MIRREM

Measuring Irregular Migration

www.irregularmigration.eu

ESTIMATING IRREGULAR MIGRATION - A REVIEW OF TRADITIONAL AND INNOVATIVE DATA AND METHODS

MIrreM Working Paper 4/2023

AUTHORS:

Alejandra Rodríguez Sánchez and Jasper Tjaden

Deliverable Information:

Project Acronym:	Measuring irregular migration and related policies (MIrreM)
Project No.	101061314
WP	WP6 - Methods Innovation Lab
Deliverable Type:	Report
Deliverable Name	D6.1 Review report on innovative approaches to estimating irregular migration
Version:	2
Date:	17/10/2023
Responsible Partner:	University of Potsdam / UP
Contributing Partners:	
Authors	Alejandra Rodríguez Sánchez (UP), Jasper Tjaden (UP)
Reviewers:	Albert Kraler (UWK), Jill Ahrens (UWK), Norbert Cyrus (UOS)
Dissemination Level:	Public

Revision History:

Revision	Date	Author	Organisation	Description
0.1	30/08/2023	Alejandra Rodríguez Sánchez and Jasper Tjaden	UP	Initial draft sent for review
0.2	18/09/2023	Albert Kraler and Jill Ahrens	UWK	Review
0.3	29/09/2023	Alejandra Rodríguez Sánchez and Jasper Tjaden	UP	Pre-final draft sent for check of open issues
0.4	29/09/2023	Albert Kraler	UWK	Editorial adjustments, additional literature
1.0	01/10/2023	Alejandra Rodríguez Sánchez and Jasper Tjaden	UP	Final revisions

1.0	10/10/2023	Norbert Cyrus	UOS	Additional comments added
2.0	17/10/2023	Albert Kraler	UWK	Correction of erroneous definition and typos. Resubmission.

Executive Summary

In this paper, we provide a concise review of 20 methodological approaches to measuring irregular migration for a non-technical audience. First, we revisit established approaches covered in previous reviews (such as CLANDESTINO). Second, we review innovative approaches which apply new methodologies or make use of novel data. We include any empirically tested attempts to measure irregular migration stocks or flows. Each approach is discussed in the form of a one-page bulletin that describes the main idea behind the approach, the required data sources, the definition or coverage of irregular migrant groups, the reliability and scalability of the approach, the main assumptions of the approach as well as potential ethical issues for researchers.

We find an increasing diversity and complexity of approaches that made important advances to overcome limitations or drawbacks of traditional approaches. Despite the progress made in estimating irregular migration in the last decade, approaches remain highly fragmented depending on the type of data source used for the estimation. Some approaches can only be applied in specific countries due to data limitations. Any evaluation of the relative performance of various approaches is still missing due to the absence of comparative assessments testing various approaches for the same context. Simulation studies could offer a way forward. One trend in the last decade has been the emergence and use of digital data which offers promising avenues for future research.

TABLE OF CONTENTS

LIST OF TABLES

THE MIRREM PROJECT

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

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

1. INTRODUCTION

Why	This report has been produced in the context of the EU-funded project "Measuring irregular migration and related policies" (MIRREM). MIRREM is a project aimed at updating the seminal CLANDESTINO project – an EU–funded project (2007-2009) which reviewed data on irregular migration over a decade ago. In this report, we provide a concise review of methodological approaches to measuring irregular migration building on previous reviews provided by the CLANDESTINO project (Jandl, 2011) but complementing it with advances in methods and data. The interest in the phenomenon of irregular migration among policy makers and academics was high then and remains high now, especially after substantial migration movements to the EU since 2014 and persistent migration from Central America into the US. Estimating the size of irregular migrant populations in any given country remains a challenging endeavour. However, as this report will show, new approaches and data sources can help overcome some of the difficulties as highlighted in published research we review.
What	We review broader 'methodological approaches' rather than individual studies. Approaches considered in this report include any empirically tested attempts to measure irregular migration stocks (e.g., the total population size in a given territory during a particular year) or flows (i.e., the events that increase or decrease irregular migrant stock within a given territory during a certain period). Approaches that have been proposed in other fields, but have not yet been applied to migration, are not included in this review. In addition, we did not include approaches that have been applied to migration and mobility (such as approaches using digital data sources) but not yet to the case of *irregular* migration. Methodological approaches can be distinguished based on which type of data was used and/or the type of estimation methods or calculations that were employed.
Target audience	The report is aimed at non-technical audiences. Our ambition is to convey the underlying rationale, assumptions and limitations of each of the approaches without delving into the mathematics behind each approach. The overarching objective is to help policy makers, journalists and academics from various fields navigate this complex and ever-growing field of research.

How	The starting point for our review is a summary of traditional approaches which have been covered in previous reviews, particularly in the CLANDESTINO project (e.g., Jandl, 2011). Subsequently, we review "innovative approaches" which were not covered in prior reviews and, usually, involve either new data sources or innovative estimation procedures. To identify innovative approaches, we looked for studies published after 2010 that directly referenced any of the main studies included in that review and the review itself. In this way, we identified recent publications that were innovative in terms of data, methods, or both, with applications in multiple countries. Following this, we checked already existing overviews of methods for population size estimation methods for 'hard-to-reach' or 'hard-to-sample' human groups, such as research on drug addiction, homelessness, or human trafficking (Marpsat & Razafindratsima, 2010; Wesson, et al., 2017; Xu et al., 2022). Like in research on irregular migration, research from these fields share the challenge that their subjects are not included in official sampling frames and that the issue at hand is sensitive. We then looked for studies employing any of the identified innovative methods in the context of irregular migration. We review in total 21 approaches (11 "traditional" and 10 "innovative" ones). Each approach is discussed in the form of a one-page bulletin including a description of the main idea behind the approach, the required data sources, the definition or coverage of irregular migrant groups, the reliability and scalability of the approach, the main assumptions of the approach as well as potential ethical issues related to measurement. Each bulletin closes with a few published references. The review focusses largely on limitations inherent in the respective data sources required for each approach instead of limitations related to specific statistical estimation procedures. It Is important to note that the review does not rank or rate approaches. Each approach has severe limitations and empirical, comparative assessments are still largely absent in this literature.
Structure	Each approach is discussed in the form of a one-page bulletin including a description of the main idea behind the approach, the required data sources, the definition or coverage of irregular migrant groups, the reliability and scalability of the approach, the main assumptions of the approach as well as potential ethical issues related to measurement. Each bulletin closes with a few published example references. The review focusses largely on limitations inherent in the respective data sources required for each approach instead of limitations related to specific statistical estimation procedures. The report is structured as follows: In section 2, we explain key methodological decisions for this review including which approaches were selected and which criteria were assessed. Section 3 presents each of the approaches, and section 4 discusses broader trends and future developments.

2. METHODS

2.1. Selection of approaches

Traditional vs innovative approaches

In total, we have selected 21 approaches for this review. We distinguish "traditional" from "innovative" approaches to estimating irregular migration. There is no strict dividing line between both groups. "Traditional" refers to approaches which have been included in previous reviews and are generally known to provide estimates of the irregular migrant population, especially those described in Jandl (2011). Our assessment of traditional approaches builds heavily on this previous review. Traditional approaches have been widely discussed and applied in various countries. We included previous approaches in this review for reasons of completeness and ease of navigation for the reader, but also because some work has improved upon some of the shortcomings of traditional approaches.

We defined "Innovative" approaches as those which have been applied more recently. "Innovative" methods either use novel data sources (e.g., digital data) or apply a new estimation method to standard data sources. Innovate approaches have been identified through literature review and discussions with experts[1].

Data sources

Another helpful dimension to guide readers is the main data source used in the approach. Many traditional approaches either rely heavily on administrative data sources, such as residence or housing registry; visa and residence registry, border and policy registries, work permit registries, firm-level economic data, etc.; whereas other approaches take their input from sample surveys (either general population surveys or dedicated migration surveys); censuses or, increasingly, a mix of various data sources. Innovative approaches make increasing use of digital data sources which present new challenges for inferring population estimates.

---

[1] Albert Kraler; Migration Policy Institute; Ettore Recchi; and the participants to the the MIRREM expert workshop "Innovative approaches to measuring irregular migration", which took place on the 25th of April 2022 in Brussels. For a recording of the presentations click [here](#).""

Overview

Table 1 shows a graphical illustration of the approaches we reviewed in this report including a broad grouping by type of data employed. We attempted to unite more specific methods under a common umbrella defined by the main idea of the approach rather than narrowing in on differences at the margins. For example, there are many new methods which have been developed to sample "hidden" populations. We have grouped all these approaches under "self-identification in surveys". Another example involves multiple approaches to matching individuals across administrative data sources. We grouped these approaches under the label "capture – recapture" and "multiple systems estimation", two of the most common labels for this family of approaches. There are no formal distinctions between approaches available, as a result, our decisions were mostly guided by readability and accessibility.

Table 1 ***Overview of methodological approaches covered in this review***

	Government administrative & census	
Traditional	Capture-Recapture / Multiple Systems Estimation	
	Regularisation & formal status adjustments	
	Life course events	
	Labour demand models	
	Flow-Stock Models	
	Irregular border crossings & apprehensions	
	Database systems enabling the identification of visa overstayers	
Innovative	Using consular registers of migrant communities in combination with other data-sources	
	Drivers' licenses	
	Consumption data	
	Biometrics & surveillance	
	Non-government administrative data	
Innovative	Institutional registers: College enrolment, prison population, and human trafficking	
	Missing, disappeared or deceased migrants	
	Survey data	
Traditional	Self-identification in surveys	
	Expert surveys / Delphi surveys	
Innovative	Statistical imputation	
	Retrospective surveys	
	Mixed data sources	
Traditional	Multiplier/ simple extrapolation	
	Residual estimation method	
Innovative	Alternative residual: labour force survey and social security register	
	Digital data	
Innovative	Online search behaviour	

Note: own elaboration

2.2. Review criteria

The starting point for selecting review criteria was a brief reflection on common standards in the evaluation of scientific evidence, such as PRISMA, GRADE and others (Guyatt et al., 2008; Page et al., 2021). After internal deliberation within the project team, we decided on the following set of criteria as a trade-off between keeping scope and complexity manageable and the assessment as non-technical as possible. Table 2 provides definitions of each review criteria.

Table 2 **Main dimensions for the review of methodological approaches**

Dimension	Description
Main idea	This category attempts to boil down the main underlying rationale of the methodological approach into a few sentences. We frame these basic ideas in a way to make them intelligible for non-technical readers.
Data source	This category lists the main data source (administrative, survey, census, mixed) and names examples of specific sources if applicable.
Coverage / Definition	This category describes the degree to which the approach covers all sub-groups of irregular migrants. Irregular migrants are a heterogenous, and fluid group defined by a myriad of legal, political and empirical boundaries (Kraler, 2023). The provision of high-level estimates of group sizes often deflects from the fact that such estimates only reflect one part of the overall group. Understanding who is and who is not included by specific approaches informs the type of biases present in any estimation.
Estimation assumptions	This criterion highlights the number and salience of assumptions underlying the calculation or estimation. The question we ask is what researchers would need to assume for the estimate to reflect the "true" estimate of the whole irregular migration stock or flow. The assumptions speak to the many biases and imperfections of each method as, in most cases, the assumptions are not met in practice.
Reliability	Reliability is a standard dimension for assessing the quality of research. It refers to the degree to which the approach can be applied repeatedly and provide consistent results. In the context of irregular migration estimates, we consider the likelihood of the methods providing comparable and consistent results if they were applied year to year.
Scalability	Scalability refers to the degree to which the approach can be applied in other countries. Scalability often refers to how widely certain data sources are available.
Assumptions	This criterion highlights the amount and severity of assumptions underlying the calculation or estimation. The question we ask is what researchers would need to assume for the estimate to reflect

	the "true" estimate of the whole irregular migration stock or flow. The assumptions speak to the many biases and imperfections of each method given that, in most cases, the assumptions are not met in practice.
Ethical issues	This criterion reflects potential ethical concerns for researchers when applying each respective method. It is important to note here that this report does not discuss ethical issues relating to the question of whether estimates should be provided or not given its many potential political misuses and potential harm for migrants. This discussion is captured elsewhere (Cyrus., 2023). In this report, ethical issues refer to specific considerations arising from the data source, data used, or the way the method is applied, as well as to the risks that the application of the method might entail for migrants, by, for example, revealing geographic location or specific occupations where migrants are employed, etc. Often these issues refer to unethical ways in which data is collected, linked, analysed or presented.
Examples	In this final section, we reference a few published works relevant to each approach. The selection of studies appearing here is not exhaustive, but it represents ways in which the general approaches have been used in empirical research. The listed studies provide the reader with additional material to further explore one approach in detail.

Note: Own elaboration.

3. REVIEW

3.1. Traditional approaches

The following section reviews "traditional" approaches. As mentioned in the introduction, we consider "traditional" approaches those that have been the subject of previous reviews (Jandl, 2011; Pinkerton et al., 2004). In this section, we review approaches to measuring irregular migration stocks and flows. The methods covering stocks estimate the cumulative size of the resident irregular migrant population measured at one specific moment in time (e.g., the year 2022), whereas methods to estimating flows capture the changes in stocks in a specific period (e.g., the interval 2015-2020).

3.1.1. Multiplier/ simple extrapolation

Main idea	Derive a percentage or fraction representing "irregular" migrants relative to the population of all migrants. Then apply this percentage to other contexts or larger geographic boundaries.
Data source	Different types of data sources. survey data; administrative data such as criminal records, apprehension data, hospital records, labour inspection, etc.
Coverage / Definition	The method is agnostic regarding the type of irregular migrant population to be estimated.
Estimation assumptions	The estimation is a simple extrapolation by multiplying a total population of migrants with the assumed fraction of those with irregular status. The main assumption is that the fraction of irregular migrants obtained from another population is transferrable to a different context. Alternatively, a fraction of irregular migrants is derived in one data source (e.g. crime statistics) and then extrapolated to the whole population in the country.
Reliability	The reliability of this method is low in cases when it is applied outside of the context in which it has been derived. In many cases, a percentage has been derived from one study based on one time and one location. It is then applied to other contexts which may involve different underlying populations. In cases where the fraction of

	irregular migrants in a sub-population (i.e. crime suspects) is extrapolated to the whole population, inclusion into the sub-sample is selective, non-random, and most likely, not stable across location and years.
Scalability	The method is scalable to most contexts because it is flexible based on the sources of its two main components: The number of all migrants in a location and the assumed fraction of irregular migrants (derived from a different study).
Ethical issues	There are no ethical issues beyond the high risk of misidentifying irregular migrants.
Examples	Burgers, J. (1996). Natte vingers en vuile handen; over het schatten van het aantal illegale vreemdelingen. Een reactie op Böcker en Groenendijk. Migrantenstudies, 12(1), 14–26.
	Drbohlav, D., & Lachmanová, L. (2008). Undocumented migration. Counting the Uncountable. Country report - The Czech Republic | European Website on Integration. https://ec.europa.eu/migrant-integration/library-document/undocumented-migration-counting-uncountable-country-report-czech-republic_en
	Espenshade, T. J. (1995). Using INS Border Apprehension Data to Measure the Flow of Undocumented Migrants Crossing the U.S.-Mexico Frontier. International Migration Review, 29(2), 545–565. https://doi.org/10.1177/019791839502900209
	Jandl, M. (2009). A multiplier estimate of the illegally resident third-country national population in Austria based on crime suspect data. Database on Irregular Migration. Working Paper, (2), 2.
	van Meeteren, M., Engbersen, G., & van San, M. (2007). Irreguliere immigranten in België. https://repub.eur.nl/pub/12074

3.1.2. *Capture-Recapture / Multiple Systems Estimation*

Main idea	The "Capture-Recapture" method, also known as the "Capture-Mark-Recapture" method, is a statistical technique used in ecology and population studies to estimate the size of a population based on a sample of individuals that have been "captured", marked (or tagged), released, and then captured again in a subsequent sample. In the context of irregular migration, the basic approach uses one data source which can identify irregular migrant status (i.e., police or border apprehensions) and then uses one or more data sources of the

general population to assess how many irregular migrants appear in them by matching individuals across registers. The percentage of detected migrants (from the first data source) of the total population (in the second source) is used to extrapolate the total number of irregular migrants in the population via multiplying the share with the total number of detected irregular migrants from the first source.

There are several variations of this approach with increasing degree of complexity, including Multiple Systems Estimation (UNODC, 2022), repeated capture-recapture (Van Der Heijden et al., 2003), and random-effects mixed modelling (Zhang, 2008). All approaches share the general empirical strategy consisting of linking individuals across two or more administrative data sources.

Data source	Two or more "linkable" administrative data sources with at least one containing information on irregular status (for example, a police records and work permit registry).
Coverage / Definition	The type of irregular migrant population which is identifiable in this approach depends entirely on the definition used in the administrative data sources which record "irregular" migrant status. Administrative data sources are usually designed to implement government programs or to enforce laws. As such they do not reflect a random sample of residents in a location and, as a result, likely only capture a selective subsample of irregular migrants in a country. For example, crime suspect data (Vogel, 2015; Jandl, 2011) have limited coverage of minors.
Estimation assumptions	The approach rests on several assumptions, all of which are highly problematic in the context of irregular migration.

1) All data must relate to persons belonging to a "closed" population. Real populations, however, change constantly through births, deaths, immigration and emigration.

2) Each person must have the same chance to be included in each list. Depending on the administrative sources used, this is unlikely for irregular migrants. Irregular migrants, for example, may be more likely to appear in a health register than in police data (given that they actively attempt to abstain from enforcement authorities).

3) Inclusion in one list must be independent of inclusion in another list. This is also unlikely in the context of irregular migration. If individuals become known to the authorities, they are more likely to be "visible" in other administrative registers compared to those who have never contacted authorities at all.

Reliability	If the same data sources are available, the approach provides consistent and comparable results given that definitions and data inclusion criteria remain the same.
Scalability	The approach is easily scalable in countries with administrative records which can link individuals across registries. Many countries lack high-quality registries and others prevent data linkage due to data protection reasons.
Ethical issues	Linking personal information across administrative data sources is commonly done without informed consent of the individual which introduces issues regarding data privacy and data protection. For these reasons, some countries prohibit record linkage.
Examples	Bird, S. M., & King, R. (2018). Multiple Systems Estimation (or Capture-Recapture Estimation) to Inform Public Policy. Annual Review of Statistics and Its Application, 5(1), 95–118. https://doi.org/10.1146/annurev-statistics-031017-100641
	UNODC. (2022). Monitoring human trafficking prevalence through multiple system estimation. UNITED NATIONS. https://www.unodc.org/documents/data-and-analysis/tip/2022/MSE_TIP_UNODC_ENG.pdf
	Van Der Heijden, P. G. M., Cruyff, M., & Van Houwelingen, H. C. (2003). Estimating the Size of a Criminal Population from Police Records Using the Truncated Poisson Regression Model. Statistica Neerlandica, 57(3), 289–304. https://doi.org/10.1111/1467-9574.00232
	Zhang, L.-C. (2008, June 23). Developing methods for determining the number of unauthorized foreigners in Norway. Ssb.No. https://www.ssb.no/en/befolkning/artikler-og-publikasjoner/developing-methods-for-determining-the-number-of-unauthorized-foreigners-in-norway--18906

3.1.3. Regularization & formal status adjustments

Main idea	A government regularization program (also known as amnesty programs) provides one-off access to regular residence or work permits to a defined group of the "irregular" population. As a side-effect of this social policy, information on the potential size of the irregular migrant population is obtained. Case-by-case status adjustments are also referred to as 'silent regularization' processes,

given that they are implemented on an ongoing basis without formal announcement.

Data source	Administrative data – usually residence permit applications.
Coverage / Definition	Regularization programs often have eligibility criteria. This means that only certain types of irregular migrants can apply. Common criteria are a minimum period of residence in the country and/ or employment. In some cases, even eligible migrants may decide not to apply fearing rejection of their application and subsequent removal.
Estimation assumptions	The estimate of the size of the irregular migration population resulting from the number of applications for legal residence rests on the assumption that a) people don't travel to the country to apply for the program, b) people do not apply multiple times, c) people with legal status do not apply, and d) all eligible persons apply. Regarding d), it is likely that some eligible persons do not apply due to fear of deportation in case of failed application. Assumptions b) and c) are easy to adjust for if data availability permits. Assumption a), on the other hand, is harder to ascertain, however, almost all "regularisation programmes" require proof of duration of residence.
Reliability	One-off regularization programs may have low reliability in terms of comparability over time because each program is usually designed differently, thus yielding different estimates. Status adjustment policies may have higher comparability across time if the eligibility criteria remain the same.
Scalability	The approach is scalable to any country, however, politically, regularization programs face opposition for various reasons. One of the main concerns is that "regularization programs" attract more migrants to the country.
Ethical issues	Ethical issues may arise when irregular migrants who apply for the regularization, but do not receive legal status, face risk of expulsion or other negative consequences. In the context of research studies on regularisation, studies may uncover informal and unlawful practices, for example if uncovering the widespread use of 'fake contracts', which could include the corrupt practices such as the sale of employment contracts or strategies to circumvent narrow eligibility criteria, e.g. if migrants and employers in any other economic sectors use channels for domestic workers or agricultural workers, hoping to change to an appropriate permit upon expiry of the original residence permit issued at regularisation. Such evidence may be used for implementing stricter regularisation policies.

Examples	Baldwin-Edwards, M., & Kraler, A. (2009). REGINE - Regularisations in Europe. Study on practices in the area of regularisation of illegally staying third-country nationals in the Member States of the EU. International Centre for Migration Policy Development (ICMPD). https://ec.europa.eu/migrant-integration/sites/default/files/2009-04/docl_8193_345982803.pdf
	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
	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
	Sunderhaus, S. (2007). Regularization programs for undocumented migrants. Migration Letters, 4(1), 65–76. https://www.ceeol.com/search/article-detail?id=493672

3.1.4.   Residual estimation methods

Main idea	The main idea behind this approach is to subtract an estimate of the legal migrants from an estimate of the total migrant population counted in the census or an alternative data source, and which includes irregular migrants. The remaining "residual" is considered the size of the irregular migrant population.
Data source	Census or complete register (see, for example, the '*padron*' system in Spain); one additional register or survey including legal migrants only.
Coverage / Definition	In theory, this approach should cover all types of irregular migrant populations if they have been included in the general population survey or census.
Estimation assumptions	The key assumption of this approach is that one data source (e.g., census or large population survey), indeed, covers all migrants in the country (regardless of their status) and that the other sources capture all legal migrants. The coverage error of the census, for example, can impact estimates obtained through the residual method. Migrants, for several reasons, may not be included in official sampling frames for censuses or surveys. The approach can adjust estimates for undercounting in the census, however, uncertainty remains whether additional groups have been left out – a critical

point in this approach. If the source identifying legal migrants is based on a survey, issues surrounding response bias may apply. For example, irregular migrants may falsely report a legal status. Lastly, the approach requires assumptions about return migration as well as mortality. It is possible that a percentage of the "residual" irregular population already left the country without appearing in any of the data sources.

Reliability	The approach provides reliable results if data sources remain consistent over time.
Scalability	This approach is not easily scalable as many countries do not have the required data sources.
Ethical issues	Ethical issues are less pronounced compared to other approaches because the residual method does not allow to identify specific individuals with irregular status but rather to estimate group sizes by state or country.
Examples	Baker, B. (2021). Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2015–January 2018 (Population Statistics). Department of Homeland Security's Office of Immigration Statistics. https://www.dhs.gov/sites/default/files/publications/immigration-statistics/Pop_Estimate/UnauthImmigrant/unauthorized_immigrant_population_estimates_2015_-_2018.pdf
	Jolly, A., Thomas, S., & Stanyer, J. (2020). London's children and young people who are not British citizens: A profile. https://doi.org/10.13140/RG.2.2.13839.12964
	Passel, J. S. (2006). The Size and Characteristics of the Unauthorized Migrant Population in the U.S. Estimates Based on the March 2005 Current Population Survey. Pew Hispanic Center. https://www.procon.org/wp-content/uploads/sites/40/immidoc6.pdf
	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
	Warren, R. (2021). In 2019, the US Undocumented Population Continued a Decade-Long Decline and the Foreign-Born Population Neared Zero Growth. Journal on Migration and Human Security, 9(1), 31–43. https://doi.org/10.1177/2331502421993746

3.1.5.   *Self-identification in surveys*

Main idea	Surveys may include questions eliciting the legal status of survey respondents. This can either be done through direct questions (such as "Do you have a legal residence permit") or indirectly through asking about rights associated with legal status.
Data source	General population surveys or dedicated migration surveys
Coverage / Definition	Surveys are flexible in terms of the type of irregular migration they cover. Coverage depends entirely on the sampling strategy. However, more vulnerable populations are commonly harder to survey (Reichel & Morales, 2017). Persons on the move, without residence or those fearing prosecution may be less likely to participate.
Estimation assumptions	A key assumption is that irregular migrants participate in surveys. For groups fearing deportation or those most mobile, this is unlikely to be the case. In addition, the approach assumes that proper sampling strategies provide probability samples which can be generalized to the overall population. In practice, this standard is difficult to reach. However, several sampling approaches have been developed particularly for "hidden" or "hard-to-reach" populations (Reichel & Morales, 2017), including venue-based sampling (Baio et al., 2011) and network-based sampling (Heckathorn & Cameron, 2017), among others. Lastly, one important assumption is that individuals who participate in the survey provide truthful answers (response bias). Given that irregular status is a highly sensitive question, respondents may not provide accurate answers. Some methods, such as list experiments, have been developed to mitigate response biases (e.g., McKenzie & Siegel, 2013).
Reliability	If sampling strategy remains the same and there are no changes in the behaviour of the population, the method can produce reliable estimates.
Scalability	Resources and technical expertise permitting, the approach can be scaled to any country or region. However, accuracy of the estimates may vary depending on the scale and quality of available sampling frames (such as registers, censuses etc.).
Ethical issues	Ethical issues may arise if informed consent is missing, or data is not sufficiently anonymized.
Examples	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. https://doi.org/10.1111/imre.12059

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. https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/centre-sampling-technique-in-foreign-migration-surveys-a-methodological-note.pdf

Huddleston, T., & Tjaden, J. (2012). Immigrant Citizens Survey. How immigrants experience integration in 15 European cities. King Baudouin Foundation and MPI. https://www.om.acm.gov.pt/documents/58428/296070/immigrant+citizens+survey.pdf/a3e84f37-5825-48f0-b4aa-125970595cab

Martí, M., & Ródenas, C. (2012). Measuring International Migration through Sample Surveys: Some Lessons from the Spanish Case. Population, 67(3), 435–463. https://www.cairn-int.info/journal-population-2012-3-page-435.htm

McKenzie, D., & Siegel, M. (2013). Eliciting illegal migration rates through list randomization. Migration Studies, 1(3), 276–291. https://doi.org/10.1093/migration/mnt018

Reichel, D., & Morales, L. (2017). Surveying immigrants without sampling frames – evaluating the success of alternative field methods. Comparative Migration Studies, 5(1), 1. https://doi.org/10.1186/s40878-016-0044-9

3.1.6. Expert surveys / Delphi surveys

Main idea	Instead of directly interviewing the target population, experts are surveyed and asked to estimate the population size of irregular migrants in a certain area. In the Delphi survey approach, experts are surveyed multiple times, each time receiving responses from their peers. Experts are supposed to reach consensus through feedback loops and discussion.
Data source	Primary data collection based on survey with pool of experts.
Coverage / Definition	Coverage of irregular migrant population is flexible and depends entirely on design of the questionnaire.
Estimation assumptions	The key assumption is that experts have more knowledge, first-hand experience and access to varied data sources. As a result, they make more accurate predictions. However, research has shown repeatedly that experts are often not more likely to accurately estimate events than non-experts (Tetlock, 2017; Grossmann et al., 2023).

Reliability	The reliability of this approach is low given that results vary substantially depending on the set of experts involved.
Scalability	The approach is easily scalable as many countries rely on expert judgement for other policy areas.
Ethical issues	Ethical issues may apply if experts have an incentive to exaggerate the size of the irregular migrant populations as they benefit personally from more salience of the issue.
Examples	Acostamadiedo, E., Sohst, R. R., Tjaden, J., Groenewold, G., & de Valk, H. (2020). Assessing Immigration Scenarios for the European Union in 2030: Relevant, Realistic and Reliable? International Organization for Migration. https://publications.iom.int/books/assessing-immigration-scenarios-european-union-2030
	Drbohlav, D., Dzúrová, D., Čermák, Z., Janská, E., Čermáková, D., & Medová, L. (2008). Immigrants' irregular economic activities in the Czech Republic (a multi-approach study). Transfer: European Review of Labour and Research, 14(4), 639–652. https://doi.org/10.1177/102425890801400409
	Grossmann, I., Rotella, A., Hutcherson, C. A., Sharpinskyi, K., Varnum, M. E. W., Achter, S., Dhami, M. K., Guo, X. E., Kara-Yakoubian, M., Mandel, D. R., Raes, L., Tay, L., Vie, A., Wagner, L., Adamkovic, M., Arami, A., Arriaga, P., Bandara, K., Baník, G., … The Forecasting Collaborative. (2023). Insights into the accuracy of social scientists' forecasts of societal change. Nature Human Behaviour, 7(4), 484–501. https://doi.org/10.1038/s41562-022-01517-1
	Jandl, M., Hollomey, C., & Stepien, A. (2007). Migration and irregular work in Austria: results of a Delphi-study. (No. 90; International Migration Papers). International Labour Organization and International Centre for Migration Policy. https://www.icmpd.org/file/download/47428/file/Migration%2520and%2520Irregular%2520Work%2520in%2520Austria-%2520Delphi%2520study.pdf
	Tetlock, P. E. (2017). Expert Political Judgment: How Good Is It? How Can We Know? - New Edition. Princeton University Press. https://doi.org/10.1515/9781400888818

3.1.7.   Life course events

Main idea	The main idea of this approach is to compare the actual probability of certain demographic life course events for migrants with the expected probability of such events in non-migrant populations. Demographic events are, for example, birth, deaths and hospitalizations. Demographic research has established stable and known probabilities of such events for many groups, including migrants. If, for example, the mortality rate for a certain migrant group exceeds the projected rate, one can assume that 'unregistered' or irregular migrants were present who drove the excess rates.
Data source	Administrative data sources (i.e., mortality register; hospitalization records) of the total population (including regular and irregular migrants).
Coverage / Definition	In principle, this approach can cover all irregular migrant groups who experience life events which are recorded in administrative registers.
Estimation assumptions	One assumption is that the rate at which demographic events occur for irregular migrants is of a comparable magnitude as the rate at which events occur for regular migrants from the same country and demographic group. Given that irregular migrants are a vulnerable group which is often excluded from health services or social assistance, the probabilities of life events may be different. Another key assumption is that the main life events of irregular migrants are recorded in mortality or hospital registers. It is possible that irregular migrants do not visit the hospitals altogether due to risk of detection or that they return to their home countries at older ages.
Reliability	Given the robustness of known risks of demographic events such as death or hospitalization, the approach is highly reliable if data sources are consistently available over time. For young irregular migrants, due to the low number of deaths in such age groups, the method is less reliable.
Scalability	The approach is scalable to any country which has a detailed mortality or hospitalization registry.
Ethical issues	Ethical issues are less pronounced because the method does not identify irregular persons individually but provides a group size estimate.
Examples	Robinson, J. G. (1980). Estimating the Approximate Size of The Illegal Alien Population in the United States by the Comparative

Trend Analysis of Age-Specific Death Rates. Demography, 17(2), 159–176. https://doi.org/10.2307/2061057

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

Wanner, P. (2002). Compter les clandestins. Méthodes d'estimation de la population sans autorisation de séjour à partir des données sur la population (No. 13; Discussion Paper). Swiss Forum for Migration Studies. http://www.unine.ch/files/live/sites/sfm/files/shared/pub/dp/dp13.pdf

3.1.8. *Labour demand models*

Main idea	The main idea is to estimate the labour demand for certain sectors based on reported economic output (i.e., how many workers are required to produce a given number of units of output). The estimate of the irregular migrant worker is the difference between the estimated labour demand and the reported number of foreign workers (Hess, 2006). A similar approach is based on estimating the size of the "shadow economy" in combination with assumptions of the share of migrants working irregularly (Schneider & Enste, 2000).
Data source	Administrative data – Firm-level economic indicators
Coverage / Definition	This approach only covers employed migrant workers in certain sectors.
Estimation assumptions	Inferring underreported migrant labour from labour demand assumes that firms in the same sector have the same level of productivity, i.e., that firms need a similar number of workers to produce the same amount of output.
Reliability	If reporting standards remain similar over time, the method should produce reliable results.
Scalability	The approach is scalable to any context where data on firm-level economic productivity is reported continuously.
Ethical issues	Ethical issues are less pronounced because the method does not identify irregular persons individually but provides a group size estimate.

| Examples | 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 |
| | Schneider, F., & Enste, D. H. (2000). Shadow Economies: Size, Causes, and Consequences. Journal of Economic Literature, 38(1), 77–114. https://doi.org/10.1257/jel.38.1.77 |

3.1.9. Flow-Stock Models

Main idea	The size of the irregular migrant population can be represented as its initial value (e.g., in any given year) plus cumulative inflows (visa overstayers, irregular border crossings/apprehensions) minus cumulative outflows (deportations, voluntary emigration, mortality, status changes).
Data source	Various administrative data sources including on visa overstayers, irregular border crossings/apprehensions, deportations, voluntary emigration, mortality, status changes.
Coverage / Definition	In theory, the approach has a broad coverage if everyone that arrived and did not leave adds to the existing "stock" of irregular migrants regardless of their specific sub-population.
Estimation assumptions	One key assumption is that administrative data sources can capture all in- and outflows into the country and all status changes. It is well known, for example, that emigration is notoriously under-measured (Willekens et al., 2016). In addition, the method requires a baseline size of the irregular migrant population for a specific year. Depending on how this baseline is derived, more assumptions apply.
Reliability	Assuming data sources are available consistently, the approach is reliable.
Scalability	Scalability is an issue because very few countries in the world have the breadth of administrative data sources required to capture all types of inflows and outflows.
Ethical issues	One ethical limitation is the risk of overestimating irregular migration because of applying assumptions which enjoy limited agreement among experts, such as the base stock level that is assumed by such models.

Examples	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
	Rodilitz, S., & Kaplan, E. H. (2021). Snapshot Models of Undocumented Immigration. Risk Analysis, 41(9), 1643–1661. https://doi.org/10.1111/risa.13658
	Sheldon, G. (2002). Forschungsdesign zur Bestimmung des Ausmasses und der Struktur der Schwarzarbeit in der Schweiz [Bericht zu einem Forschungsauftrag des Bundesamtes für Ausländerfragen, des Bundesamtes für Flüchtlinge und des Staatssekretariats für Wirtschaft]. Forschungstelle für Arbeitsmarkt- und Industrieökonomik, Universität Basel.
	Tapinos, G. (1999). Clandestine Immigration: Economic and Political Issues. Organization for Economic Co-operation and Development. In Trends in International Migration (1st ed., pp. 229–251). OECD. https://www.oecd.org/migration/mig/2717683.pdf
	Willekens, F., Massey, D., Raymer, J., & Beauchemin, C. (2016). International migration under the microscope. Science, 352(6288), 897–899. https://doi.org/10.1126/science.aaf6545

3.1.10.   Irregular border crossings & apprehensions

Main idea	Irregular migration flows are measured as the number of individuals detected at the border or apprehended after entry. Irregular migrants are defined in this case as individuals crossing the border or residing in a location without a valid visa, residence permit, or alternative entry permit.
Data source	Administrative data from border enforcement agencies and police.
Coverage / Definition	This approach covers all irregular migrants who enter or reside irregularly. This excludes irregular migrants who entered regularly and overstayed their visa but are unlikely to be checked after entry.
Estimation assumptions	The assumption is that detections at the border or apprehensions of irregular residents capture all irregular migrants. However, many irregular entries at the border likely go unrecorded and police checks are highly selective. Irregular migrants actively try to avoid capture on either administrative data source. In addition, some migrants may enter irregularly with legitimate reasons, and subsequently apply for

	asylum, thus receiving a legal residence status for the duration of their asylum claim assessment.
Reliability	Given that the scope of border checks and police checks vary substantially over time and location, the approach likely has low reliability. Policing and enforcement efforts depend on current events and institutional capacity.
Scalability	The approach is scalable to any country with immigration and border control data as well as police data.
Ethical issues	Ethical issues may arise if border crossing attempts are pushed back illegally and do not enter official statistics. Additional issues may arise if police checks are predominantly targeting certain communities suspected of irregularly.
Examples	FRONTEX. (2022, October 7). Risk Analysis for 2022/2023. Publications. https://frontex.europa.eu/publications/risk-analysis-for-2022-2023-RfJIVQ
	Massey, D. S., & Singer, A. (1995). New Estimates of Undocumented Mexican Migration and the Probability of Apprehension. Demography, 32(2), 203–213. https://doi.org/10.2307/2061740

3.1.11. *Database systems enabling the identification of visa overstayers*

Main idea	The main approach is to estimate the number of irregular migrants by subtracting the number of legal exits from the sum of legal entries and status changers recorded in dedicated databases. For example, the Arrival and Departure Information System (ADIS) in the US, allows for a match of departure and status change records to arrival records collected during the admission process of passengers coming through air, land (i.e., excluding irregular land crossings), or sea traffic. Visa overstayers are thus calculated as those who entered legally but have not left the country - at least not known to the authorities - and have not obtained another legal status. In part, the approach could be based, to some extent, on a fusion of biometric and biographic verification of travellers at entry and exit to detect those who overstay.
Data source	Administrative data sources, particularly entry/ exit visas and number of residence-permits.
Coverage / Definition	This approach covers legal entrants to a country and is usually linked to a travel authorisation system recording travellers and recording them as departed upon exit. Travel authorisation systems are usually

also linked to visa and other databases. In the European context, the planned entry-exit system/ ETIAS will be limited to short term visitors and will exclude persons with longer term national visas or residence permits. This indicator does not capture irregular migrants who have entered the country irregularly without a visa (for example, by land or sea).

Estimation assumptions	One of the main assumptions is that exits, or emigration is recorded accurately. However, in case of emigration over land borders emigration could remain undercounted which would falsely inflate the size of the visa overstayers. Importantly, a perfect record linkage between entries and exits must be in place.
Reliability	If data sources are measured consistently, this approach has high reliability.
Scalability	This approach has limited scalability because many countries cannot link entry data with residence permits and exit data. In the EU, efforts are under way to create an entry-exit system recording visa-free travellers. Like in the US, such a system could be used to estimate overstayers.
Ethical issues	Ethical issues may result when status adjustments are not recorded accurately, and regular migrants are recorded as visa overstayers. This issue applies in countries where various administrative sources are not linked.
Examples	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
	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
	Morita, K., & Sassen, S. (1994). The New Illegal Immigration in Japan, 1980-1992. International Migration Review, 28(1), 153. https://doi.org/10.2307/2547030
	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
	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

3.2. Innovative approaches

In this section, we introduce innovative approaches to estimate irregular migration. These approaches are innovative in that they employ innovative data, innovative methods or both.

3.2.1. Using consular registers of migrant communities in combination with other data-sources

Main idea	This approach is based on the counts of the number of cards issues by consulates in destination countries in combination with population survey data. For example, Mexican consulates register nationals in the U.S. and issue cards by local areas. The consulate counts are used to estimate the total population of resident Mexican nationals. The number of "irregular" migrants from Mexico is then estimated by applying the percentage of "undocumented" migrants based on the American Community Survey.
Data source	Consular registry (e.g., Mexico's Matrícula Consular Program by the Instituto de los Mexicanos en el Exterior)
Coverage / Definition	Only irregular migrants who register directly with the consulate
Estimation assumptions	The main assumption is that most irregular migrants will register for this card. The actual people who register, however, are not a random cross-section of all undocumented migrants. The people more likely to register are better off than undocumented migrants in more unstable conditions. In fact, in the case of the U.S., the number of people registered here is a very low number compared to more traditional estimates of Mexican irregular migrants obtained by the residual method. Furthermore, extrapolating this number to other migrant groups from Latin America or the rest of the world assumes their behaviour is comparable. Applicants for a card must document their nationality, their identity, and current place of residence, so migrants without documents are unable to register through this system. Cards are valid for five years and can be renewed. All applicants are undocumented given that no other benefits exist for documented migrants.
Reliability	Research has found little reliability when this method is used to estimate the numbers of irregular migrants from various countries. For foreign-born from countries outside of Latin America and the Caribbean the predictions employing the association between card holders and ACS estimates of migrant stocks is not reliable. The estimates for the Mexican

group are 35% higher than the estimated number employing the residual method.

Scalability	If such a register system exists for other nationalities in other destination countries, with respect, for example, to voting systems, it might be possible to apply a similar approach.
Ethical issues	Data from such consular programs could be used to initiate law enforcement actions against irregular migrants if used by authorities. The method might be used to also locate specific areas where irregular migrants are concentrated. In addition, consular programmes may strain diplomatic relations between the origin and destination country.
Examples	Bhandari, 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 Massey, D. S., Rugh, J. S., & Pren, K. A. (2010). The Geography of Undocumented Mexican Migration. Mexican Studies/Estudios Mexicanos, 26(1), 129–152. https://doi.org/10.1525/msem.2010.26.1.129

3.2.2.   Drivers license data

Main idea	A comparison of the number of drivers licenses before and after law changes that grant irregular migrants access to drivers licenses (e.g., the AB60 reform in California, US).
Data source	A county's share of drivers license that were issued to irregular immigrants.
Coverage / Definition	The approach covers only those irregular migrants who drive cars and who would apply for drivers licenses when eligible.
Estimation assumptions	This approach assumes that eligible irregular migrants apply for a drivers license.
Reliability	This is a reliable method in the specific context of states with the US that grant these licenses. Since such reforms are likely to be adopted only once, repeated estimates in one country are not possible.
Scalability	This approach cannot be scaled up to other contexts that lack such laws and settings.
Ethical issues	Access to a drivers license and a car are important prerequisites for obtaining a job in the US. The use of driving license data to estimate irregular migration might discourage migrants from applying for such

	forms of documentation if there is risk in being identified. As reviewed below in section 3.2 for the approach "Biometrics and surveillance", in some states within the US, immigration authorities are making use of such data to identify irregular migrants and prosecute them.
Examples	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

3.2.3.   *Institutional registers: College enrolment, prison population, human trafficking*

Main idea	Various institutional registers include information on legal status as this information may be relevant for these institutions' mandates. Examples include universities, prisons, and official registers of victims of human trafficking. In the US context, universities capture information on the legal status of the migrant student body. This is done through the admission process. Students who do not present any documentation on legal status are assumed to be irregular migrants, often receiving advice on how to legalize their situation. Analogously, data on the prison population also contains information of migrants who committed a crime. Whether the migrant is found to be irregular or has a legal status is often determined, though not necessarily available for research. A direct count of the number of students or the prison population is possible based on this type of institutional register data. In some countries, non-governmental organizations supporting victims of human trafficking in human beings[2] provide data on the number of assisted individuals. This data has been used to provide stock estimates. One famous example is the organization Polaris in the US (DiRienzo, 2022).
Data source	Register data from colleges and universities (e.g., in the US, see CUNY) or statistics of criminality by legal status (e.g., in Texas). In some countries, data on trafficking victims are collected by non-governmental organisation data source capturing the number of persons who have been assisted or have been reported as victims of human trafficking.
Coverage / Definition	Only the population that establishes some contact with those institutions, university applicants, for example, is covered. In combination with the multiplier method, however, this information could

---

[2] According to Article 3 (1) of the Protocol to the Convention against Transnational Organised Crime trafficking in human beings is defined as "the recruitment, transportation, transfer, harbouring or receipt of persons, by means of the threat or use of force or other forms of coercion, of abduction, of fraud, of deception, of the abuse of power or of a position of vulnerability or of the giving or receiving of payments or benefits to achieve the consent of a person having control over another person, for the purpose of exploitation."

help in estimating the number of irregular migrants but only for a relatively selective group of migrants (i.e., young and educated group of irregular migrants, or those more likely to commit crimes, or being victims of human trafficking such as those involved in prostitution or sex work).

Estimation assumptions	The approach assumes that irregular migrants either disclose their status in their university application, for example, or that their status can be inferred based on omission of information on their status (i.e., when no other documentation is provided). Equally important, for the case of human trafficking data the main assumption would be that reported or detected victims of human trafficking represent an accurate estimate of the total population of this subgroup. This is not the case given the illegal and sensitive nature of human trafficking. The approach assumes, for example, that victims who find themselves in a situation of trafficking are aware of it and would seek support from non-governmental organizations.
Reliability	Data on the legal status of college students seems quite reliable since enrolment is based on complete records that are recorded consistently over time; similarly for registers of the prison population. However, estimates of the stock of human trafficking victims likely suffers from low reliability because the data relies on persons reporting their own situation or by detection from enforcement agencies – both of which are likely to vary over time and location.
Scalability	Other educational systems might not make it possible for irregular migrants to enrol at universities. Statistics on the prison population, crime statistics or other comparable institutional registers may be reliable sources for identifying migrants in an irregular situation, and in combination with the multiplier or capture/re-capture approaches could serve as a basis for broader population estimates.
Ethical issues	Data protection issues apply to the group of college students seeking higher education and employing such data for estimates might deter people with an irregular status to seek education. Depending on the level at which data on human trafficking is reported, data could be misused to reveal the identity of individuals who have fallen victim to traffickers. This may expose them to harm by the traffickers.
Examples	Bilger, V., Borkert, M., Cibea, A., Dzhengozova, M., Hollomey, C., & Rogoz, M. (2010). Study on the assessment of the extent of different types of human trafficking in EU countries. https://doi.org/10.13140/RG.2.2.35994.72647
	DiRienzo, C. E. (2022). Human Trafficking: What the New IOM Dataset Reveals. Journal of Human Trafficking, 8(3), 294–308. https://doi.org/10.1080/23322705.2020.1808383

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

Light, M. T., He, J., & Robey, J. P. (2020). Comparing crime rates between undocumented immigrants, legal immigrants, and native-born US citizens in Texas. Proceedings of the National Academy of Sciences, 117(51), 32340–32347. https://doi.org/10.1073/pnas.2014704117

3.2.4.   Missing, disappeared or deceased migrants

Main idea	The main approach is to record persons who have gone missing, who disappeared or who deceased during their attempts to reach another country irregularly.
Data source	Cumulated data from media reports, government reports, NGO reports, released, for example, by projects such as UNITED for Intercultural Action, the Migrant Files, and IOM's Missing Migrants Project.
Coverage / Definition	This approach only covers a small sub-group of irregular migrants, i.e., those *en-route* who were recorded as missing, disappeared or deceased. It does not include those who disappeared but were not recorded and those who successfully migrated but were not recorded.
Estimation assumptions	The count is based on reported cases and thus assumes limited underreporting. However, it is reasonable to assume that most migrant deaths and disappearance (for example, in the Sahara Desert or the US-Mexico border) are never reported. Some government and media reports estimate the number of deceased or disappeared migrants based on guesses from eyewitnesses who travelled with a larger group. Estimates of group sizes in these situations may be inaccurate.
Reliability	The data relies heavily on reporting from multiple sets of changing sources. Studies have shown that heightened enforcement has been associated with higher amounts of recorded deaths, likely because of increased surveillance.
Scalability	Scalability is an issue as data collection relies on multiple state- and non-state actors. State actors may change operations or refrain from surveillance due to political reasons. Non-state actors often lack resources and access to achieve broad coverage.
Ethical issues	Ethical issues arise because data collection often relies on institutions with vested interests. Government agencies may have incentives to

	underreport and reduce surveillance for political reasons, whereas non-governmental agencies might have incentives to overreport to raise awareness and funding.
Examples	Black, J. (2020). "No one talks about what it's really like" – risks faced by migrants in the Sahara Desert. In Migration in West and North Africa and across the Mediterranean (pp. 49–161). IOM. Retrieved 28 September 2023, from https://publications.iom.int/books/migration-west-and-north-africa-and-across-mediterranean-chapter-12
	Brian, T., & Laczko, F. (Eds.). (2014). Fatal journeys: tracking lives lost during migration. International Organization for Migration (IOM).
	Carling, J. (2007). Migration Control and Migrant Fatalities at the Spanish-African Borders. International Migration Review, 41(2), 316–343. https://doi.org/10.1111/j.1747-7379.2007.00070.x
	Eschbach, K., Hagan, J., Rodriguez, N., Hernández-León, R., & Bailey, S. (1999). Death at the Border. International Migration Review, 33(2), 430–454. https://doi.org/10.1177/019791839903300206
	Garcia Borja, A., & Black, J. (2022). Measuring migrant deaths and disappearances. 66. https://www.fmreview.org/issue66/garciaborja-black

3.2.5. Statistical imputation

Main idea	In the imputation approach, information from a given source that captures the legal status of migrants through direct questions is used to build a predictive model that serves to predict irregular status for respondents in another larger survey that is presumed to be a probabilistic sample of the country, including irregular migrants, but that did not include specific survey questions allowing for their identification. There are various kinds of predictive models used (e.g., logical edits, logistic regression, or machine learning applications.)
Data source	Two large, independent surveys, one containing direct questions on regular residence status, such as the Survey of Income and Program Participation (SIPP), and another larger survey, such as the American Community Survey (ACS) where the legal status of migrants can be imputed, and population sizes estimated.
Coverage / Definition	The coverage is based on how many immigrants the larger survey captures. For example, the ACS captures immigrants without complete information - did not report a legal status, did not have the US citizenship and were foreign born in one survey - that could not be logically or statistically assigned to a legal status. These migrants are matched with a donor sample member where the legal status is directly observed (e.g., SIPP).

Estimation assumptions	Imputation of missing answers assumes that missingness occurs at random and can be explained by observable factors. In addition, the approach assumes unbiased sampling and no response bias to items on legal status from respondents.
Reliability	The imputation model plays a critical role here. Different models will yield different estimates, but this can be examined. Lack of overlap between the donor sample – one where irregular status is directly measured – and the larger survey, in terms of the variables allowing for the identification of residence status, as well as potential for response biases in any of the two surveys, can also affect the reliability of the estimates. This approach does not take into consideration non sampling error from item nonresponse; there is no way of knowing whether there have been changes in the undocumented population with this method. The method assumes that the larger sample (e.g., the ACS) is a probabilistic sample of irregular migrants.
Scalability	The scalability of this approach is limited due to the absence of a donor sample where immigrants are identified according to their residence status.
Ethical issues	Respondents to the ACS do not consent to their information being used for this purpose. The government could use the imputation exercise to prosecute migrants.
Examples	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
	Capps, R., Bachmeier, J. D., & Van Hook, J. (2018). Estimating the Characteristics of Unauthorized Immigrants Using U.S. Census Data: Combined Sample Multiple Imputation. The ANNALS of the American Academy of Political and Social Science, 677(1), 165–179. https://doi.org/10.1177/0002716218767383
	Mira, A. F., & Bollinger, C. (2023). Credible Interval Estimates of the Size and Legal Composition of the US Foreign-Born Population. ASSA Annual Meeting. https://www.aeaweb.org/conference/2023/program/powerpoint/AyE4 aKyr
	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

3.2.6. Retrospective surveys

Main idea	This method is based on the concept of direct questions in surveys (see section 3.1 for the method "Self-identification in surveys"), however, the survey targets legal migrants who report on their previous residence statuses retrospectively.
Data source	Sample surveys with migrants.
Coverage / Definition	Coverage is only limited by the target population in the respective survey.
Estimation assumptions	The same assumptions apply which applied to section 3.1 for the method "Self-identification in surveys", employing direct survey questions. This approach, however, makes two additional assumptions: First, it assumes that persons can correctly recall previous legal residence statuses. Second, it assumes that legal migrants provide truthful answers about previous irregular status in the past.
Reliability	The approach is reliable if features of the survey design remain consistent across locations and time.
Scalability	Depending on resources and available sampling frames, this method could be scaled up to any country to measure irregular migration retrospectively.
Ethical issues	Ethical issues may apply if the government agencies obtain access to individual-level data and revoke the legal status of migrants based on their survey responses regarding past spells in an irregular status. In addition, certain migrant groups may face stigmatization as their previous irregular status becomes known.
Examples	Larson, E. M., & Droitcour, J. A. (2012). The Grouped Answer Method for Estimating Immigration Status: Analysis of Data from the 2004 General Social Survey. In N. Hoque & D. A. Swanson (Eds.), Opportunities and Challenges for Applied Demography in the 21st Century (pp. 311–334). Springer Netherlands. https://doi.org/10.1007/978-94-007-2297-2_17
	Liu, M.-M. (2015). Becoming a Man: Legal status, Networks and Male Migration between Senegal and Europe (No. 38; MAFE Working Papers). INED. https://mafeproject.site.ined.fr/fichier/rte/29/wp%2038%20bis.pdf
	Medová, L., & Drbohlav, D. (2013). ESTIMATING THE SIZE OF THE IRREGULAR MIGRANT POPULATION IN PRAGUE - AN ALTERNATIVE APPROACH: ESTIMATING THE MIGRANT POPULATION IN PRAGUE. Tijdschrift Voor Economische En Sociale Geografie, 104(1), 75–89. https://doi.org/10.1111/j.1467-9663.2012.00737.x
	Molinari, R., Impicciatore, R., & Ortensi, L. E. (2023). Traces in the shadow: Occupational outcomes of previously undocumented migrants

in Italy. International Migration, imig.13144.
https://doi.org/10.1111/imig.13144

3.2.7.   Alternative residual: labour force survey and social security register

Main idea	Comparing estimates based on multiple data sources: a registry and survey. For example, microdata from labour force surveys and official social security registers can provide an estimation of the sectoral composition of the immigrant workers. This comparison can reveal higher numbers of people employed in one economic sector, as per the survey, in comparison to officially registered immigrant employees, which could hint at the presence of irregular migrant workers in specific economic sectors.
Data source	The Labour Force Survey (LFS) and the Social Security Administrative Register
Coverage / Definition	Only employed irregular migrants are captured but it is not possible to distinguish different types of irregular migrants.
Estimation assumptions	This method assumes that labour force surveys do constitute random samples capturing enough of the irregular migrant population, which may not be the case given that many sampling frames exclude irregular migrants. Furthermore, registers are more reliable for some economic sectors than others, depending on how much informal labour is present in them.
Reliability	This approach would have a high reliability depending on whether there are substantial changes in the sampling frame of the LFS. Equally so, changes in the register of social security payments might affect the reliability of this approach. However, validity will depend on the comparability between the two data sources' design and scope.
Scalability	This method can be scaled up to other countries where the LFS is used and in combination with the registers of workers paying social security.
Ethical issues	Once sectors with high prevalence of irregular migration are identified, this information could lead targeted reinforcements in policing and therefore prosecution of individuals. If the estimates are inaccurate, this could lead to erroneous persecution of legal migrants.
Examples	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

3.2.8.   *Consumption data*

Main idea	In the consumption data-based estimation approaches, the first step is to estimate the expected relationship between food consumption (e.g., rice or any other type of high frequency and almost universally consumed good in each country) and economic development levels over time. In a second step, the difference between actual (recorded) rice consumption, for example, and that predicted by the rice-development relationship and total official population can yield a residual quantity estimate of rice consumption attributable to the irregular population. In the final third step, this residual quantity is converted into irregular person estimates, using various alternative assumptions for undocumented persons' consumption patterns.
Data source	Consumption per capita, census counts, and some country-level adjustment variables and price data.
Coverage / Definition	All groups of irregular migrants are covered by the consumption data, but this approach would not allow for a distinction of the different groups of irregular migrants.
Estimation assumptions	The consumption goods or items must be consumed by everyone in the respective country, such that a consumption unit per capital estimate is plausible. Consumption of the item must be measured reliably. Additionally, the census is of a good quality such that any undercount results from people unwilling to be identified and not from other failures in the statistical process.
Reliability	Consumption data is highly reliable. The reliability of the approach, however, will depend on measures of economic development and population size estimations that are used to obtain expected levels of consumption. Moreover, depending on the context, other types of consumption could be more useful. For example, water or electricity consumption or even waste production might be more appropriate in Europe.
Scalability	The approach can be applied in other countries employing other types of consumption.
Ethical issues	Ethical issues in this case are less pronounced because the approach does not allow to identify individuals.
Examples	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

3.2.9.   Biometrics & surveillance

Main idea	This approach is based on the use of socio-technical systems based on artificial intelligence that enable facial recognition in combination with biometric information in official data sources which, when used in combination, can allow for the identification of irregular migrants, as individuals without required documentation do not appear registered in biometric government data bases. For this approach to serve the estimation of irregular migrants, the method should be applied at a larger scale, looking not for specific individuals but extending it to the complete database.
Data source	In the US context, the Department of Motor Vehicles (DMV) data base, in combination with, official biometric data from the Department of Homeland Security (DHS). Also, Amazon company 'Ring Doorbell' video service and facial recognition software *'Rekognition'* could be used for this, though currently this company only partners with police departments for other law enforcement purposes.
Coverage / Definition	Covers all potential irregular immigrants with a driving license who drive and can be stopped by the police in over a dozen states within the US.
Estimation assumptions	This is not a method to per se estimate the stock of irregular migrants, but in principle generates data that allows for the identification at a large scale of migrants without a regular status, or could generate enough data for such estimations, in combination with, for example, the multiplier method. The DMV database has more current and reliable information on migrants than the DHS.
Reliability	Well documented biases in facial recognition software can hamper the accuracy and hence reliability of identification of certain groups that differ from the standard White male (Raji et al., 2022), but we do not know of any specific study of such biases on DMV data.
Scalability	The DHS and its Immigration and Customs Enforcement, ICE, office can scale such procedures. The information contained in the DMV data base could be in principle scaled up in the states within the US that grant driver licenses to undocumented immigrants, or other digital data bases which contain facial images (e.g., Instagram).
Ethical issues	Violations of irregular migrants' right of privacy. Use of race-ethnicity as the primary means of targeting and identification of individuals. There are no authorizations to mine the DMV data for identifying irregular migrants.
Examples	Knutson, A. (2021). Saving Face; The Unconstitutional Use of Facial Recognition on Undocumented Immigrants and Solutions in IP. IP Theory, 10(1). https://www.repository.law.indiana.edu/ipt/vol10/iss1/2

Molnar, P. (2019). Technology on the margins: AI and global migration management from a human rights perspective. Cambridge International Law Journal, 8(2), 305–330. https://doi.org/10.4337/cilj.2019.02.07

3.2.10.   Online search behaviour

Main idea	Online search data (e.g., Google trends data) can be used to estimate change in the volume of people who search "will I be deported". Changes over time following a specific event are analysed to estimate the effect of a migration policy change (e.g., the "Muslim ban") be used to estimate irregular migrants.
Data source	Online search data (e.g., google trends data on the frequency of searches for the query "will I be deported")
Coverage / Definition	All migrants who input the same query into Google's search engine, which, depending on the type of query, could relate to mostly Muslim migrants, as in the case of the so-called 'Muslim ban' in the US, or to other groups. Therefore, coverage is rather relative and highly uncertain if migrants with a defined legal status – as well as other non-migrants - make the query. Equally important is that the query and the underlying population are restricted to the country where law changes affecting migrants are taking place.
Estimation assumptions	The main assumption is that only irregular migrants fear deportation, which is not necessarily the case (Waldinger et al., 2023). Relatives and friends of irregular migrants could also fear the deportation of their close ones and then search for the same query.
Reliability	There is no reference point in time where searches for the term could be used as benchmark (e.g., a period without searches for this term). The reliability of the method is low if no anchoring population can be established.
Scalability	This method could be replicated in other countries where similar restrictive migration legal policy changes are implemented.
Ethical issues	Google trend searchers can be narrowed down to smaller regional units, e.g., States, Metropolitan regions, and even cities. Such data could be used to target irregular migrants living in those areas through increasing policing.
Examples	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

Chykina, V., & Crabtree, C. (2018). Using Google Trends to Measure Issue Salience for Hard-to-Survey Populations. Socius: Sociological Research for a Dynamic World, 4, 237802311876041. https://doi.org/10.1177/2378023118760414

4. DISCUSSION

In this section, we briefly summarize the broader development of approaches to estimate irregular migration, and comment on future trends according to the evaluation criteria which we applied throughout the report. In comparison to the number of approaches reviewed in Jandl, 2011, important advances were made in estimation methodologies to overcome limitations or drawbacks of traditional approaches, such as the residual method. These advances were possible thanks to the use of innovative methods, as well as the availability of more data sources (Vespe et al., 2017). Despite the progress made in estimating irregular migration in the last decade, various challenges remain. Approaches remain highly fragmented depending on the type of data source used for the estimation. Some approaches can only be applied in specific countries due to data limitations.

In the following, we comment on each of the items we reviewed for each approach.

Data sources	The data sources used for estimating irregular migration become ever more diverse including institutional data sources, consumption data, and digital data sources. We also observe an increasing trend of linking various data sources. So far, we have found few attempts at employing data from social networks, such as Facebook or Twitter, or mobile phone data or flight passenger data. We see a potential for the use of such data sources.
Coverage/ Definition	Coverage of various sub-groups of irregular migrants remains a challenge. Most approaches using population surveys or censuses aim to capture all irregular migrants. Approaches using administrative, economic, and digital data sources are usually bound to a restricted sub-group of irregular migrants. Digital sources such as google searches or social media data have even greater challenges of delineating specific groups.
Estimation assumptions	As outlined in this report, all approaches rely on a set of – often highly problematic - assumptions. While the assumptions may vary in number and scope, it is difficult to provide any assessment of how plausible these

assumptions are because benchmarks are missing. As a tentative rule of thumb, we suggest that estimation assumptions become more challenging with every additional data source that is used to produce an estimate, since each data source carries its own limitations. Key assumptions in the estimation procedure are often linked to how data has been collected. The more sources are used, the more challenging it is to defend that all assumptions are met.

Reliability

Reliability of estimates depends on the consistency of required data sources and definitions over time. Approaches using administrative data sources, which do not depend on sampling, can yield reliable results of applied consistently using the same method and sources. Approaches involving survey data, for example, face additional challenges as the population to sample from is constantly changing.

It is difficult to evaluate whether approaches have become reliable in the past 15 years because of the absence of any benchmark that new estimates could be compared to. There remains a lack of methodological work comparing various approaches in the same context to assess how results vary in accuracy and uncertainty. There is large potential for a simulation-based approach to evaluating the different methodologies, which could help in highlighting the advantages and limitations of existing approaches.

Scalability

Scalability remains an important challenge because approaches heavily depend on the availability of specific data sources, which may exist in some countries and be completely absent or only partially available in others. For example, the use of driver's license data might not be applicable in Europe, but similar data coming from initiatives like the 'Zuri-Card' for Zurich inhabitants could allow for a similar estimation. In addition, countries usually apply different definitions derived from legislation upon which the administrative data source is derived. Much of the literature is based on the US context due to the broad availability of administrative data sources, surveys and few hurdles in terms of data privacy and data protection.

Surveys are more flexible in nature and could be applied across various countries. Sampling frames are often not available or differ across countries. More cross-country research applying alternative survey methods which do not rely on standard sampling frames is needed. Certain approaches reviewed in this report rely on specific reforms or one-off regularizations which often do not scale to other contexts. The use of digital data could have the potential for a higher degree of comparability across countries, for example, employing Google Search trend data or Facebook. The data is collected globally and applied consistent definitions.

Ethical issues	This review focused on ethical issues related to the specific application of a given methodological approach for researchers, not ethical and political issues of producing and using estimates in general (Düvell et al., 2010; Van Liempt & Bilger, 2012). Ethical issues relating to methodological approaches commonly relate to issues of data protection, data privacy and the do-no-harm principle (Salah et al., 2022). The use of administrative data sources and digital data raises questions regarding missing consent by individuals and linking individual records without is illegal in countries with strong data privacy provisions. Approaches that yield broader estimates of population sizes are generally less problematic than approaches which directly identify individual migrants. Identifying personal information can be used to prosecute irregular migrants. In this context, increasing use of digital data by authorities for border control and surveillance is problematic (Broeders, 2007; Newel et al., 2017; Pötzschke & Rinken, 2022; Rampazzo et al., 2023). Guidelines for ethical collection and use of data sources are readily available (Belmont Report 1979; Dittrich, Kenneally, and others 2011). Non-government actors such as researchers need to consult these guidelines and ask for approval from ethical review boards before estimating irregular migration (Cyrus 2023).

The approaches to estimating irregular migration have become more diverse and more complex. However, since the seminal CLANDESTINO project in 2008, the field has not moved closer to a "gold standard". There remains a lack of methodological comparative research. Simulation studies could be one way forward in this regard.

Beyond the methods which were reviewed in this report, there are additional approaches which have been used to estimate "hard-to-reach" populations which could be applied to irregular migration in the future. Particularly epidemiological research (Wesson et al., 2017; Xu et al., 2022) focusing on studying small populations (e.g., population with HIV) has developed new approaches which have yet to be applied to the case of irregular migration. Similarly, the emergence of digital data will likely increase the diversity of approaches. For example, scholars have demonstrated how social media data and flight passenger data can be used to measure international mobility and migration (Tjaden, 2021; Zagheni et al., 2017). However, these approaches have not yet been used to estimate irregular migration.

References

Acostamadiedo, E., Sohst, R. R., Tjaden, J., Groenewold, G., & de Valk, H. (2020). *Assessing Immigration Scenarios for the European Union in 2030: Relevant, Realistic and Reliable?* International Organization for Migration. https://publications.iom.int/books/assessing-immigration-scenarios-european-union-2030

Ambrosini, M., & Hajer, M. H. J. (2023). *Irregular Migration: IMISCOE Short Reader*. Springer International Publishing. https://doi.org/10.1007/978-3-031-30838-3

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. https://doi.org/10.1111/imre.12059

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. https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/centre-sampling-technique-in-foreign-migration-surveys-a-methodological-note.pdf

Baker, B. (2021). *Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2015–January 2018* (Population Statistics). Department of Homeland Security's Office of Immigration Statistics. https://www.dhs.gov/sites/default/files/publications/immigration-statistics/Pop_Estimate/UnauthImmigrant/unauthorized_immigrant_population_estimates_2015_-_2018.pdf

Baldwin-Edwards, M., & Kraler, A. (2009). *REGINE Regularisations in Europe Study on practices in the area of regularisation of illegally staying third-country nationals in the Member States of the EU*. International Centre for Migration Policy Development (ICMPD). https://ec.europa.eu/migrant-integration/sites/default/files/2009-04/docl_8193_345982803.pdf

Bhandari, 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

Bilger, V., Borkert, M., Cibea, A., Dzhengozova, M., Hollomey, C., & Rogoz, M. (2010). *Study on the assessment of the extent of different types of human trafficking in EU countries*. https://doi.org/10.13140/RG.2.2.35994.72647

Bird, S. M., & King, R. (2018). Multiple Systems Estimation (or Capture-Recapture Estimation) to Inform Public Policy. *Annual Review of Statistics and Its Application*, *5*(1), 95–118. https://doi.org/10.1146/annurev-statistics-031017-100641

Black, J. (2020). "No one talks about what it's really like" – risks faced by migrants in the Sahara Desert. In *Migration in West and North Africa and across the Mediterranean* (pp. 49–161). IOM. Retrieved 28 September 2023, from

https://publications.iom.int/books/migration-west-and-north-africa-and-across-mediterranean-chapter-12

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

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

Brian, T., & Laczko, F. (Eds.). (2014). *Fatal journeys: tracking lives lost during migration*. International Organization for Migration (IOM).

Broeders, D. (2007). The New Digital Borders of Europe: EU Databases and the Surveillance of Irregular Migrants. *International Sociology*, *22*(1), 71–92. https://doi.org/10.1177/0268580907070126

Burgers, J. (1996). Natte vingers en vuile handen; over het schatten van het aantal illegale vreemdelingen. Een reactie op Böcker en Groenendijk. *Migrantenstudies*, *12*(1), 14–26.

Capps, R., Bachmeier, J. D., & Van Hook, J. (2018). Estimating the Characteristics of Unauthorized Immigrants Using U.S. Census Data: Combined Sample Multiple Imputation. *The ANNALS of the American Academy of Political and Social Science*, *677*(1), 165–179. https://doi.org/10.1177/0002716218767383

Carling, J. (2007). Migration Control and Migrant Fatalities at the Spanish-African Borders. *International Migration Review*, *41*(2), 316–343. https://doi.org/10.1111/j.1747-7379.2007.00070.x

Chykina, V., & Crabtree, C. (2018). Using Google Trends to Measure Issue Salience for Hard-to-Survey Populations. *Socius: Sociological Research for a Dynamic World*, *4*, 237802311876041. https://doi.org/10.1177/2378023118760414

Cyrus, N. (2023). Ethical Benchmarking for the Measurement of Irregular Migration, MIrreM Working Paper No. 03. Krems: University for Continuing Education Krems (Danube University Krems). DOI: 10.5281/zenodo.8389189.

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

DiRienzo, C. E. (2022). Human Trafficking: What the New IOM Dataset Reveals. *Journal of Human Trafficking*, *8*(3), 294–308. https://doi.org/10.1080/23322705.2020.1808383

Drbohlav, D., Dzúrová, D., Čermák, Z., Janská, E., Čermáková, D., & Medová, L. (2008). Immigrants' irregular economic activities in the Czech Republic (a multi-approach study). *Transfer: European Review of Labour and Research*, *14*(4), 639–652. https://doi.org/10.1177/102425890801400409

Drbohlav, D., & Lachmanová, L. (2023, September 19). *Undocumented migration. Counting the Uncountable. Country report - The Czech Republic | European Website on Integration*. https://ec.europa.eu/migrant-integration/library-

document/undocumented-migration-counting-uncountable-country-report-czech-republic_en

Düvell, F., Triandafyllidou, A., & Vollmer, B. (2010). Ethical issues in irregular migration research in Europe. *Population, Space and Place*, *16*(3), 227–239. https://doi.org/10.1002/psp.590

Eschbach, K., Hagan, J., Rodriguez, N., Hernández-León, R., & Bailey, S. (1999). Death at the Border. *International Migration Review*, *33*(2), 430–454. https://doi.org/10.1177/019791839903300206

Espenshade, T. J. (1995a). Using ins Border Apprehension Data to Measure the Flow of Undocumented Migrants Crossing the U.S.-Mexico Frontier. *International Migration Review*, *29*(2), 545–565. https://doi.org/10.1177/019791839502900209

Espenshade, T. J. (1995b). Using ins Border Apprehension Data to Measure the Flow of Undocumented Migrants Crossing the U.S.-Mexico Frontier. *International Migration Review*, *29*(2), 545–565. https://doi.org/10.1177/019791839502900209

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

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

Garcia Borja, A., & Black, J. (2022). *Measuring migrant deaths and disappearances*. *66*. https://www.fmreview.org/issue66/garciaborja-black

Gelman, A., & Hennig, C. (2017). Beyond Subjective and Objective in Statistics. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *180*(4), 967–1033. https://doi.org/10.1111/rssa.12276

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

Gonzales, R. G. (2011). Learning to Be Illegal: Undocumented Youth and Shifting Legal Contexts in the Transition to Adulthood. *American Sociological Review*, *76*(4), 602–619. https://doi.org/10.1177/0003122411411901

Grossmann, I., Rotella, A., Hutcherson, C. A., Sharpinskyi, K., Varnum, M. E. W., Achter, S., Dhami, M. K., Guo, X. E., Kara-Yakoubian, M., Mandel, D. R., Raes, L., Tay, L., Vie, A., Wagner, L., Adamkovic, M., Arami, A., Arriaga, P., Bandara, K., Baník, G., … The Forecasting Collaborative. (2023). Insights into the accuracy of social scientists' forecasts of societal change. *Nature Human Behaviour*, *7*(4), 484–501. https://doi.org/10.1038/s41562-022-01517-1

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

Heckathorn, D. D., & Cameron, C. J. (2017). Network Sampling: From Snowball and Multiplicity to Respondent-Driven Sampling. *Annual Review of Sociology, 43*(1), 101–119. https://doi.org/10.1146/annurev-soc-060116-053556

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

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

Huddleston, T., & Tjaden, J. (2012). *Immigrant Citizens Survey. How immigrants experience integration in 15 European cities*. King Baudouin Foundation and MPI. https://www.om.acm.gov.pt/documents/58428/296070/immigrant+citizens+survey.pdf/a3e84f37-5825-48f0-b4aa-125970595cab

Jandl, M. (2009). A multiplier estimate of the illegally resident third-country national population in Austria based on crime suspect data. Database on Irregular Migration. Working Paper, (2), 2.

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

Jandl, M., Hollomey, C., & Stepien, A. (2007). *Migration and irregular work in Austria: results of a Delphi-study.* (No. 90; International Migration Papers). International Labour Organization and International Centre for Migration Policy. https://www.icmpd.org/file/download/47428/file/Migration%2520and%2520Irregular%2520Work%2520in%2520Austria-%2520Delphi%2520study.pdf

Jolly, A., Thomas, S., & Stanyer, J. (2020). *London's children and young people who are not British citizens: A profile*. https://doi.org/10.13140/RG.2.2.13839.12964

Knutson, A. (2021). Saving Face; The Unconstitutional Use of Facial Recognition on Undocumented Immigrants and Solutions in IP. *IP Theory*, *10*(1). https://www.repository.law.indiana.edu/ipt/vol10/iss1/2

Koser, K. (2010). Dimensions and dynamics of irregular migration. *Population, Space and Place*, *16*(3), 181–193. https://doi.org/10.1002/psp.587

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. (2023). *Taxonomy of migrant irregularity*. https://doi.org/10.5281/ZENODO.7875543

Larson, E. M., & Droitcour, J. A. (2012). The Grouped Answer Method for Estimating Immigration Status: Analysis of Data from the 2004 General Social Survey. In N. Hoque & D. A. Swanson (Eds.), *Opportunities and Challenges for Applied Demography in the 21st Century* (pp. 311–334). Springer Netherlands. https://doi.org/10.1007/978-94-007-2297-2_17

Light, M. T., He, J., & Robey, J. P. (2020). Comparing crime rates between undocumented immigrants, legal immigrants, and native-born US citizens in Texas. *Proceedings of the National Academy of Sciences*, *117*(51), 32340–32347. https://doi.org/10.1073/pnas.2014704117

Liu, M.-M. (2015). *Becoming a Man: Legal status, Networks and Male Migration between Senegal and Europe* (No. 38; MAFE Working Papers). INED. https://mafeproject.site.ined.fr/fichier/rte/29/wp%2038%20bis.pdf

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

Marpsat, M., & Razafindratsima, N. (2010). Survey Methods for Hard-to-Reach Populations: Introduction to the Special Issue. *Methodological Innovations Online*, *5*(2), 3.1-16. https://doi.org/10.4256/mio.2010.0014

Martí, M., & Ródenas, C. (2012). Measuring International Migration through Sample Surveys: Some Lessons from the Spanish Case. *Population*, *67*(3), 435–463. https://www.cairn-int.info/journal-population-2012-3-page-435.htm

Massey, D. S., Rugh, J. S., & Pren, K. A. (2010). The Geography of Undocumented Mexican Migration. *Mexican Studies/Estudios Mexicanos*, *26*(1), 129–152. https://doi.org/10.1525/msem.2010.26.1.129

Massey, D. S., & Singer, A. (1995). New Estimates of Undocumented Mexican Migration and the Probability of Apprehension. *Demography*, *32*(2), 203–213. https://doi.org/10.2307/2061740

McKenzie, D., & Siegel, M. (2013). Eliciting illegal migration rates through list randomization. *Migration Studies*, *1*(3), 276–291. https://doi.org/10.1093/migration/mnt018

Medová, L., & Drbohlav, D. (2013). ESTIMATING THE SIZE OF THE IRREGULAR MIGRANT POPULATION IN PRAGUE - AN ALTERNATIVE APPROACH: ESTIMATING THE MIGRANT POPULATION IN PRAGUE. *Tijdschrift Voor Economische En Sociale Geografie*, *104*(1), 75–89. https://doi.org/10.1111/j.1467-9663.2012.00737.x

Mira, A. F., & Bollinger, C. (2023). Credible Interval Estimates of the Size and Legal Composition of the US Foreign-Born Population. *ASSA Annual Meeting*. https://www.aeaweb.org/conference/2023/program/powerpoint/AyE4aKyr

Molinari, R., Impicciatore, R., & Ortensi, L. E. (2023). Traces in the shadow: Occupational outcomes of previously undocumented migrants in Italy. *International Migration*, imig.13144. https://doi.org/10.1111/imig.13144

Molnar, P. (2019). Technology on the margins: AI and global migration management from a human rights perspective. *Cambridge International Law Journal*, *8*(2), 305–330. https://doi.org/10.4337/cilj.2019.02.07

Morita, K., & Sassen, S. (1994). The New Illegal Immigration in Japan, 1980-1992. *International Migration Review*, *28*(1), 153. https://doi.org/10.2307/2547030

Newell, B. C., Gomez, R., & Guajardo, V. E. (2017). Sensors, Cameras, and the New 'Normal' in Clandestine Migration: How Undocumented Migrants Experience Surveillance at the U.S.-Mexico Border. *Surveillance & Society*, *15*(1), 21–41. https://doi.org/10.24908/ss.v15i1.5604

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

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

Passel, J. S. (2006). *The Size and Characteristics of the Unauthorized Migrant Population in the U.S. Estimates Based on the March 2005 Current Population Survey*. Pew Hispanic Center. https://www.procon.org/wp-content/uploads/sites/40/immidoc6.pdf

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

Pötzschke, S., & Rinken, S. (2022). Introduction: Using Innovative Technologies to Tackle Methodological Challenges in Migration Research. In S. Pötzschke & S. Rinken (Eds.), *Migration Research in a Digitized World* (pp. 1–11). Springer International Publishing. https://doi.org/10.1007/978-3-031-01319-5_1

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2020). Saving Face: Investigating the Ethical Concerns of Facial Recognition Auditing. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 145–151. https://doi.org/10.1145/3375627.3375820

Rampazzo, F., Rango, M., & Weber, I. (2023). New Migration Data: Challenges and Opportunities. In E. Bertoni, M. Fontana, L. Gabrielli, S. Signorelli, & M. Vespe (Eds.), *Handbook of Computational Social Science for Policy* (pp. 345–359). Springer International Publishing. https://doi.org/10.1007/978-3-031-16624-2_18

Reichel, D., & Morales, L. (2017). Surveying immigrants without sampling frames – evaluating the success of alternative field methods. *Comparative Migration Studies, 5*(1), 1. https://doi.org/10.1186/s40878-016-0044-9

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

Robinson, J. G. (1980). Estimating the Approximate Size of The Illegal Alien Population in the United States by the Comparative Trend Analysis of Age-Specific Death Rates. *Demography*, *17*(2), 159–176. https://doi.org/10.2307/2061057

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

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

Salah, A. A., Canca, C., & Bariş, E. (2022). Ethical and Legal Concerns on Data Science for Large-Scale Human Mobility. In A. A. Salah, C. Canca, & E. Bariş, *Data Science for Migration and Mobility* (pp. 24–48). British Academy. https://doi.org/10.5871/bacad/9780197267103.003.0002

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

Schneider, F., & Enste, D. H. (2000). Shadow Economies: Size, Causes, and Consequences. *Journal of Economic Literature*, *38*(1), 77–114. https://doi.org/10.1257/jel.38.1.77

Sheldon, G. (2002). *Forschungsdesign zur Bestimmung des Ausmasses und der Struktur der Schwarzarbeit in der Schweiz* [Bericht zu einem Forschungsauftrag des Bundesamtes für Ausländerfragen, des Bundesamtes für Flüchtlinge und des Staatssekretariats für Wirtschaft]. Forschungstelle für Arbeitsmarkt- und Industrieökonomik, Universität Basel.

Sunderhaus, S. (2007). Regularization programs for undocumented migrants. *Migration Letters*, *4*(1), 65–76. https://www.ceeol.com/search/article-detail?id=493672

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

Tapinos, G. (1999). Clandestine Immigration: Economic and Political Issues. Organization for Economic Co-operation and Development. In *Trends in International Migration* (1st ed., pp. 229–251). OECD. https://www.oecd.org/migration/mig/2717683.pdf

Tetlock, P. E. (2017). *Expert Political Judgment: How Good Is It? How Can We Know? - New Edition*. Princeton University Press. https://doi.org/10.1515/9781400888818

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

UNODC. (2022). *Monitoring human trafficking prevalence through multiple system estimation*. UNITED NATIONS. https://www.unodc.org/documents/data-and-analysis/tip/2022/MSE_TIP_UNODC_ENG.pdf

Van Der Heijden, P. G. M., Cruyff, M., & Van Houwelingen, H. C. (2003). Estimating the Size of a Criminal Population from Police Records Using the Truncated Poisson Regression Model. *Statistica Neerlandica*, *57*(3), 289–304. https://doi.org/10.1111/1467-9574.00232

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

Van Liempt, I., & Bilger, V. (2012). Ethical Challenges in Research with Vulnerable Migrants. In C. Vargas-Silva (Ed.), *Handbook of Research Methods in Migration*. Edward Elgar Publishing. https://doi.org/10.4337/9781781005231.00031

van Liempt, I., Schapendonk, J., & Campos-Delgado, A. (Eds.). (2023). Research Handbook on Irregular Migration. In *Research Handbook on Irregular Migration*. Edward Elgar Publishing. https://www.elgaronline.com/edcollbook/book/9781800377509/9781800377509.xml

van Meeteren, M., Engbersen, G., & van San, M. (2007). *Irreguliere immigranten in België*. https://repub.eur.nl/pub/12074

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

Waldinger, R., Hoffmann, N. I., & Lai, T. (2023). Differentiated legality: understanding the sources of immigrants' deportation fear. *Ethnic and Racial Studies*, 1–24. https://doi.org/10.1080/01419870.2023.2235419

Wanner, P. (2002). *Compter les clandestins. Méthodes d'estimation de la population sans autorisation de séjour à partir des données sur la population* (No. 13; Discussion Paper). Swiss Forum for Migration Studies. http://www.unine.ch/files/live/sites/sfm/files/shared/pub/dp/dp13.pdf

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

Warren, R. (2021). In 2019, the US Undocumented Population Continued a Decade-Long Decline and the Foreign-Born Population Neared Zero Growth. *Journal on Migration and Human Security*, *9*(1), 31–43. https://doi.org/10.1177/2331502421993746

Wesson, P., Reingold, A., & McFarland, W. (2017). Theoretical and Empirical Comparisons of Methods to Estimate the Size of Hard-to-Reach Populations: A Systematic Review. *AIDS and Behavior*, *21*(7), 2188–2206. https://doi.org/10.1007/s10461-017-1678-9

Willekens, F., Massey, D., Raymer, J., & Beauchemin, C. (2016). International migration under the microscope. *Science*, *352*(6288), 897–899. https://doi.org/10.1126/science.aaf6545

Xu, C., Jing, F., Lu, Y., Ni, Y., Tucker, J., Wu, D., Zhou, Y., Ong, J., Zhang, Q., & Tang, W. (2022). Summarizing methods for estimating population size for key populations: a global

scoping review for human immunodeficiency virus research. *AIDS Research and Therapy*, *19*(1), 9. https://doi.org/10.1186/s12981-022-00434-7

Zagheni, E., Weber, I., & Gummadi, K. (2017). Leveraging Facebook's Advertising Platform to Monitor Stocks of Migrants: Leveraging Facebook's Advertising Platform. *Population and Development Review*, *43*(4), 721–734. https://doi.org/10.1111/padr.12102

Zhang, L.-C. (2008, June 23). *Developing methods for determining the number of unauthorized foreigners in Norway*. Ssb.No. https://www.ssb.no/en/befolkning/artikler-og-publikasjoner/developing-methods-for-determining-the-number-of-unauthorized-foreigners-in-norway--18906

ABOUT THE AUTHORS

Alejandra Rodríguez Sánchez is a postdoc at the University of Potsdam, Potsdam, Germany. Jasper Tjaden is Professor for Applied Social Research and Public Policy at the Economic and Social Science Department of the University of Potsdam, Potsdam, Germany.

COPYRIGHT CLAUSE

THE MIRREM CONSORTIUM

University for Continuing Education Krems (Coordinator)

European University Institute

University of Osnabrück

University of Maastricht

University of Turku

Complutense University Madrid

Hellenic Foundation for European and Foreign Policy (ELIAMEP)

University of Milan

University of Potsdam

Platform for International Cooperation on Undocumented Migration (PICUM)

International Centre for Migration Policy Development (ICMPD)

Migration Policy Institute Europe (MPI-E)

University of Warsaw

Vrije Universiteit Brussel

Associated Partners:

Toronto Metropolitan University

University of Leicester

University of Oxford