a novel method, Savatic et al. (2024) use data on asylum adjudications across 31 European states to divide Frontex data on IBCs into those who would likely obtain refugee status (or not) given their nationality. The average acceptance rate is weighted given the number of first instance asylum decisions by nationality made in each of the 31 states. First instance data are used to ensure comparability given that asylum appeals procedures vary across states; using these data generates a conservative estimate of asylum acceptances as only rejections are overturned. This division of IBCs reveals that, between 2009-2021, 55.4% can be considered "likely refugees," a proportion that rises to 75.5% at the peak of arrivals in 2015. With most IBCs representing forced migration flows considering the asylum policies implemented domestically within Europe, the use of data on border crossings as an objective measure of "illegal" migration is misplaced.

Overall, this analysis exposes how data can be – and are – deployed to further certain public narratives and thereby represent political constructions rather than objective truths. In the case of data on border-crossings collected by law enforcement agencies such as Frontex, narratives of "illegal" migration flows construct an understanding of border crossings as something which requires a securitized response – one that law enforcement bodies can provide. Alternative labelling such as "forced" migration would imply that humanitarian responses to migration flows would be more appropriate. Thus, it is imperative for news media, researchers, and all other public authorities to adopt a critical approach to data, questioning what they represent and what purpose they serve for those who collect and publish them.

References

Savatic, F., Thiollet, H., Mesnard, A., Senne, J.-N., & Jaulin, T. (2024). Borders Start With Numbers: How Migration Data Create "Fake Illegals". International Migration Review, 59(3), 1432-1463. https://doi.org/10.1177/01979183231222169 (Original work published 2025)

United Nations (1951/1967). Convention relating to the status of refugees and protocol relating to the status of refugees. United Nations Treaty Series, 189 U.N.T.S. 137; 606 U.N.T.S. 267. https://www.unhcr.org/media/1951-refugee-convention-and-1967-protocol-relating-status-refugees

4 See https://www.frontex.europa.eu/what-we-do/monitoring-and-risk-analysis/migratory-map/

Conclusion

While there is still hardly any data available for demographic flows, available EU-level indicators for irregular flows are generally of good quality, particularly with respect to accessibility and documentation. But there are, unsurprisingly, some limitations in terms of validity and reliability. In terms of external validity, the data available often only describe an aspect of the phenomenon of irregular migration instead of being representative of the whole (e.g., asylum data only capture status-related flows). Among others, there are also issues with double-counting[5] or missing data, particularly when disaggregating by age and sex, which pose a challenge to measurement precision. As for internal validity, it is difficult to independently assess since the data are generated by bureaucracies with limited oversight; the indicators used are cross-sectional and not linked in any way; and there are not many opportunities to cross-validate the numbers with other information. Eurostat and EU Agencies work hard to harmonise data collection among member states, but currently, limitations continue to be evident, particularly with regard to double- or under-counting, geographical and temporal comparability (including time lags), and finally, interoperability across EU systems.

Good quality data are essential for effective migration governance. On the one hand, it can be in the best interests of irregular migrants to be counted, particularly if they need protection. However, the same data can also be used for the enforcement of migration legislation, including apprehension, detention, or deportation. As such, the interest in enhancing data collection on irregular migration and generating estimates must be carefully weighed against privacy considerations and societal interests. This balance is crucial so as not to impede trust on the part of irregular migrants and hinder the public service mission of providers or support groups, civil servants, and other street-level bureaucrats who regularly come into contact with them. In view of these, we recommend the following main ways forward to advance research on irregular migration flows and to prevent misuse of migration data:

- Define irregularity well and, when needed, be clear about different types of irregularity;
- Continue improving data quality for (selected) flow indicators, for example, by investing more resources into quality checks and making cohort data across multiple indicators available (without compromising privacy considerations);
- Acknowledge that supplemental qualitative information is essential for the validation and triangulation of quantitative data; incorporating qualitative studies[6] into the collection of migration data should be the norm; and finally,
- Consider using accessible informational resources, such as educational videos, to mitigate the misuse of migration data for political purposes; knowing the importance of a fact-based discourse can help ensure that statistics on migrant populations are not manipulated or misrepresented to serve political agendas.

The salience and problematisation of irregular migration in policy and everyday discourse increase the risks associated with the use of irregular migration data for political purposes. The potential for misuse[7] cannot be underestimated – from the presentation of statistics to the utilisation of such statistics in political decisions and policymaking. Immigration, particularly irregular migration, has become a divisive, even polarising topic. As such, all the more is good quality data – accurate, frequent and timely – of critical importance.

5 For example, if an individual attempts to cross the border multiple times within a short period, each attempt is likely recorded as a separate incident. There is also potential double counting between indicators as one person might generate a detection at one border, then an application for asylum, then a withdrawal, then another detection at another border, another asylum application, a Dublin hit, a negative asylum decision etc. All these data concerning one individual may be recorded within a year on Eurostat.

6 For example, conducting anonymous interviews with irregular migrants themselves and collecting testimonies that describe their situations and intentions in more detail rather than relying solely on a simple counting exercise.

7 See for example, ECRE. (2022). Asylum statistics and the need for protection in Europe: Updated Factsheet https://ecre.org/wp-content/uploads/2022/12/Asylum-statistics-and-the-need-for-protection-in-Europe-final.pdf. Also, Mouzourakis, M. (2014). 'Wrong number?' The Use and Misuse of Asylum Data in the European Union. https://www.ceps.eu/ceps-publications/wrong-number-use-and-misuse-asylum-data-european-union/

References

Ahmad-Yar, A. W., & Bircan, T. (2021). Anatomy of a misfit: International migration statistics. Sustainability, 13(7), 4032. https://doi.org/10.3390/su13074032

Bijak, J., Disney, G., Findlay, A. M., Forster, J. J., Smith, P. W., & Wiśniowski, A. (2019). Assessing time series models for forecasting international migration: Lessons from the United Kingdom. Journal of Forecasting, 38(5), 470-487. https://doi.org/10.1002/for.2576

Cantat, C., Pécoud, A., & Thiollet, H. (2023). Migration as Crisis. American Behavioral Scientist, 1-23. https://doi.org/10.1177/00027642231182889

CLANDESTINO Project (2009). CLANDESTINO Project Final Report. https://cordis.europa.eu/docs/publications/1266/126625701-6_en.pdf

De Genova, N. P. (2002). Migrant "illegality" and deportability in everyday life. Annual Review of Anthropology, 31(1), 419-447. https://doi.org/10.1146/annurev.anthro.31.040402.085432

ECRE (2022). Asylum statistics and the need for protection in Europe: Updated Factsheet. https://ecre.org/wp-content/uploads/2022/12/Asylum-statistics-and-the-need-for-protection-in-Europe-final.pdf

Eurostat. (n.d.). Enforcement of immigration legislation statistics introduced. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Enforcement_of_immigration_legislation_statistics_introduced

Jandl, M. (2008). Methods for estimating stocks and flows of irregular migrants. Chapter 3 of CLANDESTINO Report on methodological issues. https://migrant-integration.ec.europa.eu/library-document/clandestino-project-report-methodological-issues_en

Jandl, M., Vogel, D., & Iglicka, K. (2008). Report on methodological issues (Research Paper, CLANDESTINO Undocumented Migration: Counting the Uncountable, Issue. https://migrant-integration.ec.europa.eu/library-document/clandestino-project-report-methodological-issues_en

McAuliffe, M., & Ruhs, M. (2017). World Migration Report 2018. https://www.iom.int/sites/g/files/tmzbdl486/files/country/docs/china/r5_world_migration_report_2018_en.pdf

McAuliffe, M., & Sawyer, A. (2021). The roles and limitations of data science in understanding international migration flows and human mobility. In M. McAuliffe (Ed.), Research Handbook on International Migration and Digital Technology (pp. 42-57). Edward Elgar Publishing.

Mouzourakis, M. (2014). 'Wrong number?' The Use and Misuse of Asylum Data in the European Union. https://www.ceps.eu/ceps-publications/wrong-number-use-and-misuse-asylum-data-european-union/

Sassen, S. (1999). Guests and Aliens. The New Press.

Siruno, L., Leerkes, A., Hendow, M., & Brunovská, E. (2024). MIrreM Working Paper on Irregular Migration Flows. University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.10702228

Vogel, D., Jandl, M., Kraler, A., & Vogel, D. (2008). Report on methodological issues (Report prepared for the research project CLANDESTINO Undocumented Migration: Counting the Uncountable. Data and Trends across Europe funded under the 6th Framework Programme of the European Union, Issue. https://migrant-integration.ec.europa.eu/library-document/clandestino-project-report-methodological-issues_en

Yildiz, D., & Abel, G. (2021). Migration stocks and flows: data concepts, availability and comparability. In M. McAuliffe (Ed.), Research handbook on international migration and digital technology (pp. 29-41). Edward Elgar Publishing.

Box 8.2: Understanding asylum data in the context of irregular and regular migration

Teddy Wilkin and Petya Alexandrova

Data on asylum applications are used widely used as indicators of mixed migration to and within the EU+.[8] Yet interpreting these figures in relation to irregular and regular migration requires careful nuance. Many asylum seekers cross borders undetected, some enter legally, and others apply repeatedly in the same country or move between EU+ countries. This complexity creates challenges for measurement, interpretation and policy.

As of mid-2025, there were 1.3 million asylum applications in the EU+ still awaiting a final decision. This highlights the scale of people currently staying with unresolved legal status—many of whom may eventually find themselves in an irregular position if their claim is rejected. In 2024, EU+ countries issued around a third of a million negative asylum decisions. While some appeal such decisions, many abscond and remain without legal residence.

Visa policy provides a direct link between asylum and regular migration. In 2024, around a quarter of all asylum applications in the EU+ were lodged by persons originating from visa-exempt countries. Such persons can enter the EU for touristic reasons without needing to apply for visa. Many do so, and then claim asylum. Conversely, those from visa-obliged countries may apply for a visa and then arrive regularly and apply for asylum. The share of visa holders among asylum applicants is quite important in some EU+ countries.

However, irregular entry remains extremely important for asylum applications. EUAA estimates suggest that in 2024, detected illegal border-crossings by land and sea accounted for about 1 in 7 asylum applications overall, rising to a third of all asylum applications in frontline Member States. However, these only reflect actual detections at the border. Undetected irregular arrivals are, by definition, not counted—meaning any analysis based solely on detections risks underestimating the scale. This makes it even more important to triangulate asylum data with other sources.

Asylum applications can also reveal secondary movements—people applying sequentially in more than one EU+ country or applying in EU+ countries other than the one they initially entered. In 2024, nearly 150,000 decisions were issued in response to outgoing Dublin requests, which, we estimate, relates to about 14% of total applications. Such requests are made under the Dublin III Regulation which establishes which Member State is responsible for examining an asylum application. Most of these requests were for reasons related to secondary movements. Even persons with refugee status have been known to move and reapply elsewhere. Data from Eurodac, the EU's biometric database for asylum and irregular entry, provide additional insights. In 2023, there were more than 276,000 instances of asylum applications being linked to recent irregular border-crossings. Just over half applied for asylum in the same Member State where they were detected, while the rest applied for asylum in another Member State. These matches illustrate the link between irregular entry and asylum applications, but the Eurodac data have limitations including potential double counting, the exclusion of children under 14, and the lack of breakdowns by nationality.

Repeated asylum applications add another layer of complexity. According to eu-LISA,[9] only 55% of applications lodged in 2023 were first-time claims, indicating that nearly half of all applicants had already lodged previous asylum applications somewhere in the EU+. EUAA estimates suggest that nearly a tenth were individuals reapplying in the same EU+ country (in both 2023 and 2024), often after remaining in the country for an extended period—typically in an irregular or tolerated status.

8 EU+ = EU Member States plus Norway and Switzerland

9 E-LISA stands for the European Union Agency for the Operational Management of Large-Scale IT.

Asylum data can also reflect demographic patterns. Some applicants are children born in the EU+ to an asylum-seeking parent, in some cases making up more than 10% of all applicants. These figures reflect how status can persist intergenerationally without clear legal resolution.

Looking ahead, under the Interoperability Regulation, the Central Repository for Reporting and Statistics (CRRS), currently under development, is expected to deliver cross-system statistics that will significantly improve our understanding of these dynamics. It will enable anonymous tracking across databases and provide more precise insights into how individuals move through stages of irregular entry, legal stay, asylum and status withdrawal.

In short, asylum data provide a valuable but incomplete window into migration stocks. They reflect both regular and irregular situations but must be interpreted with care. Analysts should consider visa status, secondary movements, repeated applications and related demographic patterns. When triangulated with detections at the border, visa records, and Dublin statistics, asylum data help clarify not only the scale of irregular presence, but also how individuals engage with EU+ migration and protection frameworks.

Box 8.3: Understanding 4Mi data

Francesco Teo Ficcarello

What is 4Mi?

4Mi, developed by the Mixed Migration Centre (MMC), is an innovative and global data collection platform[10] that provides independent and in-depth insights into the experiences of migrants moving along mixed migration routes. Since 2014, 4Mi has become the world's largest globally comparable primary data collection system focused specifically on people on the move, with more than 130,000 interviews conducted in over 30 countries across Africa, Asia, Europe and Latin America.

Rationale and scope

4Mi was created to fill a major evidence gap around the realities faced by migrants and refugees in transit—populations often invisible in traditional migration statistics due to their irregular or non-camp-based status. The initiative contributes directly to the Global Compact for Migration's Objective 1 by providing accurate and disaggregated data for informed policy-making.

Methodology

Data are collected through one-on-one, structured interviews using standardized yet adaptable survey tools, administered by a large network community-based enumerators (approx. 130 as of 2025), most of whom are migrants or refugees themselves. These enumerators are embedded in local contexts, ensuring high trust and access to hard-to-reach groups. In countries where MMC is not established, data are collected through local partners rooted in the countries. Surveys are quantitative, enabling statistically robust analysis, but also include open-ended questions to capture personal narratives. Sampling is purposive, with enumerators operating in migration hubs identified through scoping and mapping exercises. While not statistically representative, the data are highly indicative, enabling rich, contextual understanding of profiles, drivers, journeys,, vulnerabilities, and aspirations of people on the move.

10 See https://mixedmigration.org/4mi/4mi-faq/

Tools and flexibility

The model is flexible, allowing for add-ons on topics such as youth migration, climate mobility, and urban integration. Innovations include longitudinal follow-ups, remote data collection, and interactive dashboards for public data exploration. This flexibility was key to rapidly launching Covid-19-specific modules, through which 25,500 interviews were conducted in 2020 alone.

Data use and outputs

4Mi data feed into MMC's research publications,[11] interactive dashboards[12] and presentations towards evidence-based programming and policy-making. The data are also shared with partners such as UN agencies and NGOs under data-sharing agreements. Outputs include statistical analyses in the form of research reports, briefing papers, snapshots, infographics and policy briefings, as well as real-time response tools for humanitarian actors.

A unique complement to flow data

By providing in-depth, qualitative insights into the human dimension of migration, 4Mi complements other data collection and flow monitoring systems (e.g., IOM's DTM), which focus more on volumes. 4Mi captures lived experiences, decisions, and risks in a globally comparable format, enabling cross-regional and route-based analysis. Its integration of quantitative scale with qualitative depth ensures that the perspectives of (irregular) migrants—often missing from mainstream migration discourse—are not only heard but systematically analyzed. In doing so, 4Mi plays a vital role in providing an evidence base for the development of more humane, inclusive, and responsive migration policy and practice worldwide.[13]

11 See https://mixedmigration.org/resources/

12 See https://mixedmigration.org/4mi/4mi-interactive/

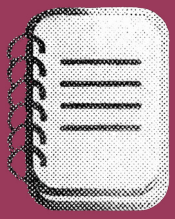
13 More information on 4Mi can be found at https://mixedmigration.org/wp-content/uploads/2021/08/4Mi-Introduction.pdf

Chapter 9

Irregular migration and informal work

Aslı Salihoğlu and Carlos Vargas-Silva

Irregular migration and informal work



Key points

- Public debate often conflates irregular migration with informal work. Varying definitions of both make it difficult to compare research or draw clear policy conclusions.
- Both irregular migration and informal work are hard to measure. This makes it especially difficult to accurately assess how they intersect.
- This chapter presents a straightforward, flexible and scalable framework for estimating this overlap. Using data sources commonly available in many countries, it can generate plausible ranges for the number of irregular migrants working in the informal economy.

Introduction: Politically important, conceptually confused, and empirically challenging

The media has placed a lot of attention on the role irregular migrants in the informal economy. Headlines suggests that "Migrants scrape by in underground economy", "Migrants will keep coming as long as we offer illegal jobs" and "We don't need French lessons on the black economy and illegal workers". Yet, there is a lack of general understanding of how concepts such as the "underground", "black" or "shadow" economy relate to migration.

The systematic study of irregular migrants in the informal economy is beset by conceptual ambiguities and methodological challenges. Definitions of economic informality and migrant irregularity remain indefinite as scholars and policy professionals continue to reformulate them iteratively. Any synthesis and cohesive knowledge accumulation are hampered by the competing definitions used in this large body of research. Furthermore, data unavailability and unreliability hinder efforts to estimate the size of the informal economy and irregular migrant stocks across different contexts and time periods in a consistent fashion. Since both phenomena are hard to capture statistically, the accurate assessment of their intersection is doubly challenging. Overall, confusion regarding the concepts themselves and obstacles to measurement impede research and policymaking regarding the participation of irregular migrants in the informal economy.

Who are the irregular migrant workers in the informal economy?

Our framework includes two types of firms, registered and unregistered, where foreign nationals (with and without right of residence), denizens and nationals of the country can take up employment. The informal economy is composed of those a) working in unregistered firms, and b) working in registered firms, but not abiding by all regulations (e.g., getting paid "off the books").

For nationals and denizens, the formal versus informal economy placement depends on individuals' workplace registration and conditions of employment. For others, their residence status plays a role in placing them in the formal or informal economy. Foreign nationals without a right of residence are categorized as irregular migrants regardless of their employment status apart from asylum seekers who have obtained a work permit. Foreign nationals with terminable right of residence are categorized as regular non-denizen migrants if they are active in the formal economy or not working. Conversely, their participation in the informal economy is an indicator of migrant irregularity.

See Figure 9.1 for the main framework of analysis and for a more detailed breakdown of the concepts under discussion, please refer to Salihoğlu and Vargas-Silva (2024).

<table>
<tr><th></th><th colspan="2">Registered firms</th><th rowspan="2">Unregistered firms</th></tr>
<tr><th></th><th>Conditions met</th><th>Conditions not met</th></tr>
<tr><td>Nationals and denizens</td><td>National and denizen workers in the formal economy</td><td colspan="2">National and denizen workers in the informal economy</td></tr>
<tr><td>Foreign nationals with terminable right of residence</td><td>Regular migrant workers in the formal economy</td><td>Foreign nationals with terminable right of residence working in registered firms, but violating their conditions of stay otherwise</td><td>Foreign nationals with terminable right of residence working in unregistered firms</td></tr>
<tr><td>Foreign nationals without right of residence</td><td>Not applicable</td><td>Foreign nationals without right of residence working in registered firms</td><td>Foreign nationals without right of residence working in unregistered firms</td></tr>
</table>

Irregular migrant workers in the informal economy

Figure 9.1: Framework for analysis

How to estimate the number of irregular migrant workers in the informal economy?

To measure the participation of irregular migrants in the informal economy, the intersection of two estimates, namely that of irregular migrants and the informal economy, should be calculated.

We present an approach to estimate this intersection. The method yields estimate ranges, each fitted with a minimum and a maximum figure generated through a procedure subject to context-specific conditionalities. It involves the following 4 steps:

1. Start with a full dataset from a survey representative of a labour market. Drop those who are not in employment.
2. Drop all employed nationals and denizens. Use different definitions of denizenship to set several thresholds of estimation that decrease in their degree of conservativeness and gradually constrain the number of observations categorized as non-denizen migrants in the dataset.
3. Generate a maximum and a minimum estimate of the number of irregular migrant workers. This relies on using survey variables that proxy economic in/formality in standalone or combinatory fashion. A separate minimum-maximum estimate range is generated per denizenship threshold as defined in Step (2).
4. Lastly, the observations that remain under the maximum and minimum specifications per denizenship threshold are multiplied with their corresponding survey weights to generate estimate ranges for irregular migrant worker populations.

For further details of the estimation see Salihoğlu and Vargas-Silva (2025).

What are the limitations of this approach?

Our approach is straightforward, flexible and scalable by design. It can be applied to most surveys with supply-side information on the labour force, including labour force surveys, censuses and living conditions surveys, all of which are traditional and relatively standardized data sources available in many countries.

Yet, the performance of our method is constrained by the underlying microdata that it draws on. Survey samples may suffer from self-selection bias. Irregular migrants tend to have lower participation rates in surveys than regular migrants due to concerns over visibility to state authorities. They may also have accommodation arrangements that leave them out of the sampling frame altogether, such as newly arrived working tourists who stay in hotels. These factors hamper our method's ability to provide a full accounting of the scope and distribution of irregular migrant worker profiles identifiable in survey data.

Implications

Unsettled academic and policy debates regarding the definition of the informal economy have hindered clearheaded analysis and policymaking vis-à-vis the participation of irregular migrants therein. Our approach has pinned down a working definition for the informal economy in order to identify, characterize, and quantitatively measure this phenomenon.

This method is intended as a starting point for researchers to adapt our approach to their national setting. The method offers a framework for quantifying the participation of migrants in host labour markets in a holistic manner. Even in countries where labour and/or firm informality levels are low, migrants are likely to be overrepresented in the informal labour force and their economic contributions consequently not acknowledged in national statistics.

References

Salihoğlu, A., & Vargas-Silva, C. (2024). Irregular Migrants in the Informal Economy: Conceptual Issues and its Measurement. MIrreM Working Paper 5. Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/ZENODO.10569402

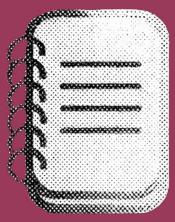
Salihoğlu, A., & Vargas-Silva, C. (2025). Measuring the Participation of Irregular Migrants in the Informal Economy. MIrreM Briefing Paper. Krems: University for Continuing Education Krems (Danube University Krems). https://doi.org/10.5281/zenodo.14809000

Chapter 10

Surveying irregular migrants: Challenges and approaches

Rocco Molinari and Livia Elisa Ortensi

Surveying irregular migrants: Challenges and approaches



Key points

- Irregular migrants are difficult to capture in statistics because of their absence from official sampling frames, mobility, and fear of detection. Surveying them requires tailored approaches, including non-probability sampling, trust-building strategies and ethical safeguards.
- This chapter reviews three types of surveys that can yield data on irregular migrants: those that explicitly include them in the sampling design, those that target applicants of regularisation programmes, and retrospective surveys that reconstruct past legal trajectories.
- Drawing on examples from France, Italy, Spain and the United States, this chapter shows how innovative designs and context-specific adaptations can improve coverage and data quality.
- Each approach has its own strengths and limitations. A combination of methods, applied thoughtfully, is needed to strengthen the evidence base and support more accurate data collection and analysis.

Introduction

Understanding irregular migration processes is crucial in contexts where legal barriers to long-term immigration are prominent, such as Western migrant-receiving countries. Policymakers not only need techniques to estimate irregular migration flows and stocks, but also data on the lived experiences of undocumented migrants. This includes how legal status interacts with various dimensions of settlement (e.g., health, labour market, family formation, crime, attitudes). Surveying undocumented migrants is one way to investigate these issues. However, while traditional migration surveys are already challenging (Vickstrom and Beauchemin, 2024), these challenges are amplified when the target population lacks legal status, due to structural, methodological, and ethical issues that distinguish this population from most others.

A fundamental difficulty is that irregular migrants are not generally included in official population registers or sampling frames, leading

to identification challenges for researchers. Without a known universe from which to draw a representative sample, it is not possible to apply standard probability sampling methods. Moreover, the lives of irregular migrants tend to be embedded in informal networks and practices. Mistrust can be a pervasive issue: irregular migrants often avoid contact with entities perceived as linked to official institutions due to fear of detection, detention, or deportation. This leads to high levels of non-response and answers shaped by mistrust, especially if anonymity is not fully guaranteed. Building trust requires time, cultural sensitivity, and in many cases, collaboration with community-based actors or mediators.

Even when undocumented migrants are—whether by design or by chance—included in a survey sample, legal status is rarely collected, and if it is, the data are often unreliable. High mobility and precarious living conditions further complicate data collection. Frequent changes in housing and employment, geographic mobility, and periods of complete inaccessibility due to informal work patterns make it extremely difficult to trace respondents over time, particularly in longitudinal studies (Peitz et al., 2024). Finally, undocumented migrants are likely to differ from the other migrants on the basis of observable and unobservable characteristics. This selectivity can affect the representativeness of any resulting sample.

Altogether, these factors combine to make irregular migrants one of the most difficult populations to study using conventional social science methods. Accurately capturing their living conditions requires not only adapted methodological tools, but also a deep ethical commitment to protection, confidentiality, and respectful engagement. For all these reasons, surveys that include undocumented migrants are generally scarce, small, locally based, and targeted to specific migrant subgroups (Bachmeier et al., 2014). However, a limited number of studies have succeeded in targeting undocumented migrants or including them within broader samples of migrant populations. This chapter examines the most commonly used approaches to surveying undocumented migrants and reviews promising practices. Although most of the research has traditionally been conducted in the United States, the chapter places greater emphasis on Europe, where several innovative approaches have recently emerged.

What types of irregular migration surveys are there?

Surveys that include information on the life conditions of current or former undocumented migrants can be broadly grouped into three main categories, based on their methodological approach and target population:

1. The first category comprises surveys that explicitly include undocumented migrants in their sampling design. These are the only surveys that can be used to understand the life conditions of current irregular migrants. They typically compare irregular with regular migrants. These surveys use specific data collection techniques—such as centre-based sampling or other network-based methods—designed also to reach undocumented individuals, or they rely on existing sources that indirectly capture segments of the undocumented population without targeting them explicitly.
2. A second category consists of surveys conducted in the context of regularisation programmes. These surveys focus on people applying for legal status and often gather information on their legal trajectories and socio-economic conditions. Some include a longitudinal component, following applicants over time to assess the impact of regularisation on their lives.
3. A third type includes retrospective surveys conducted with migrants who currently hold a legal status, but which collect data on their past experiences of irregularity, thereby reconstructing their legal trajectory and capturing temporary phases of undocumented residence. These surveys can be used to understand the situation of migrants who have recently regularised and to understand the medium and long-term consequences of irregularity among regularised migrants.

Surveys that explicitly include undocumented migrants in their sampling design

Some surveys designed to collect information on undocumented migrants avoid the use of a conventional sampling frame altogether. A leading example in Europe is the Regional Observatory for Integration and Multiethnicity (ORIM) in Lombardy, Italy. Active from 2001 to 2021, the program collected data on the living conditions of people with a migration background. Explicit efforts were made to include irregular migrants, who – particularly in ORIM's early years – made up a substantial portion of the foreign-origin population in the area.

Every year, ORIM conducted retrospective, face-to-face interviews with a representative sample of foreign residents in the region using the Centre Sampling Technique (CST; see Box 10.1; Baio et al., 2011). A cornerstone of the ORIM model was its participatory and inclusive approach to fieldwork: interviews were conducted by trained cultural-linguistic mediators of migrant background, enhancing trust and communication, which was particularly important when engaging with undocumented individuals.

Over the course of two decades, ORIM generated a unique cross-sectional data series that supported academic research and informed evidence-based policies in integration, social inclusion, and rights protection. Although the program was discontinued in 2021, it has remained a methodological benchmark for research on hard-to-reach populations and a model for how undocumented migrants can be ethically and effectively surveyed. CST has also been used at the national level in Italy and outside the Italian context (e.g. the Immigrant Citizenship Survey ICS).

Box 10.1: The Centre Sampling Technique

Rocco Molinari and Livia Elisa Ortensi

The **Centre Sampling Technique (CST)** is a probabilistic sampling method developed to reach hard-to-survey populations, particularly undocumented migrants who are typically excluded from standard household surveys due to the lack of a sampling frame. The method was first implemented systematically in Italy. CST is based on the idea that migrants—regardless of their legal status—tend to frequent specific **centres** or **aggregation points** in their everyday lives, such as religious institutions, cultural and community associations, consulates, NGOs, migrant help desks, public spaces, and informal meeting places. The method proceeds in three stages.

First, a **mapping phase** is conducted to identify and classify existing centres that are expected to be regularly visited by the target population within the geographic area of interest. Centres are categorised by type (e.g., religious, cultural, associative, consular), estimated relevance (e.g., estimated average attendance) and population specificity (e.g., open to all migrants or nationality-specific), and then stratified accordingly. Then, **a sample of centres** is drawn, and some individuals are selected in each centre either randomly (e.g., systematic sampling upon entry) or via controlled quota sampling if the flow is not randomizable. The unit of analysis is the individual migrant.

After the end of the interview phase, weights are calculated based on the number of centres attended and their importance, which allows for correcting potential overrepresentation of more socially active individuals.

Some surveys have successfully reached undocumented migrants by exploiting administrative sources that, by their nature, include them. One prominent example in Europe is the **Spanish National Immigrant Survey** (**ENI**; Reher and Requena, 2009), carried out by Spain's National Statistics Institute (INE) in 2006–07. The ENI drew its sample from the municipal population register (*Padrón Municipal*), which grants all registered residents—including irregular migrants—access to public health care and other services and is considered representative of immigrants living in

Spain irrespective of their legal status. It collected information on the type of respondents' residence permit and immigration status (e.g., asylum applicant).

Similarly, Germany's **IAB-BAMF-SOEP**[1] (see Box 10.2) and the Feasibility Study on the Im-/Mobility of Rejected Asylum Seekers (**MIMAP**; Stache et al., 2024) include groups such as rejected asylum seekers with temporary suspension of removal ('Duldung'), capturing segments of the population who experience forms of de facto irregularity. The MIMAP Survey, in particular, was explicitly designed to target irregular migrants through its sampling strategy and questionnaire items.

Box 10.2: Surveying irregular migrants with an existing sampling frame – The IAB-BAMF-SOEP survey of refugees

Randy Stache

As in any survey, a suitable sampling frame that includes the entire target population and enables sample selection as well as contact details is crucial for reliable survey data collection on irregular migrants and for generalizing empirical results. In Germany, the Central Register of Foreigners (see Chapter 7) offers such a sampling frame for subgroups of irregular migrants, enabling representative samples and the use of traditional survey methods. Since 2016, **the IAB-BAMF-SOEP Survey of Refugees** is annually surveying refugees who arrived in Germany since 2013 in a panel study, regardless of the outcome of their asylum procedures. As a result, the data include irregular migrants known to the authorities whose deportation has been temporarily suspended (tolerated/*Duldung*).

The dataset offers several advantages to analyse the living situation of irregular migrants: 1) **Accessibility** to external researchers via a data usage agreement. 2) **Broad thematic coverage**, including migration trajectories, housing, employment, language acquisition, health, attitudes, religion. 3) **Longitudinal design**, allowing for the observation of individual developments over time. 4) A **heterogeneous group** of irregular migrants in terms of age, gender country of origin, and other characteristics. 5) **Comparative potential**, enabling systematic analyses of differences between individuals with tolerated status and other groups (recognized refugees or migrants and natives - when using the compatible SOEP-CORE and IAB-SOEP MIG data), and the identification of influencing factors across domains.

However, when using the data for research on irregular migrants some limitations arise: 1) The dataset **includes only a specific subgroup** of irregular migrants – those with tolerated status following an asylum application. Additionally, this group tends to participate less often in follow-up surveys and had higher non-response. 2) Additionally, **not all topics are covered in every survey wave**. 3) As a result, **representativeness and reliable estimations may be limited** for certain research questions. However, statistical techniques such as weighting, pooling of waves, or propensity score matching can help mitigate vthese issues. 4) There is **inherent selectivity**: irregular migrants who have returned, moved to another country, or gone into hiding are not captured in the data. 5) Some **questions central to the lived experiences of irregular migrants** – such as work permits, life in irregularity, coping with the threat of deportation, or expectations regarding their country of origin – are either absent or not asked in a way that avoids possible bias, like social desirability.

1 This survey is undertaken by the Research Centre of the Federal Office for Migration and Refugees (BAMF-FZ) in cooperation with the Institute for Employment Research (IAB) and the Socio-Economic Panel (SOEP) at German Institute for Economic Research (DIW Berlin). Further information can be found at https://www.diw.de/en/diw_01.c.930532.en/iab-bamf-soep_survey_of_refugees.html

The **Brief Analysis 3/2024** published by the Research Centre of the Federal Office for Migration and Refugees illustrates how this data can be used to **study the living conditions of tolerated persons** in comparison to recognized refugees, using propensity score matching. The comparison shows that both groups are similarly integrated in terms of language skills and employment. However, the tolerated are more likely to live in shared accommodations and report much lower life satisfaction, which further declines over time (Stache, 2024).

References:

Stache, R. (2024). Auswirkungen einer Duldung auf Lebenssituation und Lebenszufriedenheit. (BAMF-Kurzanalyse, 3-2024). Nürnberg: BAMF.
https://doi.org/10.48570/bamf.fz.ka.03/2024.d.2024.duldung.1.0

In the US, nationally representative surveys have been used to identify 'likely undocumented' immigrants through imputation. For example, using the **Survey on Income and Program Participation (SIPP)**, a longitudinal study investigating occupational-related aspects in the US, some scholars exploited limited information on visa status (concerning citizenship and legal permanent resident (LPR) status) and participation in welfare programs to infer immigrant respondents' current legal status (Hall et al., 2010). Other studies have developed imputation methods based on observable characteristics unrelated to legal status, which have been applied to the **Current Population Survey (CPS)**, the American labour force survey (Passel and Cohen, 2014).

Surveys targeted to applicants of regularisation programmes

Surveys targeting applicants of regularisation programmes are a key source of empirical evidence on migrants who have experienced irregularity. However, they only capture information on those who successfully applied, and therefore exclude non-applicants or rejected cases. These surveys are typically conducted in the process of major legalisation programmes and are designed to capture individuals' socioeconomic characteristics, labour market trajectories, and integration patterns.

One of the most prominent examples is the **Legalized Population Survey (LPS)**, a longitudinal survey launched in the US after the 1986 Immigration Reform and Control Act (IRCA), which granted legal status to nearly 2.7 million undocumented migrants. Conducted in two waves, the LPS collected detailed data on pre- and post-legalisation employment, mobility, income, and legal trajectories, and remains a foundational source for studying the economic impacts of legalisation. The first wave of the survey (LPS1) gathered data from 6,193 individuals who had applied for temporary residence status by January 31, 1989. Respondents were asked to report their employment status during the week preceding the submission of their amnesty application. In the second wave (LPS2), conducted in 1992, a follow-up was carried out with 4,012 participants from LPS1 who had since obtained lawful permanent residence. While the sample is not representative of all individuals who received amnesty under IRCA, the longitudinal design remains a major strength for analysing changes in employment outcomes over time, specifically around the critical transition from undocumented to legal status.

Another smaller scale example is the **Parchemins Study**, a prospective, mixed-methods panel survey conducted alongside Operation Papyrus, the 2017–2018 regularisation scheme for undocumented economic migrants in the Swiss canton of Geneva. It tracked approximately 400 individuals up to 3 years after regularisation, focussing on the effects of regularisation on their health and well-being (Lives Centre, 2020).

Retrospective surveys on migrants who currently hold a legal status collecting data on their past experiences of irregularity

A third type of survey focuses on the past irregular experiences of migrants who now hold legal status. By working with immigrants holding legal status, these surveys simplify sampling design, but rely on respondents' recall and willingness to disclose prior undocumented residence through direct questions (e.g., 'Have you ever been irregular?') and collecting information on how their legal status changed over time (e.g., the types and timings of residence permits).

Examples include the **Social Condition and Integration of Foreign Citizens (SCIF)** survey, conducted by the Italian National Statistical Office (Istat) in 2011-12, and **Trajectories and Origins 2 (TeO2)**, carried out by the French Institute for Demographic Studies (INED) and the National Institute of Statistics and Economic Studies (INSEE) in 2019-20 (see Box 10.3). One of the main limitations of these studies lies in their exclusive focus on the initial phase of irregularity (i.e., between arrival in the destination country and the acquisition of a first permit) without reconstructing respondents' full legal status trajectory. To address this limitation, one could extend the time frame by combining retrospective questions about past legal status with longitudinal or prospective data that track respondents over time.

Box 10.3: Reliability in measuring migrants' legal trajectories and experiences of irregularity in a retrospective survey: The case of "Trajectories and Origins 2"

Julia Descamps

In a retrospective survey, how much can we rely on the data collected on legal status and past episodes of irregularity? Drawing on the example of the French Trajectories and Origins survey (Ined, INSEE, 2019-2020), the potential biases were considered (Descamps, 2024). Two of these are particularly challenging in the context of surveying irregular migration. Memory bias, which occurs when the content of a response depends on the ability to recall information, could affect migrants with insecure and bumpy legal trajectory. Social desirability bias, a tendency to present oneself in a favorable light to others, might be more prevalent among migrants who have experienced irregularity, an experience on the legal margins, therefore particularly sensitive. Those biases are tested using TeO2 survey, by examining the non-response rates, and quantifying the under-reporting of irregularity, on a sample of 7,057 immigrants arrived to France after the age of 18.

Non-response to the question "Have you ever been irregular?" is low (1%), and does not increase with the length of time since arrival, unlike the non-response rate on the first legal permit in France. Regarding irregularity, memory bias appears to be minimal: respondents found it more difficult to recall events from the early stages of their legal journey, but were less hesitant when it came to irregularity.

The length of time that respondents declare they spent as irregular migrants is then compared with a proxy for irregular status on entry: the time it took them to obtain their first residence permit (from the year they entered France to the year they obtained their first residence permit). Positive differences between the two figures (reported time with undocumented status inferior to time before first residence permit obtained) are taken as evidence of under-reporting of periods of irregular status by respondents. Taking only those respondents with a gap between accessing France and obtaining their first permit – who could therefore underreport this situation – 70% of cases match within one year. The proportion of under-reported irregularity is 27%. This rate is an estimate of the social desirability bias. This bias appears to be more prevalent among educated migrants. The feeling of downward social mobility associated with irregular status, stronger when the social status in the home country is high, can lead respondents to regain control over their migratory narrative. The same is true of asylum applicants who were denied refugee status: they also tend to under-report irregularity. Their experience of administrative domination could lead them to modify their account of their irregular status. Social desirability bias could also overlap with memory bias, with partial answers being due to the often precarious and rocky migration trajectories of asylum seekers.

These results highlight the importance of statistically surveying migrants about their various legal statuses and experiences of irregularity. Particular attention should be paid to the effects of categorisation and the leeway it provides.

References:

Descamps, J. (2024). Can We See Their ID? Measuring Immigrants' Legal Trajectory: Lessons From a French Survey. International Migration Review, 0(0). https://doi.org/10.1177/01979183241295995

Another notable example is the ELIPA 2 French panel, conducted by the Ministère de l'Intérieur et des Outre-Mer in three waves (2019, 2020, and 2022) with a representative sample of immigrants who obtained their first residence permit in France in 2018. In addition to other topics, the survey collected both retrospective and ongoing information on the administrative process of respondents, allowing researchers to reconstruct their legal status trajectories over a four-year period.

A common limitation of these surveys is that they only include immigrants who have obtained legal status at some point, thereby excluding those who remain undocumented. However, retrospective surveys also offer several advantages. First, instead of treating legal status as a fixed condition, they make it possible to investigate specific phases of irregularity, which is particularly valuable in contexts characterised by recurrent regularisations. Second, by relying on large samples and rich questionnaires, they enable long-term analyses of the consequences of irregular status over multiple time periods and dimensions of migrants' lives.

11 See https://mixedmigration.org/resources/

12 See https://mixedmigration.org/4mi/4mi-interactive/

13 More information on 4Mi can be found at https://mixedmigration.org/wp-content/uploads/2021/08/4Mi-Introduction.pdf

Conclusion

Efforts to survey irregular migrants will always face trade-offs between coverage, data quality, and ethical safeguards. No single method can fully overcome the challenges of sampling, trust, and mobility, so mixed approaches tailored to specific contexts are essential. Well-designed surveys can generate robust evidence to inform more balanced debates and better-targeted policies, but only if they are grounded in careful methodological choices and genuine engagement with the communities concerned.

References

Bachmeier, J.D., Van Hook, J. & Bean, F.D. (2014) Can we measure Immigrants' legal status? Lessons from two U.S. surveys. International Migration Review, 48(2), 538–566

Baio, G., Blangiardo, G. C., & Blangiardo, M. (2011). Centre Sampling Technique in foreign migration surveys: A methodological note. Journal of Official Statistics, 27(3), 451–465.

Hall, M., Greenman, E. & Farkas, G. (2010) Legal status and wage disparities for Mexican immigrants. Social Forces, 89(2), 491–513.

Lives Centre (2020). "Parchemins" Project - Assessing the health and well-being of undocumented migrants. Online Resource https://www.centre-lives.ch/en/project/parchemins-project-assessing-health-and-well-being-undocumented-migrants?chapter=195-project-description-parchemins

Passel, J.S. & Cohn, D. (2014) Unauthorised immigrant totals rise in 7 states, fall in 14: decline in those from Mexico fuels Most states decreases. Washington, DC: Pew Research Centre's Hispanic Trends Project. November.

Peitz, L. and Stache, R. & Johnson, L. (2024) How to Survey Hard-to-Reach Populations: A Practical Guide to App-Based Respondent-Driven Sampling. Robert Schuman Centre for Advanced Studies Research Paper No. 2024/23, http://dx.doi.org/10.2139/ssrn.4901241

Reher, D., & Requena, M. (2009). The National Immigrant Survey of Spain: A New Data Source for Migration Studies in Europe. Demographic Research, 20(12): 253–278. https://doi.org/10.4054/DemRes.2009.20.12

Stache, R., Peitz, L., & Johnson, L. (2024). The MIMAP Survey on Im-/Mobility Aspirations of Rejected Asylum Seekers: Survey Instruments & Codebook. Nürnberg: Bundesamt für Migration und Flüchtlinge. https://doi.org/10.48570/bamf.fz.fragebogen.mimap.en.2024.1.0

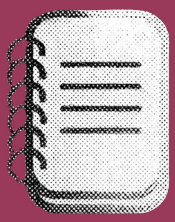
Vickstrom, E. & Beauchemin, C. (2024). Quantitative surveys on migration. In Sciortino, G., Cvajner, M., & Kivisto, P. (Eds.) Research Handbook on the Sociology of Migration, Edward-Elgar, pp 227-242.

Chapter 11

Towards the more effective use of irregular migration data

Adèle Appriou, Jasmijn Slootjes and Ravenna Sohst

Towards the more effective use of irregular migration data



Key points

- Irregular migration data is used to inform decisions across sectors—from service planning and public debate to policy design—but remains poorly aligned with the needs of those expected to use it. This chapter examines how different stakeholders rely on the data, and where gaps emerge.
- We introduce the data pathway to show how data on irregular migration are defined, collected, shared, and interpreted. At each stage, competing mandates, vague concepts, and inconsistent practices risk undermining coherence and usability.
- Five core challenges, such as gaps in availability, ambiguous definitions and low data literacy, limit the value of the data. Legal and institutional barriers, particularly a lack of knowledge about the implication of the GDPR, further constrain responsible sharing and can erode trust.
- Improving data use requires more than technical or methodological improvements. It calls for clearer definitions, accessible documentation, stronger privacy protections, and sustained investment in the capacity of those who collect and use the data.

How is data on irregular migration used by stakeholders?

As more data on irregular migration becomes available across different policy areas, there is a growing interest in stakeholders' data needs and use. Far from being a neutral resource, such data shapes decisions, drives public narratives, and influences outcomes across a wide range of domains, including:

- **Decision-making**: Helping policymakers weigh trade-offs and design effective responses.
- **Strategic and operational planning**: Enabling authorities and civil society to anticipate service needs and better predict migration flows
- **Identifying needs and policy gaps**: Informing policy agendas by highlighting the needs of irregular migrants.

- **Good governance and policy learning**: Enabling policy evaluation, accountability, and transparency.
- **Policy outcomes**: Shaping both individual lives and broader systems through resource allocation and programme design.
- **Fostering dialogue and innovation**: Informing public debate and driving new approaches.
- **Research**: Serving as the basis for analysis, theory testing, and evidence-building.

This chapter examines the main challenges surrounding the collection, interpretation, and use of irregular migration data. It also highlights promising practices and emerging solutions aimed at improving data use and supporting more informed policy responses.

How data on irregular migrants and migration are shaped along the data pathway

A variety of stakeholders—each with their own objectives and priorities—shape data along the 'data pathway'. Before data can be collected, key concepts such as 'irregular migration' must be defined, after which the data collectors decide how to measure this quantitatively. Data are then collected, shared, accessed, interpreted, and disseminated (see Figure 11.1). Notably, this data pathway is not always a linear process; steps may be skipped or repeated.

Throughout this process, barriers can emerge that stand in the way of the effective collection and use of irregular migration data. Obstacles that arise earlier on, for example unclear or inconsistent definitions of irregular migration or issues related to data sharing and access, create problems down the line for actors using the data.

Irregular migration data is shaped by the mandates, interests, and priorities of the varying stakeholders at each step along the data pathway.

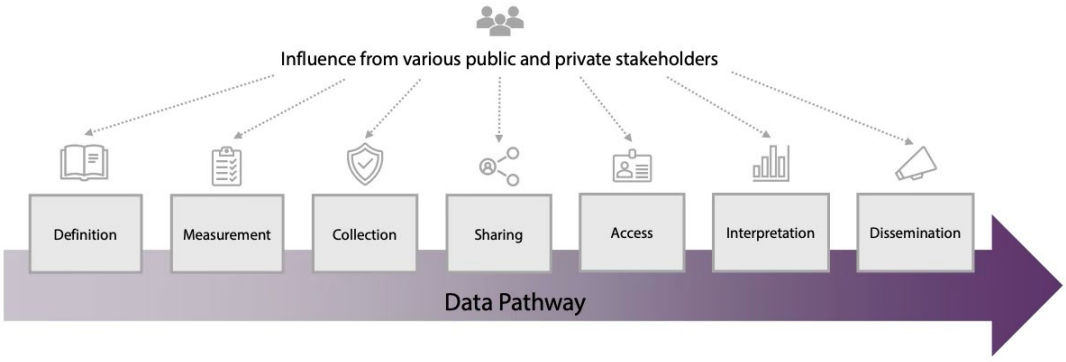


Figure 11.1: The data pathway

There are five challenges that limit the effective use of irregular migration data.

Challenge 1: Gaps in data availability

Selective and incomplete data are one of the main challenges for stakeholders seeking to use information on irregular migration. While data are predominantly available on rejected asylum claims and irregular border crossings, it lacks on other routes into irregularity (e.g., visa overstaying and being born into irregularity), secondary movements, sociodemographic characteristics, and longitudinal trends.

This is caused by:

1. A mismatch in priorities between those who collect data and those who use it. Most data on irregular migration are gathered as a byproduct of operational activities (e.g., border controls), and not to improve the broader evidence-base.
2. Practical obstacles, including irregular migrants underreporting crimes and abuse due to fear of deportation, and service providers or NGOs not registering migration status to prevent data-abuse and a chilling effect on migrants with irregular status seeking support.

Challenge 2: Data quality – ambiguous definitions and limited transparency

Even when data exists, its use is often hampered by unclear definitions and limited documentation, making it difficult to interpret or compare effectively.

- **Inconsistent definitions**: There is no universally accepted definition of irregular migration, and related terms like "undocumented" or "overstayer" are used inconsistently across contexts, reducing comparability and clarity.
- **Lack of transparency**: Critical details about how data is collected, what it measures, or how it should be interpreted are often missing or buried in annexes, preventing data users to assess quality or limitations. Inconsistent methods, outdated figures, and missing contextual details can further reduce the significance of existing datasets.

Challenge 3: Lack of data literacy

Even when data is available, it is not always used effectively. Limited data literacy among policymakers and a lack of shared language with data producers can lead to misinterpretation or mistrust. As a result, relevant data may be ignored or dismissed, reducing its potential to inform sound decision-making.

Challenge 4: Limited access – legal, technical and institutional barriers

Access to irregular migration data is often hindered by unclear legal regulations, leading some stakeholders to over- and misinterpret the guidelines set out in frameworks such as the EU General Data Protection Regulation (GDPR).

Box 11.1: GDPR and the limits of data access

Adèle Appriou, Jasmijn Slootjes and Ravenna Sohst

Legal uncertainty surrounding the GDPR — the main legal framework governing the collection, use, and protection of personal data in the EU — often creates barriers to the sharing and use of irregular migration data. While the GDPR is essential for safeguarding individuals' rights, it is not always clear how its provisions apply in specific contexts. In the absence of legal clarity, many authorities and organisations adopt a precautionary approach, limiting or avoiding data sharing altogether to reduce the perceived risk of noncompliance.

At the same time, insufficient adherence to data protection principles can result in rights violations. For instance, research by the EU Agency for Fundamental Rights has found that personal data collected during labour inspections are frequently shared with immigration enforcement authorities, despite GDPR requirements that such data should be used transparently and only for clearly defined purposes. This contravenes the principle of purpose limitation, which

prohibits the use of personal data for objectives beyond those originally specified. When these boundaries are not respected, it can erode trust, deter migrants from seeking protection or services, and ultimately undermine the effectiveness of policies intended to protect vulnerable individuals.

The example of the GDPR underscores the need for clear, practical guidance on lawful data sharing, robust privacy protections, and frameworks that both facilitate responsible data use and uphold fundamental rights.

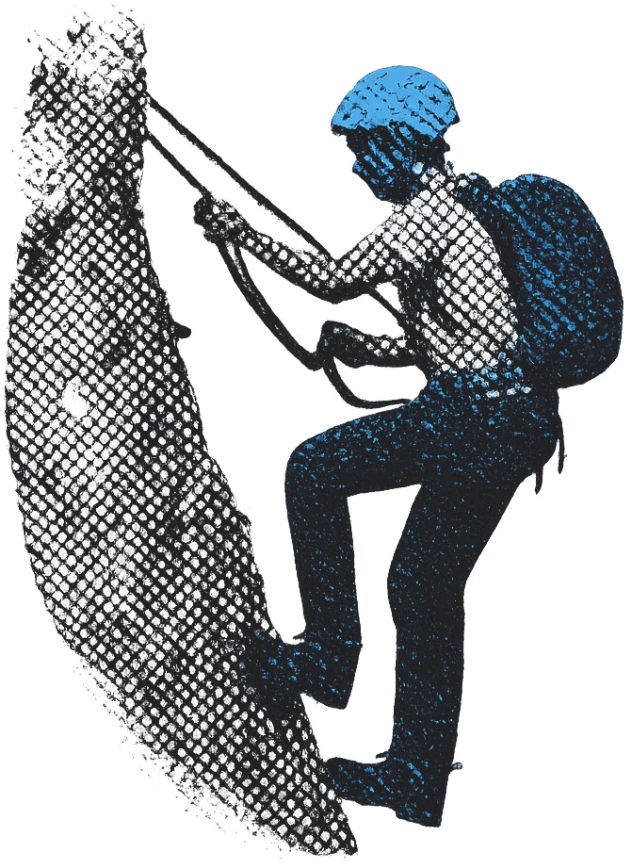
Technical barriers, such as limited interoperability between databases, and practical constraints including limited resources, staffing, and weak institutional ties between users and producers can further complicate access. In many cases, data sharing is ad-hoc and driven more by trusted relationships and political will than by established protocols or transparent collaboration.

Challenge 5: Data non-use and underuse

Even if accessible, data isn't always used to inform policymaking due to:

- Limited awareness: Data users, especially at local levels, may be unaware of existing data sources.
- Perceived irrelevance: Data may be seen as too outdated, too aggregated, and overall, not optimised for the specific needs of potential users.
- Lack of trust: Concerns about reliability, political influence, selective reporting, and opaque methodologies can fuel mistrust and lead to the dismissal of available data.
- Concerns about impact and control: Data producers may limit dissemination out of concern that information could be misused, misinterpreted, or worsen migrants' vulnerability, especially in sensitive political contexts. Without clear safeguards or control over secondary use, valuable data may remain unpublished or overlooked.

These five challenges, ranging from availability and access to trust and interpretation, can create significant barriers to the meaningful use of irregular migration data. Addressing them is essential for promoting evidence-informed policies that are more inclusive, effective, and transparent.



How to improve irregular migration data practices?

Addressing the challenges that hinder the use of irregular migration data requires targeted efforts to improve its quality, accessibility, and usability. These improvements are essential for supporting informed decisions and designing effective, targeted interventions. Several practices can be recommended:

- Harmonising definitions and measurement methodologies is crucial to enhance data compatibility across countries and systems. While multiple definitions of irregular migration will continue to serve different policy purposes, it is vital to clarify which definitions are used and why.
- Providing accessible and user-friendly data documentation can help prevent misinterpretation. Clear manuals should accompany datasets, outlining the definition of irregular migration applied, whether figures represent precise counts or estimates, potential limitations, and data quality concerns.
- Investing in capacity-building for key stakeholders involved in data collection and use is a vital step to address skill and expertise gaps. High-quality training materials, communities of practice facilitated by organisations dedicated to irregular migration can enhance data literacy. Pooled resources from Member States, research organisations and non-governmental organisations, alongside dedicated funding can encourage evidence-based policymaking.
- Strengthening data privacy safeguards is critical when enhancing data base interoperability. Linking data sources offers the potential for more comprehensive, real-time exchange of data, and responsive policy actions. However, this must be balanced with privacy protections under regulations, ensuring purpose limitation and protecting migrant rights.
- Scaling up local initiatives that can fill gaps on the profiles and circumstances of irregular migrants is essential. Local and municipal efforts offer pragmatic, frontline insights that can be expanded into national or EU-level programmes. National governments could, in turn, play a key role in creating standardized frameworks for data collection and sharing, ensuring consistency across different projects and regions.

Box 11.2: Spain's padrón system

Adèle Appriou, Jasmijn Slootjes and Ravenna Sohst

Spain's *padrón municipal de habitantes* (municipal register of inhabitants) is a notable example of how local registration systems can support the inclusion of irregular migrants while generating valuable data for public planning and service provision.

All residents, regardless of their status, are required to register with the *padrón*, which grants them access to municipal services such as education, health care, libraries, and language courses. The main advantage for irregular migrants registering in the *padrón* system is the possibility of obtaining *arraigo social* (legal residence) if they provide proof that they have lived in Spain for at least three years.

Registration requires minimal documentation—typically an ID and proof of address—which many municipalities apply flexibly to reduce barriers for irregular migrants. For instance, Barcelona actively encourages registration even for those without a fixed address, with city officials conducting field visits to verify the residence of individuals unable to provide formal proof. Another key feature of this public formation is its separation from other policy functions (e.g., immigration enforcement), which, along with outreach by civil society actors, helps build trust and encourage participation. This initiative enables the country to gather valuable information about all residents, including their age, country of origin, nationality, gender, and family or marital status.

While the *padrón* fosters inclusion, challenges remain. Registration requirements and practices vary between municipalities, with some cities facilitating registration for irregular migrants more actively than others. Issues around data accuracy—such as residents failing to de-register when they move—have also been noted. Nevertheless, Spain's *padrón* offers valuable lessons on how local initiatives can improve data collection and service access for irregular migrants.

Conclusion

While data on irregular migration has the potential to drive more effective, transparent, and responsive policymaking, this potential remains limited by persistent gaps and structural barriers. Enhancing the quality, accessibility, and responsible use of data is not only feasible, but necessary for fair and informed migration governance.

References

Delvino, N. (2017). City initiative on migrants with irregular status in Europe: Overview of good practices. Centre on Migration, Policy and Society, University of Oxford. https://www.compas.ox.ac.uk/wp-content/uploads/City-Initiative-on-Migrants-with-Irregular-Status-in-Europe-CMISE-report-November-2017-FINAL.pdf

European Union Agency for Fundamental Rights (2018). FRA Opinions Biometrics. https://fra.europa.eu/en/content/fra-opinions-biometrics

Slootjes, J., & Sohst, R. (2024). Towards the More Effective Use of Irregular Migration Data in Policymaking. Brussels: Migration Policy Institute Europe. https://doi.org/10.5281/zenodo.14276892

Slootjes, J., Sohst, R., & Kokkelmans, R. (2023). Mapping Stakeholders' Needs and Usage of Irregular Migration Data. MIrreM Briefing Paper No. D2.1. Brussels: Migration Policy Institute Europe. https://doi.org/10.5281/zenodo.7589494

Chapter 12

Progress, limits, and the need for sustained effort

Denis Kierans and Albert Kraler

Progress, limits, and the need for sustained effort

Denis Kierans and Albert Kraler

Irregular migration remains one of the most politically salient and technically challenging areas of migration data and policy in Europe. A range of stakeholders from academia to NGOs to government ministries collect and analyse data on irregular migration and are actively improving upon the evidence base in important respects. Still, quantitative information on irregular migration remains marked by significant gaps, inconsistencies, and contested interpretations. Too often, service provision, public discourse and decisions on migration management are made with reference to numbers that are partial, outdated, or biased, and presented without clear explanation of their scope and limitations. These problems persist in part because there is no European body tasked with sustaining cooperation, building capacity, or coordinating knowledge on irregular migration data. Overcoming this gap is essential if progress is to become cumulative rather than fragmented and short-lived.

Many of the implications set out in this Handbook will be familiar to those who have worked on improving statistics on irregular migration – or migration more generally – for years. Calls for clearer concepts, more robust quality assessments, scalable methodologies, greater transparency, stronger ethical safeguards, and closer alignment between data producers and users are not new. Yet to say there has been no change would overlook the progress of recent years. Across Europe, there is growing use of administrative registers to capture aspects of irregular migration and produce publicly available analysis; greater openness to innovative estimation methods; more ambitious and thoughtful surveys to boost coverage of hard-to-reach populations; and increased awareness of the importance of trust-based engagement.

The examples featured in this Handbook illustrate that such progress is possible and can be sustained. Spain's *padrón* system continues to register all residents regardless of status, enabling both service provision and valuable local-level statistics. The United Kingdom's Home Office publishes regular operational statistics and analysis on irregular arrivals and enforcement activity, providing an accessible view of specific flow indicators. Austria's Austrian Micro Data Centre offers a model for privacy-compliant linkage of administrative datasets to support longitudinal analysis. Italy's Regional Observatory for Integration and Multiethnicity (ORIM) survey in Lombardy demonstrates how inclusive, community-engaged data collection can be maintained over decades. The Mixed Migration Centre's global 4Mi survey shows how community-based enumerators can gather detailed information from people on the move at a global scale.

Various innovative methods have worked well in specific settings. In some places, promising approaches remain at the pilot level, dependent on individual champions, short-term funding, or local conditions unlikely to be replicable elsewhere. We also recognise that many of the implications set out in this Handbook are necessarily broad. We resist detailed prescription because context matters: initiatives need to be adapted to local legal frameworks, institutional arrangements and operational realities, drawing on the expertise of those who work closest to the data and the communities the data concern. There is a balance

to strike. Too much rigidity makes it difficult for a concept or method to travel; too much generality reduces its practical value. The case studies in this Handbook, particularly the innovative estimation methods and data-collection practices, are intended to provide concrete examples that, while rooted in specific contexts, are well suited to adaptation – if not immediately, then over time – into other settings.

Notwithstanding this need for techniques to be context-appropriate, it is clear that one thing which could help maintain and grow this body of innovation would be more coordination at the European level. The work documented in this Handbook sits within a longer trajectory of European and international efforts to improve irregular migration data. Earlier initiatives – in particular the CLANDESTINO project, which ended in 2009[1] – laid the groundwork for this Handbook and many of the good practices highlighted in it. While the past 15 years have seen new innovations and pilots, the lack of a consistent, Europe-wide mechanism for maintaining and building on these advances has limited their cumulative impact. A sustained, coordinated investment over that period would have undoubtedly produced a more harmonised, institutionalised, and widely adopted set of approaches across the continent.[2]

Such coordination would need to be mindful of ethical considerations. Techniques such as probabilistic matching of administrative records, capture–recapture analysis, mixed-method survey designs, and the integration of digital trace data have broad applicability when adapted with care. Local and municipal practices that build trust, such as the inclusivity of Spain's *padrón* regardless of migration status or the Regional Observatory for Integration and Multiethnicity's (ORIM) use of cultural mediators, show that the findings from integration research can go hand in hand with data quality. These case studies also underscore a broader point: effective irregular migration data systems are as much about relationships, governance, and institutional trust as they are about statistical methods.

Looking ahead, one of the hopes for this Handbook is that it will help to spur greater Europe-wide coordination on irregular migration data. This should take place under the leadership of key stakeholders, such as Eurostat, the Directorate General for Migration and Home Affairs of the European Commission (DG Home), the European Border and Coast Guard Agency (Frontex), the EU Agency for Asylum (EUAA) and the EU Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom, Security and Justice (eu-LISA). This type of leadership, coupled with long-term funding, would go a long way to ensuring that improvements in irregular migration data are sustained rather than episodic. This should include knowledge exchange and technical cooperation between researchers, NSOs, ministries and international organisations from different countries.

Irregular migration will never be fully knowable. Uncertainty is inherent in a phenomenon shaped by mobility, with strong incentives to remain "under the radar," and prone to shifting policies

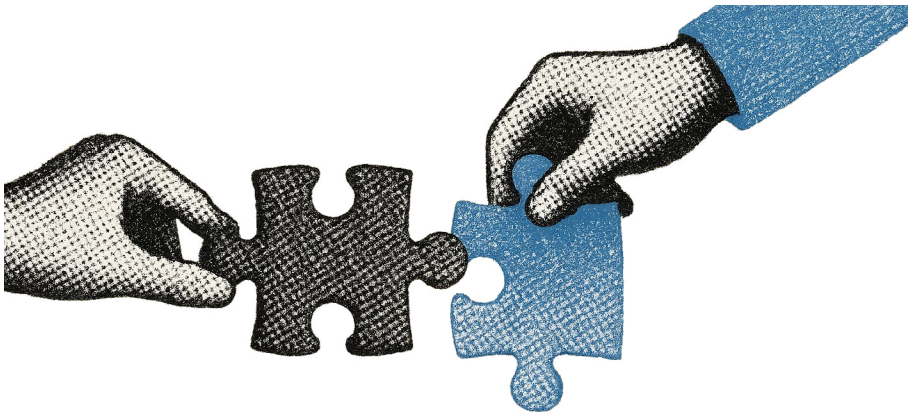
1 See https://cordis.europa.eu/project/id/44103. Results are also available from the archived project website at https://www.uni-bremen.de/fb12/irregular-migration-1.

2 The CLANDESTINO team sought additional funding from the European Commission, proposing a cooperation with the European Migration Network to undertake regular updates of the CLANDESTINO database and undertake related analyses, but their attempt was unsuccessful.

and legal frameworks. More information does not automatically reduce uncertainty or lead to greater insights. In some cases, an abundance of data can be more damaging than a scarcity of information, fuelling misinterpretation, selective use, or misplaced confidence in the numbers. The aim is not to eliminate uncertainty, but to manage it, grounding policy and public debate in evidence that is as reliable, transparent, and context-aware as possible.

The examples and approaches in this Handbook show that there is significant potential in activities already under way, that better data are achievable, and that their careful use can strengthen both understanding and governance of irregular migration. The challenge is to move from promising but isolated or short-lived initiatives towards a Europe-wide infrastructure for irregular migration data that is durable, well-resourced and collaborative. This will not be achieved quickly, but the building blocks already exist, providing a foundation for longerterm investment in more informed, transparent and credible data on irregular migration.

List of contributors



List of contributors

Petya Alexandrova, European Union Agency for Asylum, Malta

Jill Ahrens, Department for Migration and Globalisation, University for Continuing Education (Danube University) Krems, Austria

Adèle Appriou, Migration Policy Institute Europe, Belgium

Tuba Bircan, Brussels Institute for Social and Population Studies, Vrije Universiteit Brussel (VUB), Belgium

Julia Black, Missing Migrants Project, International Organization for Migration, Germany

Francesco Teo Ficcarello, Mixed Migration Centre Global, Uganda

Roberto Forin, Mixed Migration Centre Europe, Switzerland

Denis Kierans, Centre on Migration, Policy and Society (COMPAS), University of Oxford, United Kingdom

Albert Kraler, Department for Migration and Globalisation, University for Continuing Education (Danube University) Krems, Austria

Frank Laczko, former Head of IOM's Global Migration Data Analysis Centre (GMDAC), Germany, and series editor, Routledge books on "Global Migration Issues"

Arjen Leerkes, Maastricht Graduate School of Governance, Maastricht University, Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, and Research and Documentation Centre (WODC) of the Dutch Ministry of Justice and Security

Naomi Lindt, UNICEF and International Data Alliance for Children on the Move (IDAC), United States

Felipe Malle, National Migration Service, Chile

Marco Marsili, National Institute of Statistics, Italy

Rocco Molinari, Department of Statistical Sciences “Paolo Fortunati”, University of Bologna, Italy

Livia Elisa Ortensi, Department of Statistical Sciences “Paolo Fortunati”, University of Bologna, Italy, University of Bologna, Italy

Sebastián Palmas, UNICEF and International Data Alliance for Children on the Move (IDAC), Spain

Laura Peitz, Migration, Asylum and Integration Research Centre, Federal Office for Migration and Refugees (BAMF), Germany

Marzia Rango, UNICEF and International Data Alliance for Children on the Move (IDAC), Italy and the United States

Julibeth Rodríguez, National Migration Service, Chile

Alejandra Rodríguez-Sánchez, Department of Administrative and Political Sciences, University of Potsdam, Germany

Aslı Salihoğlu, Centre on Migration, Policy and Society (COMPAS), Oxford

Filip Savatic, Sciences Po, France

Randy Stache, Federal Office for Migration and Refugees (BAMF), Germany

Jon Simmons, UK Home Office Analysis & Insight directorate, United Kingdom

Ann Singleton, School for Policy Studies, University of Bristol, United Kingdom

Lalaine Siruno, United Nations University Maastricht Economic and Social Research Institute on Innovation and Technology and Maastricht University, the Netherlands

Jasmijn Slootjes, Migration Policy Institute Europe, Belgium

Ravenna Sohst, Migration Policy Institute Europe, Belgium

Lucy Swinnerton, UK Home Office Analysis & Insight directorate, United Kingdom

Jasper Tjaden, Department of Administrative and Political Sciences, University of Potsdam, Germany

Carlos Vargas-Silva, Centre on Migration, Policy and Society (COMPAS), University of Oxford, United Kingdom

Teddy Wilkin, European Union Agency for Asylum, Malta

Danzhen You, UNICEF and International Data Alliance for Children on the Move (IDAC), United States

This Handbook brings together concepts, findings, methods, and case studies to offer a clear, practical understanding of irregular migration data. It addresses the challenges of conceptualising, measuring, interrogating, and using data on one of Europe's most politically sensitive migration issues. Drawing on examples from across Europe and beyond, it provides guidance on concepts and definitions, ethics, estimation methods, data innovation, and policy application. It is designed to support policymakers, practitioners and researchers seeking more informed, transparent, and coordinated approaches to irregular migration data.

MIRREM
Measuring Irregular Migration

Funded by the European Union
UK Research and Innovation
Toronto Metropolitan University
Canada Excellence Research Chair in Migration & Integration
University of Krems Press