Research and Documentation Centre

Memorandum 2022-1

Intergovernmental relations and return

Part 1: Measuring enforced return in Europe

An assessment of the validity and reliability of EU data on orders to leave and the return of third country nationals

Research and knowledge centre for
the Ministry of Justice and Security

Memorandum 2022-1

Intergovernmental relations and return

Part 1: Measuring enforced return in Europe

An assessment of the validity and reliability of EU data on orders to leave and the return of third country nationals

M. Maliepaard
M. van der Meer
A.S. Leerkes
M. Ramdin

Memorandum
This series comprises overviews of studies carried out by or for the WODC Research and Documentation Centre. Inclusion in the series does not mean that the sheet's contents reflect the viewpoint of the Dutch Minister of Justice and Security.

Acknowledgements

Each year the Member States of the European Union issue around 500,000 return decisions to persons who do not, or no longer, have legal stay. A return decision requires the person to leave the territory of the state issuing the return decision and to go to a country where he/she does have legal stay, usually his/her country of citizenship. If persons do not leave themselves, they risk being returned by force. The implementation of assisted and forced return often requires cooperation by the countries of citizenship of the person receiving the return decision, and thus partially depends on the intergovernmental relations between EU+ (EU Member States plus Norway, Switzerland, and the United Kingdom) and non-EU+ countries. The WODC has conducted three interrelated studies on the influence of these relations on return:

1 A preparatory study to critically assess the validity and reliability of the European data on enforced return.
2 A quantitative longitudinal analysis of the effects of different types of EU-wide and bilateral intergovernmental return frameworks (e.g., re-admission agreements, Mobility Partnerships) on the registered rates of enforced return from the EU+ countries to the non-EU+ countries.
3 A pilot for the Netherlands and Norway to qualitatively explore the role of return frameworks and other (inter)governmental return strategies during concrete procedures to implement assisted and forced return to Afghanistan, Iran, and Iraq, and to identify possible opportunities for mutual learning between EU+ countries.

The first two studies were carried out by the WODC. For the comparative pilot, the WODC collaborated with the Norwegian Institute for Social Research (ISF).

During the project, the researchers were assisted by an Advisory Committee that consisted of: Prof. dr. Mirjam van Reisen (Tilburg University, chair), Prof. dr. Mathias Czaika (University for Continuing Education Krems), Dr. Natascha Wagner (Radboud University), and Maykel Bouma, LL.M. (Ministry of Justice and Security). I would like to express my gratitude to all the members of the advisory committee for their valuable feedback. In addition, I would like to thank ISF for its gracious offer to collaborate on the comparative pilot. I would like to thank all the professionals from various organisations at the EU-level, the Netherlands, and Norway for their participation in the first study and the comparative pilot.

Lastly, many thanks as well to WODC-colleagues Nikolaj Tollenaar and Sanne Boschman, who were very helpful in thinking along during the analytical stage of the second study, and to Marita Kok for editing the reports.

Prof. dr. G.J.L.M. (Gerty) Lensvelt-Mulders
Director WODC

Table of contents

1 Introduction

Effective return of migrants who do not, or no longer, have a legal right to stay has been an important policy focus in the European Union for many years (for an overview, see EMN, 2017; Eisele, 2020). In 2008, the EU laid down common EU rules on return in its Return Directive (2008/115/EC), which aimed to 'provide for clear, transparent and fair treatment of irregularly staying migrants, while fully respecting the human rights and fundamental freedoms of the persons concerned'. The Directive stipulates that Member States, as a rule, are to issue a return decision to all third country nationals (TCNs) who are apprehended on the grounds of illegal stay, and to all rejected asylum seekers who are no longer in a regular asylum procedure. With the return decision, the issuing state informs the TCN that they are legally required to leave the territory. Persons who do not leave voluntarily following a return decision may be returned by force. The Return Directive classifies return as voluntary if states do not directly apply physical force; a definition that also includes returns under the threat of deportation. Defined as such, both forced and voluntary return can be considered forms of *enforced return* as most persons involved would not have chosen to return to their country of origin had they been allowed to stay. Voluntary return is the preferred outcome for EU Member States as it is associated with lower human, financial, and political costs. Voluntary returns can be assisted by organisations like IOM, although TCNs may also return of their own accord (unassisted voluntary return). The Return Directive explicitly states that forced return must be treated as a last resort.

The EU – including individual Member States like the Netherlands – has taken various measures to increase the return of TCNs who do not have legal stay. These include measures to make such stays less attractive, but also measures to facilitate voluntary return using re-integration schemes (cf. Leerkes & Van Houte, 2020). Additionally, the EU and individual EU Member States have invested heavily in establishing return frameworks with a range of third countries (Cassarino, 2007; Vara, 2019). These frameworks differ in content and degree of formality. Frameworks are sometimes linked with development aid, trade, or the facilitation of legal migration or mobility. There are also frameworks that are not explicitly linked to other issues (cf. Jurje & Lavenex, 2014). Some frameworks are bilateral between an EU Member State and a third country, while other frameworks with third countries are EU-wide (e.g. the 'EU-Turkey deal', EU Re-admission Agreements, EU Mobility Partnership).

Return also has a significant place in the EU Migration and Asylum Pact, which was presented in September 2020 by the European Commission as a new 'comprehensive approach to migration management'. It contains provisions to strengthen return and aims to more closely link asylum and return procedures (European Commission, n.d.). With regards to relationships with third countries, the EU wants to increase investments in partnerships with third and transit countries, identifying return as one of the 'key areas' (European Commission, 2020). The Pact also proposes that EU Member States can offer 'return sponsorships' to each other. In that case, an EU Member State would help another Member State with migrant return. The underlying assumption is that an individual European country may have special ties with a specific third country, and that it may share the benefits of that relationship with a European partner. Put differently, the pact assumes that some return corridors – which we could define as a route from a returning state to a re-admitting state – are characterised by

higher rates of cooperation on return than comparable return corridors (e.g. rates of return may be higher from France to Afghanistan than from Denmark to Afghanistan).

Despite the importance of the theme of return to the EU and its Member States – including the Netherlands – and the variety of approaches taken, there is a remarkable lack of scientific research on the effects of international relations on return, which includes research on the different bilateral and EU-wide return frameworks mentioned above. We do not really know how the return frameworks affect the number of forced and/or relatively voluntary returns. There is also no established methodology to measure levels of cooperation on return in different return corridors.

The Dutch Ministry of Justice and Security has asked the WODC to conduct a quantitative and qualitative study on the effects of international relations, which includes bilateral and EU-wide return frameworks, on rates of forced and voluntary return in the EU and the Netherlands. Additionally, the Ministry is interested in knowing whether it is possible to identify return corridors involving other EU Member States with higher rates of forced and/or voluntary return than corridors involving the Netherlands. It is hoped that the Netherlands can learn from the experiences in these corridors.

Good quality data on return of TCNs is essential to answering these questions, a sentiment echoed by the European Parliament and Council (Regulation 2020/851). Quality data on the *population at risk* is equally important: how many TCNs receive return decisions, or should receive return decisions under the Return Directive? Without that information, the numbers on enforced returns are not very informative as it is not possible to calculate *return rates* (i.e. the percentage of returns among those receiving return decisions).

Since 2008, EU Member States have been collecting information on the Enforcement of Immigration Legislation (EIL) for Eurostat, the EU agency responsible for collecting and publishing data from all EU Member States. EIL data consist of an array of statistics, including information on orders to leave issued to TCNs (which approximates the population at risk for return), as well as returns 'to a third country' following an order to leave. The European Commission (EC, 2017; EC, 2021) combines these two data sources to calculate the return rates of the EU Member States, and has used these rates to gauge the countries' 'effectiveness' in returning TCNs without legal right of residence. But do these rates reflect actual rates of return? And can they be compared across EU Member States? These are questions pertaining to the validity and reliability of the data. Methodologically, a measure has high *validity* if it reflects the phenomenon that it is believed to measure, and high reliability if a repeated measure of the same phenomenon would yield the same value (De Vaus, 2001). If 5% of the Afghans who received a return decision by Spain in 2017 actually returned voluntarily to Afghanistan within say 12 months of the return decision, it would be problematic if the measured rate (i.e., the number of voluntary returns from Spain to Afghanistan in 2017 divided by the number of its return decisions involving Afghans in 2017) would indicate that it is 40%. If such biases occur occasionally, and if we also observe return rates that are too low (e.g., 1%), the measure has low *reliability*. If certain upward or downward biases occur systematically, the measure has low *validity*. Additionally, if the measured return rate for the Afghanistan-France return corridor is 20% against 10% for the Afghanistan-Germany corridor, it would be problematic if this difference were completely due to differences between France and Germany in how

return and/or return decisions involving Afghans are defined, used, and registered; comparability across EU Member States can be seen as an aspect of reliability.

In Part 1 of the study, the results of which are reported here, we explore the validity and reliability of Eurostat return data. The aim is to investigate whether and how Eurostat return data can be used to answer research questions on factors influencing return rates across EU Member States and across third countries. We pose two research questions:

1 What can be said about the validity and reliability of the EU data on returns and return decisions?
2 If there are issues pertaining to the validity and reliability of EU data on returns and return decisions, what methodologies are suitable to research the effects of return frameworks on return outcomes and/or to identify differences between comparable corridors in the level and/or type (e.g., forced vs. voluntary) of return?

To answer the first question, we studied policy documents and literature on the implementation of the Return Directive across Member States, and on (differences in) registration of return. To complement the existing literature and policy documents, we conducted four interviews with experts in the field of return data, and assessed the Eurostat data (see Box 1). The second question was answered by reflecting on the main limitations identified, and by experimenting with a number of statistical methodologies.

In the next section, we will provide an overview of the issues we identified as threats to the validity and reliability – including the international comparability – of the Eurostat EIL data. In each subsequent section we will go deeper into these issues. Both potential causes as well as potential ways of dealing with these issues to enhance the validity and reliability of the return rate as a measure of enforced return will be discussed. In the conclusion, we also formulate some suggestions to improve the validity and reliability of the EU return data.

1.1 A schematic overview of the issues

This paragraph contains a schematic overview of the issues we identified as having impact on the validity and reliability of the Eurostat EIL data and on which we will elaborate in the following chapters (see Table 1). The discussion focuses on two Eurostat variables that may be used to calculate return rates: 'returns to a third country following an order to leave' (the numerator of the return rate) and 'third country nationals ordered to leave' (the denominator of the return rate). In this factsheet we argue that differences between Member States that reduce the validity and reliability of the EIL return data can be divided into those related to (1) ways of registering by EU Member States; (2) lack of data; (3) the data structure itself; and (4) more substantive underlying (policy) differences in approaches to migrant return. Issues relating to the validity and comparability of the data thus originate from different (types of) sources. These issues also affect the return rates based on the Eurostat EIL data.

Table 1 An overview of the issues identified, the nature of the threat they pose, their impact on the return rate, and ways forward

Issue	Threat to validity	Threat to comparability	Impact on return rate	Way forward / take-home message
Differences in registration • Double counting of orders to leave • Counting returns of people without an order to leave	 X X		 Underestimation Overestimation	Improve availability and quality of metadata.
Lack of data • Unknown for most EU Countries whether 'return to a third country' is to country of citizenship or to other third country (e.g., transit country) • Unregistered return	 X X		 - Underestimation	Estimate proportion of TCNs returning to their own country of origin; Drop third countries whose subjects frequently return to other countries; Drop TCNs who 'return' to EU countries from the numerator and denominator when calculating return rate. Be aware that 'return' refers only to registered return.
Data structure (no cohort data)	X		Over- or underestimation	Smoothing of return rates over time Select cases (drop small cases).
Differences between Member States in the conditions under which migrants get an order to leave (Directive leaves Members State discretion)		X	Depends	Avoid comparison of return rates across countries; deviation method preferred. Use number of rejected asylum seekers for selected countries as indicator of population at risk.
Differences between Member States in whether or not they include unassisted voluntary return (Eurostat guidelines leave Member State discretion)		X	Including unassisted return boosts the return rate relative to countries that do not report such returns	Improve information in metadata on inclusion of voluntary return. Avoid comparison of return rates across countries; deviation method preferred.

Box 1 A note on methodology

We combined several methods to answer our research questions. In addition to our desk research on legal and policy documents and the Eurostat metadata, in March and April 2021, we conducted four interviews with experts working in the field of data on return. One interview was a group interview with three experts working in the same organisation but with different expertise. The respondents were approached via our network within the Ministry of Justice and Security. Questions for each interview were developed beforehand and were tailored to the information we expected to obtain given the experts' positions in the field. All interviews were digital video calls and lasted for approximately one hour. For three of the four interviews, permission was given to record the conversation. An interview report was made of each interview. In some cases, additional clarifying questions were asked via e-mail. All references to the interviews made in this factsheet were sent to the interviewees for a factual check beforehand. It should be kept in mind that only a few experts were interviewed; the interviews mainly served for additional contextual information to better understand the Eurostat EIL return data. We also investigated the data directly, by combining the datasets migr_eiord en migr_eirtn for the period 2007-2018. We assessed how orders to leave and return relate to one another across countries over time. For more details, see Box 2.

2 Issues when working with the EIL data

To assess whether return decisions indeed lead to returns, and to estimate the effectiveness of different return frameworks, it is necessary – albeit not sufficient – to be able to assess what proportion of third country nationals who are supposed to return to their countries of origin actually do so. In the past, the Eurostat EIL data have been used by the European Commission (EC, 2017; EC, 2021), among others, to calculate and compare return rates. To construct a rate of enforced return, for each year, the (registered) enforced returns are assessed in relation to the population at risk of return in a given country (usually based on how many orders to leave were issued). In Eurostat, the data on orders to leave and data on enforced return are compiled separately (the 'migr_eiord' and 'migr_eirtn' datasets, respectively).

The migr_eiord data provide information on the number of third country nationals who have been given an order to leave. Specifically, 'Third country nationals found to be illegally present who are subject to an administrative or judicial decision or act stating that their stay is illegal and imposing an obligation to leave the territory of the Member State (see Art. 7.1 (a) of the Council Regulation (EC) no 862/2007[1])' (Eurostat, 2021). The migr_eirtn data indicate the number of third country nationals who return, following an order to leave. More specifically: 'Third country nationals who have in fact left the territory of the Member State, following an administrative or judicial decision or act stating that their stay is illegal and imposing an obligation to leave the territory (see Art. 7.1 (b) of the Council Regulation (EC) no 862/2007[2]). On a voluntary basis Member States provide Eurostat with a subcategory which relates to third country nationals returned to a third country only. The EIL statistics based on Council Regulation (EC) no 862/2007 include forced returns and assisted voluntary returns. Unassisted voluntary returns are included where these are reliably recorded' (Eurostat metadata EIL).

The metadata indicate that neither migr_eiord nor migr_eirtn data include persons who are transferred from one Member State to another under the mechanism established by the Dublin Regulation. According to the metadata, each person should only be counted once within the reference period.

Of course, working with any data source requires mindfulness of the fact that there is never a one-to-one relationship between the phenomena themselves and the data collected about these phenomena. There is always a certain amount of noise present. We will now address the most important issues faced when working with these two data sources, as well as the return rate that we have based on a combination of both. We will discuss issues related to registration, lack of information, related to the data structure, and country differences in Member States' implementation of the Return Directive when issuing return decisions.[3]

1 See Council Regulation (EC) no 862/2007.
2 See Council Regulation (EC) no 862/2007.
3 Parallel to our research activities for the current publication, Belmonte and colleagues (2021) undertook a similar effort in identifying underlying issues with the EIL data, focusing particularly on policy differences across Member States and on composition differences across Member States in the population to return.

2.1 The issue of differences in national registration systems

To maximise comparability across countries, Eurostat relies on Member States supplying data in fixed formats and following strict guidelines. Nevertheless, both the Eurostat metadata and multiple experts we interviewed point to country differences in reporting resulting from differences in national registration systems. According to one expert, differences in registration are harder to pinpoint than policy differences (see paragraph 2.2 for the latter); and indeed, it is not easy to obtain an overview of country differences in registration. Based on the literature, an assessment of the metadata provided by Eurostat, and the expert interviews, we have gained insight into some of the important differences present in the data across countries, both in the registration of 'orders to leave' (return decisions) and in the registration of enforced return itself.

Firstly, regarding registration differences in the way orders to leave are counted, Member States do not always adhere to the Eurostat technical guidelines[4] to only report the number of persons who received an order to leave, rather than the number of orders issued. Greece, Germany, Austria, Poland, and Belgium report that persons appear multiple times in the data (Eurostat, 2015). Illustrating why this happens, Eisele (2020) points to the fact that, in the past, The Netherlands issued multiple orders to leave to one person, without correcting for this when supplying the data to Eurostat.[5] Such practices conflict with the information provided by the Eurostat metadata. Counting orders to leave rather than persons receiving an order to leave negatively affects the return rate. To illustrate this, imagine that all 100 TCNs from Albania, for example, were to receive two orders to leave a Member State, and they all left within that year. In this scenario, the ratio between returns and orders to leave would be (100/200), yielding a 50% return rate. Yet, in reality, all of the TCNs returned. Unfortunately, given the fact that the size and scope of reporting differences are unknown, their effect on Member States' overall return rates cannot be ascertained. The metadata on this topic have unfortunately not been updated since 2015, and only Poland gives an estimate of the size of this issue (>20%).

A second issue centres on the way enforced return is registered. The Eurostat guidelines state that only returns following an order to leave should be reported. It seems that not all Member States abide by this guideline. For instance, Carrera (2016) states that there are indications that (particularly in the earlier years), not every TCN who returns has received an order to leave. On the basis of information provided by one of our interviewed experts, we can state that at least for the Netherlands, there are still cases in which people are included in the return data, despite not having received an order to leave.[6] Specifically, this is applicable to TCNs who are transferred to another Member State under the mechanism established by the Dublin Regulation, and people who have left with the help IOM or another NGO prior to having received a return decision. It is beyond the scope of this study to assess whether this is also the case for other countries; a lack of metadata on this topic makes it difficult to determine the scope of this issue. Including persons who return without having received an order to leave in the return data boosts return rates. To illustrate, let us say there are 100 Taiwanese who returned from the Netherlands in a given year. However, only 50 received an order to leave. If the Netherlands were to count all

4 See Eurostat technical guidelines.
5 One of the experts we interviewed informed us that from 2021 on, the data supplied to Eurostat no longer includes double orders to leave.
6 According to the information provided, this was an issue until 2020; from 2021 on, the return data should include only those who received an order to leave.

100 returns, but only issued 50 orders to leave, this would result in a return rate of (100/50) *100 = 200%.

Beyond these reporting differences between Member States, there is, as with all data, general noise. The fact that in many Member States, different organisations are involved in issuing return decisions and in registering return (EMN, 2017), further increases the risk of registration errors. Experts we interviewed also highlighted this issue. This is exacerbated by the fact that registration of these data is highly complex, given the volatility of individuals' legal statuses even *within* a year. For instance, as some respondents pointed out, the status of a return decision can change from 'removable' to 'non-removable' in case of an appeal; and return decisions may become void, for instance if an asylum status is granted following an appeal (see also European Court of Auditors, 2021; Belmonte et al., 2021). It is unlikely that countries (can) take these issues into account when supplying data to Eurostat. For the Netherlands, there is quite an elaborate system of checks to deal with potential irregularities in the data. To what extent other countries have similar systems is unfortunately beyond the scope of this study. That said, one of the experts does indicate that while some Member States, such as Sweden, also have a data quality control system, others do not.

The fact that there are issues pertaining to the quality of the data is not surprising, given the fact that data were collected by different organisations in different countries, and given the fact that the data were not collected specifically with scientific research in mind. However, a crucial problem here is that issues with data quality are obfuscated by the lack of metadata. Metadata should be presented in a form that 'facilitates proper interpretation and meaningful comparisons' (Eurostat, 2011, p. 8). Currently, however, this is not the case. Various data issues mentioned above, such as double counting or including categories that should not be included, can only be uncovered after studying literature pertaining to the EIL data and speaking to experts. While Eurostat provides national quality reports in which deviations from the guidelines are mentioned, these reports are available for a limited number of countries and lacking in detail. Moreover, the quality control check Eurostat performs on the data is mainly data-technical in nature (e.g., do sub-categories add up to the reported totals), as explained by one of our respondents. Based on Eurostat's information, it is impossible to ascertain the size and scope of reporting differences.

Lack of (quality) metadata has sometimes been attributed to a wish among government actors to obscure 'messiness' which might reflect negatively on the institute collecting or country supplying the data (e.g. Scheel & Ustek-Spalda, 2019).

2.2 The issue of lack of data

Data is, by its nature, a simplification of reality. However, the more detail is generated in the data, the more accurately the processes it approximates can be described. Detail strengthens the validity of indicators. Unfortunately, the data that Member States are currently obliged to supply to Eurostat are quite general in nature and do not include type of return or country of destination. Lacking information on the precise nature of enforced return leads to problems when estimating return rates.

Enforced return is generally subdivided into forced return, assisted voluntary return, and unassisted voluntary return. While forced and assisted voluntary return are generally registered, this is often not the case with unassisted voluntary return.

Unregistered (unassisted) return is not currently (and cannot, by its nature, be) included in the data on return, but this group is included in the data on orders to leave. An important consequence of this is the structural underestimation of any return rate, based on the relation between orders to leave and registered returns (as also noted by Pluim & Hofmann, 2016). To illustrate this, let us assume that there are 100 Somalis in the Netherlands who have received a return decision. Of these 100, 25 leave through forced return, 25 through assisted voluntary return, and 25 return of their own accord, without being registered as having returned. The actual return rate among this group is ((25+25+25)/100) = 75%; the return rate derived from Eurostat data however would be ((25+25)/100) = 50%. Unregistered return might therefore (depending on its magnitude) have a significant effect on the return rate. It is therefore prudent, when working with return rates, to indicate that they are structurally underestimated as TCNs who have returned of their own accord are not counted as having returned.

A related, and potentially more complex, problem lies in the fact that unregistered return and registered return are most likely interrelated. How they are related is however unknown. It might be that in countries that invest heavily in forced returns, people are more likely to leave of their own accord, to prevent forced return. In other words, if return decisions are hardly enforced, voluntary return becomes less 'attractive'. At the same time, the higher the voluntary returns, the lower the population at risk of enforced return.

We secondly lack important information about the country to which TCNs actually return. The data provided by Eurostat differentiate return to another Member State, and return to 'a third country'. This third country is not necessarily the country of citizenship (or birth) of the returnee – it could, for instance, be a transit country they passed on the way to Europe. This can create an issue, depending on the question one wants to address. For questions specifically pertaining to international relations, one would ideally want to know who is 'returning' and to where. While there is some information on this in the Eurostat data (variable *migr_eirt_des*), this information is only available for 19 countries and only for a limited number of years. For the countries for which we do have this information, we can see that although the majority of TCNs do return to their country of origin, there are specific nationalities whose people frequently return to other countries. To redress the issue of not knowing the country of return, one might drop these groups from the data altogether. Ideally, however, one would analyse data based on country of return rather than 'third country'.

2.3 The issue of the data structure

A final and crucial issue is the fact that the data are not cohort data. The data Eurostat provides are based on persons, but are aggregated by third country, EU Member State, and year. This is done separately for the number of TCNs ordered to leave and the number of TCNs who returned following an order to leave. Because of this, it is impossible to follow individuals over time. This affects the return rate. The Return Directive (Art. 7 sub 1) states that orders to leave 'shall provide for a period for voluntary departure of between seven and thirty days'. In some instances, Member States may choose to grant a shorter period or no period at all (Art. 7 sub 4). Despite this relatively brief period granted to TCNs to organise voluntary departure, organising enforced return may take (much) longer. It is therefore likely that there are TCNs who

receive an order to leave in one year and then leave in the next (as recognised in Eurostat, 2021). Given the data structure, return rates will be structurally miscalculated, with returns being underestimated in year T, and overestimated in T+1 (Eisele, 2020; European Court of Auditors, 2021). The scale of this issue cannot be ascertained based on the Eurostat data alone, nor can cases in which people receive an order to leave in one year and return in another year be removed from the data. As a result, the odds of being returned can never be fully known, given that there will always be a margin of error in the data.

An analysis of the 2019 return data for the Netherlands – executed upon the request of the authors – indicates that in this subset of the data, approximately two thirds of returns happened in the year the order to leave was issued. In 2019, the average time between the order to leave and the return was approximately six months (Personal communication Ministry of Justice and Security, July 2021). This gives some indication of the size of the issue at hand, though we cannot establish how representative this is for other years or other Member States. The longer countries take to organise return, the larger this problem will be.

The issues discussed in this chapter are general in nature, and as such affect all countries. Although this happens in different ways and to different degrees across countries, the issues reduce the data quality across the board, threatening the validity of the data and of any return rate based on these data. Because of these limitations, the return rate has been labelled as a 'misleading' indicator (Eisele, 2020); and its explanatory power dubbed 'inherently limited' (Pluim & Hofmann, 2016). In the following chapter, we will discuss a different type of issue, which mainly threatens the reliability of the data.

3 Threats to the intra-EU comparability of the data

The previous section made clear that one must be cautious in using the Eurostat EIL data to assess the return rate of third country nationals without legal right to stay. In this section, we move beyond the data itself to the return policies underlying the data. We look at the way EU policy on the return of third country nationals is implemented by different Member States, and the way this implementation affects the reliability of the ensuing data. We focus in particular on comparability across Member States.

3.1 Country differences following from discretion in the Return Directive

With the establishment of the Return Directive in 2008, the EU has attempted to harmonise Member States' return policies. Among other things, the Directive stipulates the conditions under which TCNs in EU Member States should be issued a return decision (which means they are ordered to leave), after which enforced return (either assisted or forced) can be set in motion. Return decisions should be issued to TCNs who are apprehended on the grounds of illegal stay, as well as to rejected asylum seekers who are no longer in a regular asylum procedure. However, the Directive also leaves some discretionary power to the Member States when it comes to applying the Directive to certain groups of TCNs as explained in Article 2(2). Member States have the choice not to apply the Directive to third country nationals who are refused entry at the border; who are apprehended while illegally crossing the border; who are subject to return as a criminal law sanction; or who are subjects of extradition procedures. In an important study, the European Migration Network (EMN, 2017) assessed the implementation of the Return Directive in different EU+ countries. The report gives information about 22 EU Member States; Bulgaria, Denmark, Poland, Portugal, and Romania were not covered. It is clear from the report that many EU countries use the discretion present in the Return Directive, but each Member State does so in its own specific way. As Table 2 shows, certain groups of TCNs will receive an order to leave in some Member States, but not in others. In addition to differences that follow directly from the discretion in the Directive, there are also some differences across Member States in the issuing of return decisions to TCNs who are illegally present (EMN, 2017). For instance, Member States have different practices in issuing orders to leave when the whereabouts of TCNs are unknown; when TCNs do not have valid identity or travel documents; and when TCNs are found to have been staying illegally in the country at an exit check (see Table 2).

Another difference in policy between Member States documented by EMN (2017) pertains to whether the return decision is issued together with the decision to end the legal stay of a TCN. According to EMN (2017, p. 22), in 'at least four countries' these two decisions are not issued at the same time; the return decision is issued at a later moment, for example after all asylum appeals are exhausted (Ireland[7]). Related to this, according to one expert we interviewed, it appears that some Member States only issue a return decision shortly before a TCN actually returns. Although there is no hard evidence for this, the latter point would provide a strong artificial boost to the return rate.

[7] It must be noted here that Ireland is not bound by the Return Directive, but has implemented similar provisions (EMN, 2017).

Table 2 Overview of exceptions to issuing return decisions in the EU Member States

Member State	TCNs are NOT issued a return decision when they ...						
	Are subject to a refusal of entry	Are apprehended or intercepted while irregularly crossing the external border	Are subject to return as a criminal law sanction or as a consequences of a criminal law sanction	Are the subject of an extradition procedure	Do not have a known address [whereabouts unknown]	Lack identification or travel document	Are detected during an exit check
The Netherlands	X						
Austria	X	X	X				
Belgium	X	X	X	X			X
Croatia							
Cyprus			X				X
Czech Republic	X	X	X	X	X	X	
Estonia							
Finland							
France	X	X	X				X
Germany	X	X		X			X
Greece		X			X	X	X
Hungary		X			X		
Italy	n.a.	n.a.	n.a.	n.a.	X	X	X
Lithuania	X	X	X	X	X		
Luxembourg	X		X	X			
Latvia	X	X	X	X	X		

Member State	TCNs are NOT issued a return decision when they …						
	Are subject to a refusal of entry	Are apprehended or intercepted while irregularly crossing the external border	Are subject to return as a criminal law sanction or as a consequences of a criminal law sanction	Are the subject of an extradition procedure	Do not have a known address [whereabouts unknown]	Lack identification or travel document	Are detected during an exit check
Malta	X	X	X		X	X	
Slovenia		X	X				X
Slovakia							
Spain	X	X	X				
Sweden	X		X	X			
United Kingdom	n.a.	n.a.	n.a.	n.a.			X

Source: EMN (2017, p. 16-17, 20)

The Member State differences discussed in this paragraph point to the fact that although the EU-wide Return Directive is in place, countries vary significantly in the way they implement it. These country differences in issuing return decisions not only affect the number of return decisions issued each year, but most likely also affect the return rate. For instance, if TCNs apprehended or refused at Schengen borders receive a return decision, they will most likely return immediately to their country of origin, or to a 'third country' (i.e., non-EU+ country). Countries that do issue orders to leave to TCNs who have been refused entry at the border, will therefore most likely have a higher return rate than countries that do not. The same goes for countries that issue return decisions to those whose illegal stay is only detected as they exit the country. Particularly when border refusals make up a large proportion of the total number of return decisions in these Member States, the effect on the return rate is likely to be sizeable; since these people are returned immediately, the return rate will be relatively high. On the other hand, several states issue return decisions to people whose whereabouts are unknown, whose identity is not established, and who do not have valid documents (EMN, 2017; Eisele, 2020). These orders to leave can rarely be enforced, thus bringing down the return rate of these countries (Eisele, 2020, p. 65).

Thus, as differences in the implementation of the Return Directive result in differences in the data registration, they result in less comparable 'third country nationals ordered to leave' data, and, indirectly, less comparable return rates. Given the varying ways in which Member States implement the Return Directive, interpreting return rate as 'effectiveness' of return across countries is highly problematic. Comparing average return rates across countries should therefore be avoided. This was also concluded by some of the experts we interviewed, with one of them stating that comparing numbers of return decisions issued by EU Member States is 'hardly possible without many disclaimers'. It is important to note that the Return Directive was established in 2008 and implemented by the EU Member States afterwards (2011 in the Netherlands). It is likely that return policies across Member States were (even) less uniform before the Return Directive.

3.2 Country differences arising from discretion in Eurostat guidelines

As we have seen above, discretion in the Return Directive diminishes the comparability of data across Member States. Comparability of the data turns out to be further compromised by discretion in the Eurostat guidelines for supplying data. Though the guidelines are strict in many regards, when it comes to return, it grants Member States discretion in deciding whether or not to include voluntary unassisted return in the data. Member States are requested to supply information including both forced and voluntary return, whereby return is defined as '*Third* country nationals who have in fact left the territory of the Member State, following an administrative or judicial decision or act stating that their stay is illegal and imposing an obligation to leave the territory'. In addition, Member States may include unassisted voluntary return in the data, if it is 'well documented'.[8] This discretionary space introduces noise into the return data, as reported differences do not necessarily reflect actual differences in the number of returns. It is highly likely that unassisted voluntary return does occur from Member States that do not include this type of return in the data supplied to Eurostat. The discretion of Member States to include unassisted voluntary return in the return data is important, as it affects the return rate. Those that include unassisted voluntary return will have a higher return rate. Unfortunately, Eurostat metadata do not always

[8] Enforcement of Immigration Legislation (migr_eil) (europa.eu).

specify which types of return are included in the variable. For example, Croatia, Estonia, and the Czech Republic state that they follow the Eurostat guidelines, but do not specify whether unassisted voluntary return is included in the data.[9] Metadata for other countries point to country differences in which types of return are included. For instance, Denmark and Poland state that they include both assisted and unassisted voluntary return in their data, while Belgium chose not to include unassisted voluntary return, even when such return are confirmed.[10] Moreover, national quality reports are available for only 19 countries. For Member States that do not have a quality report, which types of return they include in their return variable remains unknown.

Ideally, data on forced return, assisted voluntary return, and unassisted voluntary return would be differentiated in the data. This is not the case, however. While most Member States supply information on return for the period 2008-2019, there is no distinction between types of return. Some countries have started providing data separately on assisted and forced return, starting in 2015. Nevertheless, as this information is provided to Eurostat on a voluntary basis, this data is only available for some of the Member States, not including the Netherlands. Given the fact that unassisted voluntary and assisted voluntary return are currently not supplied to Eurostat separately, they cannot be disaggregated in the data. There are two options to deal with this issue: (1) dropping all countries that are known to include unassisted voluntary return in the return figures; or (2) looking only at forced return. An issue with the first option is that for some countries, there is no information on whether or not they include unassisted returns in the data. Option two is thus more feasible, although this significantly reduces the number of countries and years available; additionally, a return rate based on forced return only is less suitable for answering questions of how many TCNs return.

In summary, to reliably assess return rates of third country nationals from Member States, it is critical that return decisions are issued and reported in the same way, and that return is measured and reported in the same way across countries. Clearly, as we have argued in this paragraph, this is currently not the case. The country differences in implementation of the Return Directive, and differences in reporting to Eurostat, pose a significant threat to the comparability of data across Member States. Even in the exact same situation, a TCN might receive an order to leave in one Member State, but not in another. In some Member States their return would then be counted, but not in others. While this is problematic in and of itself, a second problem is the fact that these country differences are not easily gleaned from the data. Member State data are compiled into a single database, in order to facilitate comparisons across countries (and over time). While this is of course convenient, it means that differences in implementation and reporting are obscured. This is particularly problematic since Eurostat's metadata files provide insufficient insight into the country differences, as mentioned above.

[9] Croatia (europa.eu); Estonia (europa.eu); Czechia (europa.eu)
[10] Denmark (europa.eu); Poland (europa.eu); Belgium (europa.eu)

4 Working with EIL data: what can and cannot be done?

The EU and individual Member States have invested heavily in increasing the return of immigrants without a legal right to stay. From a policy perspective, assessing whether return rates have changed as a consequence of these investments is an important evaluation instrument. It is also important to evaluate which Member States are more 'effective' in returning immigrants without a right to stay, and whether this differs depending on the immigrants' countries of origin. Can these questions be answered, given the problems with the EIL data and resulting return rates?

Based on our findings, we strongly advise against comparing return rates across countries in order to answer questions on the 'effectiveness' of return.

Member State differences in the implementation of the Return Directive, as well as in the implementation of the Eurostat guidelines, strongly threaten the reliability of the data, and of return rates derived from these data. As has become clear, the available metadata and other sources are insufficiently detailed to uncover exact differences between countries regarding the 'third-country nationals ordered to leave' and the 'third-country nationals returned following an order to leave' variable. In addition, where country differences are known, for instance in terms of implementation of the Return Directive, individual Member States are shown to deviate from the directive in unique ways, making it hard to find even two Member States that apply the directive in exactly the same way. Solutions that rely on excluding certain Member States or years from the data are therefore insufficient in dealing with underlying data issues.

Clearly, there are serious threats to the validity and reliability of the data that should not be overlooked. At the same time, the EIL data is the most detailed source of information on return across Europe, and, as such, valuable. In this chapter we develop several ways in which the data might be used to answer important questions on differences between Member States in return rates, changes over time in return rates, and on explaining return rates.

4.1 How can we compare countries?

As stated above, comparing countries' return rates is problematic given a lack of comparability of data. At the same time, EU Member States need insight into the realisation of return of TCNs by other Member States, for example to further develop existing policies and strengthen collaboration. Given the importance of the topic, we have come up with two potential analysis strategies that allow for comparison across Member States, while reducing the impact of the threats to reliability mentioned above. While by no means flawless, these approaches do provide an improvement on gross comparison of return rates across Member States in a given year.

4.1.1 The deviation method

Instead of comparing Member States based on return rates, we propose what we have dubbed the 'deviation method'. One of the main threats to comparability of return rates is that Member States differ in the groups they include in the data on orders to leave; and that some count instances rather than persons. As a consequence, average

country differences in return rate do not necessarily reflect actual differences in the probability of return. Our proposed deviation method circumvents this problem by taking not the average return from Member State to third country, but the deviation from the Member State average return rate in a given year.

While this does not allow for a general comparison of return across Member States, it does allow us to draw conclusions as to which Member States have particular difficulty or are particularly 'effective' in realising return to certain third countries. Let us say the Netherlands, Belgium, and Germany all have sizeable numbers of Afghan nationals without a legal right to stay. Which country is more effective in returning this group? A fictional example illustrates how the deviation method would work. Let us assume the average return from Germany, the Netherlands and Belgium in a given year is 70%, 50% and 30% respectively; and let us say the registered return rate to Afghanistan from each of these countries is 50%. Using the deviation method, return to Afghanistan from Germany would be 'below average' (-20 percent point); from Belgium, 'above average' (+20 percent point); and from the Netherlands, 'average' (at the mean level for the Netherlands return rate in that year). It would therefore seem that Germany has particular difficulty in returning Afghan nationals, whereas Belgium does not. This information may be useful if one wants to identify return corridors with a relatively high or low number of returns (also see part 2 of the study: Leerkes, Van der Meer, Paasche and Brekke, 2022). This indicator should always be used in conjunction with the number of return decisions – return corridors are particularly (or mainly) relevant in combination with a sizeable number of return decisions issued (if there are only 10 return decisions, an above average return rate is less meaningful than if there were 100 return decisions).

Despite being an improvement on the raw return rates, deviation data should be used with caution. The deviation method assumes that 'noise' in the data for a given Member State in a given year is stable across all third countries, which is potentially not the case. Deviations are less reliable if the average return rates of a Member State are low or high as a consequence of a few specific third countries with extremely high or low registered return rates (for instance, countries bordering the EU; or conversely, third countries to which no migrants can be deported because of war). Belmonte and colleagues (2021) have proposed several potential strategies to deal with this issue. It is also advisable to use triangulation, complementing the information with quantitative or qualitative data from other sources.

4.1.2 Focusing on rejected asylum seekers

A second way to address country differences while avoiding some of the threats to the comparability of the data is the approach that has been developed by Van Houte and Leerkes in their 2019 study (Van Houte & Leerkes, 2019). One of the key threats to the comparability of the EIL data is the different ways in which countries have implemented the Return Directive, making the figures for 'number of orders to leave issued' particularly incomparable. On this basis, Van Houte and Leerkes decided to use the number of negative asylum decisions as the denominator rather than the number of return decisions issued, and to limit the analysis to a number of countries where the vast majority of returning TCNs are rejected asylum seekers (e.g. Afghanistan, Iran, Iraq, Somalia, Eritrea, Syria). For instance, this approach looks at the number of Afghani returns from the Netherlands in 2018 as a function of the number of Afghani asylum claims that were rejected in this period in the Netherlands:

'For example, we then find a "return rate" for Afghan asylum seekers in the Netherlands of 25.2%; in the 2013- 2017 period, 980 Afghans were returned to a third country from the Netherlands, while 3,785 negative decisions involving Afghans were reached, 200 applications were withdrawn and 90 Afghans were returned to another EU country (980/(3785+200-90)=.252).' (Van Houte & Leerkes, 2019, p. 23)

Focusing on rejected asylum seekers is most suitable for questions pertaining to the return of groups where most persons who are 'at risk of return' are former asylum seekers. The estimations are most reliable for third countries from which mainly asylum seekers originate, and for Member States that are not located at the EU borders. Returns from Member States at EU borders are most likely made up to a larger extent by third country nationals found to be illegally present or stopped at the border. These migrants would be registered as 'returned to a third country' even if they are not returned to their country of citizenship (but to a transit country, for instance). For other Member States, returns to a third country are more likely to be limited to the country of citizenship. The main advantage of this approach is that it does not require the use of the 'ordered to leave' variable. In addition, voluntary unassisted return to (conflict) countries from which a lot of asylum seekers originate, is likely to be very limited (see Leerkes et al., 2010).

While this method provides insights into country differences, it also suffers from several methodological issues of its own – most importantly, the assumption that virtually all returnees are rejected asylum seekers. Additionally, the validity issues mentioned in Chapter 2 still hold (see Van Houte & Leerkes (2019) and Leerkes & Van Houte (2020) for a more elaborate discussion of limitations).

4.2 How can we assess trends?

How have return rates changed over the last decade? Have return frameworks affected forced and/or assisted return rates? These are some of the important questions in policy and research on return migration. The EIL data and resulting return rates would be ideal to assess trends in return over time. However, the pressing question concerns how they can be used to reliably assess trends over time, even within the same EU Member State. We investigate two ways forward: data selection and smoothing.

The Eurostat national quality reports suggest that Member States have not significantly altered their ways of implementing the Return Directive, or of reporting data to Eurostat, in the period 2008-2019. Based on our expert interviews, we can conclude that comparing data within Member States over time is less problematic than comparing data across Member States. At the same time, the issues pertaining to data quality mentioned in Chapter 2, still hold. The fact that the data are not cohort data in particular plays a role here.

As we have seen, the lack of cohort data may lead to return rates exceeding 100%, when return takes place in a different year than that in which the order to leave was given. In addition, the fact that return decisions and returns are generally registered by different organisations, can lead to discrepancies between the two, resulting in outlying return rates. How can this issue be dealt with? First, it is important to assess to what extent this problem is apparent in the data.

Box 2 Calculating return rates based on the EIL data

In this report, when talking about the return rate, we have described it as a simple rate between number of people ordered to leave and number of people 'returned'. However, Eurostat EIL data do not technically indicate return to the country of origin; they specify only whether TCNs left the EU Member State to go to another Member State, or returned to a third country. Based on the indicators, there are three ways to calculate return rates.

The first way to calculate a return rate (let us call it RR1), is to assess the number of people ordered to leave in relation to the number of people who left the country (including both TCNs who left to third countries and TCNs who left to other EU countries). Calculated as such, the return rate gives an indication of the proportion of TCNs who left the country after an order to do so. This measure might be suitable if the aim is to investigate which proportion of TCNs ordered to leave remained in the Member State. RR1 is less suitable to look at returns to a third country. Given that RR1 includes TCNs who moved to other EU countries, this operationalisation technically does not measure 'return'.

A second option is to remove those TCNs who left to another EU Member State from the numerator (RR2). RR2 indicates return to a third country as a proportion of all migrants ordered to leave. Thus, it is an indication of return to a third country, as compared to those who did not return and those who moved to another EU country. This measure is most suitable if the aim of the study is to see how many TCNs left European territory after an order to leave.

For answering the question of what proportion of TCNs return to a third country of all of those 'at risk' of doing so, it is most appropriate to exclude those who left to another EU country (most likely migrants who are in possession of a residence permit from another EU country) from both the numerator and the denominator (RR3). This way, we only assess the proportion of TCNs who returned, as a function of those TCNs at risk of return. At first glance, it seems that some countries do not disaggregate total return and return to third countries. For Latvia and Italy, the two are equal in each of the 12 years. In countries like Malta, Romania, Greece, and Cyprus, the percentages overlap fully in most years. Given that all countries that report a 0% return to EU countries are situated at the EU borders, this might not be a registration issue. It is possible that these countries do not return any TCNs to other EU countries. In recent years, countries have increasingly differentiated between the two destinations.

Data quality issues mentioned in this report play a role in all conceptualisations of the return rate. It is also important to remember that in all conceptualisations, the category of 'non-returnees' will unfortunately also include those TCNs who returned voluntarily. In that sense, there is no 'ideal' way of measuring return rate. However, it is important to be aware of different possibilities and what research questions they might help to address. In combination with the deviation method, RR3 would in our view yield the most suitable information to evaluate corridors in which return takes place. For ways to calculate the return rate to account specifically for composition differences across member states in the population to return, see Belmonte et al. (2021).

We have made a preliminary assessment of return rates based on Eurostat data from 2008-2019, excluding those who returned to other EU countries (using 'RR3' as an indication of the return rate; see Box 2). Each 'case' in the data refers to an EU Member State – third country corridor, in a given year – for instance, Germany-Somalia in 2017. Cases with zero orders to leave in a given year are excluded from the

data; There are roughly 25,000 cases in total for which information on the return rate can be calculated.

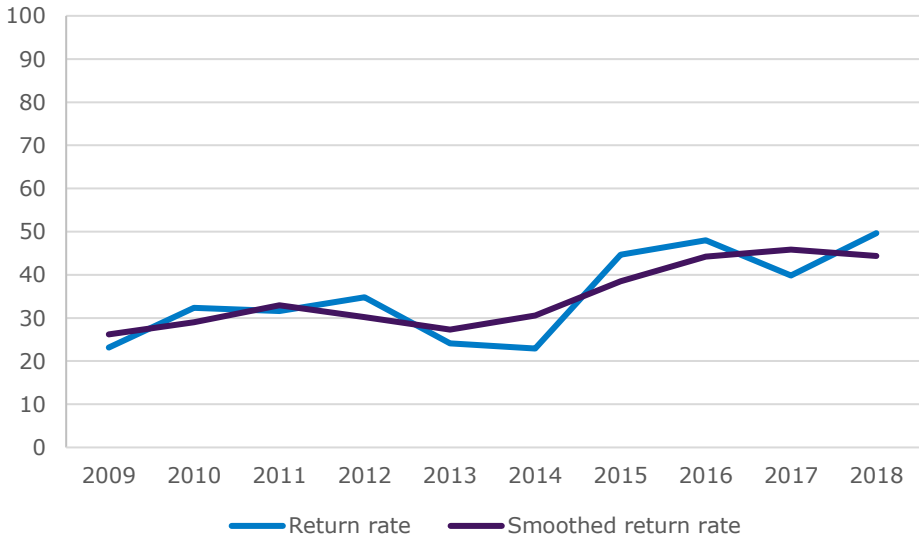
There are several instances of implausible return rates. While 92% of cases are within the 0-100% range, in 8% of cases, the return rate from an EU Member State to a third country in a given year exceeds 100%. In 2% of cases, it exceeds 200%, with outliers up to 6800%. There are strong differences across countries in terms of their return rates. The UK reports return rates of over 100% in a full 48% of cases, followed by Cyprus (28%). For the Netherlands, 13% of cases exceed 100%. Slovenia (12%) and Germany (11%) are other countries with relatively sizeable outliers. Third countries that are overrepresented in return rates exceeding 100%, are Ukraine, Moldova, Albania, Serbia, Georgia, and North Macedonia. Many of these countries are relatively close to or border the EU. However, more distant countries such as China, Malaysia, and the Philippines also rank quite highly on this list.

Clearly, the implausible return rates are not randomly distributed across EU Member States or third countries, indicating that the 'noise' in the data is not random, but particular to certain countries.

4.2.1 Smoothing and aggregation over time

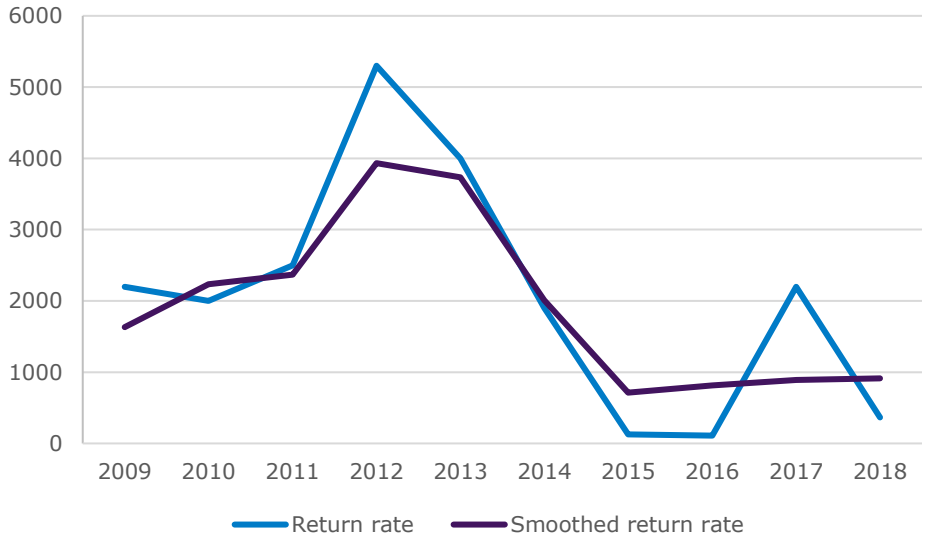
A first potential solution to the problem of implausible return rates is smoothing – a technique whereby the average rate across three years (T, T-1 and T+1) is calculated for each time point in the data. If the problem of implausible percentages is caused mostly by the lack of cohort data, smoothing should reduce these issues (see Figure 1 for smoothing applied to the data of the Netherlands).

Figure 1 Smoothed and unsmoothed return rates from the Netherlands (total) in percentages



Source: Eurostat, calculations WODC

Figure 2 Smoothed and unsmoothed return rates from the Netherlands to Nicaragua, in percentages

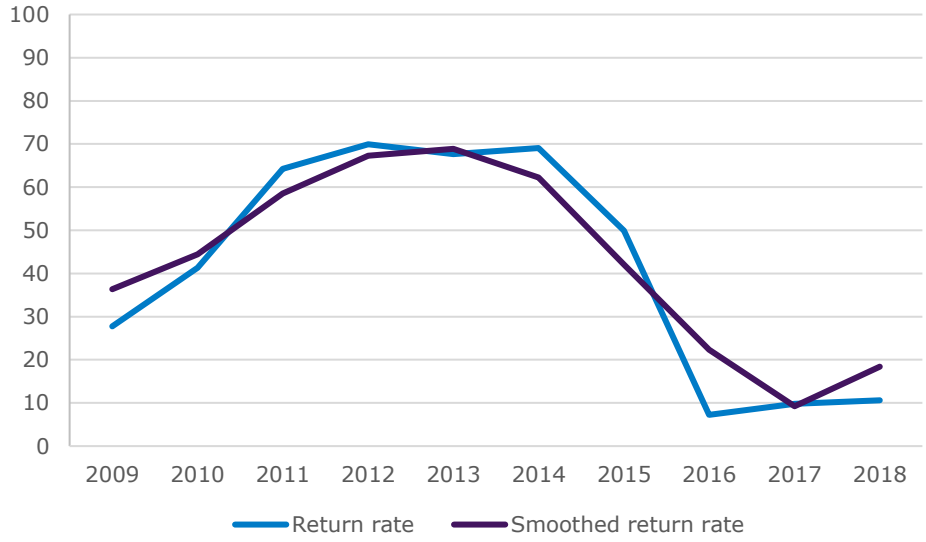


Source: Eurostat, calculations WODC

Applying smoothing techniques to the data results in a reduction of the most extreme outliers. At the same time, the average return rate is unchanged, and the standard error diminished with less than 10% (129 to 119), plus there are still some outliers of over 4000% (for an example, see Figure 2). Depending on the aim of a study, smoothing can be a way of dealing with strong fluctuations over time as it reduces variation over time (as can be seen in the reduced standard deviation on the annual return per EU country over 12 years for instance), but it does little to solve the issue of implausible rates. This is most likely because the implausible rates cannot (only) be attributed to the lack of cohort data, as evidenced by the fact that they are not randomly distributed across EU Member States and third countries.

Smoothing is also less fitting when you want to describe or study the impact of certain events on return rates, as can be seen in the case of Hungary, where return rates steeply dropped from 2015 to 2016, following the 'refugee crisis' – smoothing hides the suddenness of the drop here (see Figure 3).

Figure 3 Smoothed and unsmoothed return rates from Hungary (total), in percentages



Source: Eurostat, calculations WODC

Smoothing is therefore recommended when assessing global trends within countries over time, but it is not sufficient to solve data issues.

A similar technique to reduce the influence of statistical noise is to calculate return rates for larger periods of time (cf. Leerkes & Van Houte, 2020). By aggregating all returns over a period of multiple years, and then dividing the number by the total number of return decisions in this multi-year period, bias following from mismatching orders to leave and returns is reduced. This does of course impede studying change over time.

4.2.2 Data selection

A second potential strategy for dealing with implausible percentages is to make selections in the data. For example, one might select only those origin-destination combinations with sufficient cases; one could decide to select only third countries of which at least 10 nationals received an order to leave in a given EU Member State in a given year. For third countries with only few migrants ordered to leave by an EU Member State, implausible percentages are more likely to occur. For instance, if 3 people from Mongolia were ordered to leave the Netherlands in 2018, but they all left in 2019; and if in 2019 only 1 person from Mongolia was ordered to leave the Netherlands (and did indeed do so in that year), the return rate in 2019 would be 400% ((1/4) *100). This issue is exacerbated by the rounding of the data to the nearest 5, as Eurostat does. Because of rounding, if in a given year 7 people were ordered to leave and 8 returned, this would result in a return rate of 5/10 = 200%, instead of 7/8 = 114%. Thus, potentially, dropping small cases could work toward fixing the issue of implausible percentages. Given that return rates within dyads with very few orders to leave are not very informative (see also Belmonte et al., 2021), excluding them (regardless of whether or not the return rate is implausible) will benefit the overall quality of the analyses. It is important to note that there are very many 'small' cases in the data: 8,000 cases with 1-10 orders to leave, or approximately a third of all cases with at least 1 order to leave in a given year.

A look at the data indicates that the average number of orders to leave is indeed lower for cases with a return rate of >100% (Mean number of orders to leave = 486 vs. 1114, $p<0.001$). However, eliminating 'small' cases (up to 10 orders to leave in a given year) does not solve the issue of implausible rates; while the number of cases decreases, the proportion >100% does not change (it remains around 9%), nor does the average return rate. So, while dropping small cases will eliminate some irregularities from the data, it does not solve the issue of implausible percentages.

Data selection should not be limited to small cases. A second means of selection – which is prudent when assessing trends over time – is to exclude countries for which the data are of insufficient quality. Unfortunately, as we have argued earlier, metadata and country quality reports provide insufficient information for doing so; therefore, we can only use the data themselves as sources of information. Purely based on the data patterns, it seems that return rates from Latvia, Lithuania, Romania, Poland, and the UK should be interpreted with caution. Latvia, and to a lesser extent the other countries on the outside borders of the EU, report return rates close to or exceeding 90% on average, with Latvia even reporting a 99% return rate for three consecutive years. While it is possible that countries on the outside border of the EU have higher return rates as compared to other Member States, these outcomes warrant scrutiny. The average return rate in the UK even exceeds 100% on average in the 2008-2019 period but fluctuates strongly (up to 222% overall return rate in 2018). The UK is also overrepresented in cases with extreme percentages (>300%). It thus seems that the UK data are not reliable.

In conclusion, although the reliability issues mentioned in Chapter 3 pertain particularly to comparability across countries and less so over time, it seems that comparisons over time are quite seriously complicated by general issues of data quality and the fact that the data are not cohort data (threats to the validity), as described in Chapter 2. While we can be sure that percentages over 100 are very implausible, the fact that extremely high percentages exist in the data casts doubt on the overall quality of the data. While smoothing and data selection are theoretically good solutions, they do not solve the data issues at hand. Anyone assessing trends over time should therefore do so with caution.

4.3 How can we determine which factors affect return rates?

Researchers may want to use the EU return data to try to estimate the effects of origin-country, destination-country, and corridor-specific characteristics. It is important that such research takes the methodological limitations observed in this report into consideration. The most important limitation of the EU return data is that we cannot assume that differences between EU countries in the measured level of enforced return reflect real differences in enforced return levels, at least not when the number of return decisions is used as an indicator of the population at risk of return. This means that international variation in enforced return in the EU return data, cannot be attributed to actual differences between EU countries. While attributing differences in the levels of enforced return to differences between countries is problematic, there are ways to study the effects of origin-country, destination-country, and corridor-specific characteristics on return. Using so-called fixed effects regression models, it is possible to assess the influence of time-varying characteristics on enforced return. These models allow us to control for variables that were not or cannot be measured

(Allison, 2009). As such, they are not affected by (stable) registration differences in enforced return across EU Member States.

Let us give an example. We are interested in the effects of international relations, which include various bilateral return frameworks, on the level and type of enforced return. Fixed effects models would allow us to assess the effects of *new* bilateral re-admission agreements that were not in existence during the whole observation period. For instance, if France has a specific agreement with Afghanistan that was valid from 2014 onward, and also has relatively high rates of return to Afghanistan in that period, we can assess whether or not the higher rates of measured return are associated with the re-admission agreement. We can estimate whether the rates for the years during which the agreement was in force (2014-2019) were higher than the rates for the years it was not in force (2008-2013) by including a dummy for the presence of a re-admission agreement in a given year.

The downside of these fixed effects models is that the effects of time-invariant variables (things that do not change during the period under study) cannot be assessed. So, if you would want to study, say, the influence of having a colonial history on the return rate between two countries, or the effects of a bilateral return framework with a third country that was in force during all observation years, this approach is less fitting. In such scenarios, performing a multiple regression including EU-country dummies might to some extent be an alternative. However, this approach should be used with caution, and coefficients for the country dummies should not be interpreted as representing actual differences in enforced return. Using the same Afghanistan-France example, one might pool data on the return rates to Afghanistan for different EU countries and see whether a dummy indicating when the France-Afghanistan agreement was in place is associated with higher rates of return when the average rates of return from France, and the return rates of other EU Member States to Afghanistan, are controlled. This can be done by including a set of EU country dummies in the regression equation, possibly in addition to a set of dummies for different years so as to control general trends in measured and/or real return to Afghanistan from the EU. However, again, this approach should be used with caution. If we find an 'effect' of the Afghanistan-France framework, we cannot tell whether the effect is caused by the framework or by some other stable characteristics of the France-Afghanistan return corridor (e.g., other aspects of their international relations).

5 What do we need for (more) comparability?

As we have shown, the EIL data can be used to answer certain questions pertaining to the return of third country nationals. At the same time, it is evident that there are serious threats to the validity and reliability of the EIL data. There are several ways by which the EU might improve the comparability of EIL data across Member States.

1 A complete overview of country policy differences in implementation of the Return Directive
One important threat to the comparability of data is the fact that countries differ in the ways they implement the Return Directive. To deal with the issue of implementation differences, one approach to increase data comparability would be to make a selection of comparable countries based on the way they have implemented the Return Directive. This is not currently possible, as information is lacking for many countries (despite efforts by for instance EMN (2017) to uncover differences in implementation). To be able to make a more accurate selection, Eurostat should map out country differences in registration, in the same way EMN (2017) has previously done for a selection of countries. Such information should detail which groups of TCNs are included in the 'third-country nationals ordered to leave' variable, which will allow for more informed decisions on what Member States to include in analyses. It would be ideal, though likely unfeasible, to have separate data for every group of TCNs ordered to leave. That would enable separate analyses of groups that every EU country issues an order to leave to – such as rejected asylum seekers – allowing researchers to separate such groups from for example the derogations mentioned in Article 2(2) of the Return Directive. This would lead to more comparable return rates.

2 Complete metadata & openness about data quality issues
As we have argued above, the Eurostat metadata does not provide sufficiently detailed information on how countries differ in either implementation of the Directive, or even in how they register and report return decisions issued and actual returns. For many countries, metadata is missing altogether. Detailed metadata for each country is necessary to assess the quality and comparability of data. For instance, each report should specify the types of return included in the data. Some of the data issues mentioned in this factsheet are inherent in the fact that data are not collected with a view to the scientific study of return rates, but rather are a by-product of the primary process. The degree to which data are 'cleaned' prior to reporting to Eurostat might differ across Member States – and contamination of the data is the problem that is hardest to pinpoint, as it largely constitutes a black box. A first step to deal with this issue would be to have more detailed quality reports, in which data issues are openly shared, ideally accompanied by an estimation of the effects of these data issues for the data quality.

   Some improvement of the metadata is arguably on its way: the new EU Regulation 851/2020 on EU statistics on migration and international protection – replacing the previous Regulation 862/2007 – gives more explicit attention to the quality reports Member States must provide on the data they supply to Eurostat. Member States should take measures to make sure that the data and metadata are of sufficient quality. Moreover, according to the new regulation, Member States should inform

Eurostat 'of any relevant information or change with regard to the implementation of this Regulation that could influence the quality of the data transmitted'. While the new guidelines place more emphasis on the quality of the data, it is now up to the Member States to commit to them in practice.

3 More detailed data
The data currently available on return decisions and actual return is quite limited. Most countries do not differentiate different types of return (e.g., forced vs. voluntary), nor is there complete information on where people return to. Though some Member States do provide more detailed information on the destination, these data are still subject to severe data issues (Eurostat, 2021). The new EU Regulation 851/2020 (mentioned above) will require Member States to disaggregate the data on actual return following an order to leave not only by citizenship, but also 'the type of return and assistance received, and by the country of destination'. These amendments to the previous Regulation will already give more valuable insight into return migration from the different EU Member States. It is important that these data submit to stringent data quality checks both within and across datasets, in order to safeguard data quality. Once return is disaggregated by type of return and assistance received, it will be possible to disaggregate forced, assisted and unassisted return – if the data quality is good enough. Given that the inclusion of 'registered unassisted voluntary return' remains optional, being able to exclude this group from the data will improve comparability across countries in the return rate.

Various scholars have shown that the distinction between forced and voluntary return is not so clear-cut in reality as it may seem in registrations; scholars question how 'voluntary' voluntary return actually is (Webber, 2011, Cleton & Chauvin, 2019), or speak of 'soft deportation' (Leerkes, Van Os & Boersema, 2017) as 'voluntary' return often contains forced elements, while there may be some measure of agency and voluntariness in forced return (e.g. because immigration detainees decide to cooperate with their return, such as by providing information about their identity). It is still valuable, however, to have separate data on 'forced' and 'voluntary' returns as it could allow researchers to assess whether certain return frameworks indeed mostly affect the forced return rate, as one would expect, or whether it also affects the more voluntary forms of return.

4 Cohort data
Finally, of course, to assess the actual return rate of third country nationals, cohort data would be extremely valuable. With cohort data, individual trajectories through the 'system' can be charted. Obtaining cohort data across all Member States over time for all third country nationals is unrealistic. Nevertheless, a case study of a few countries in which a cohort of TCNs would be followed from entry into the registration systems until their exit (either when they return, or when they can no longer be located), can provide valuable additional information about the complexity of routes that TCNs follow, and also shed further light on the underlying determinants of the routes.

Clearly, there are potential avenues for improvement of the EIL data. However, these steps will take time and considerable effort on behalf of the Member States. Transparency by both Member States and data providers such as Eurostat is indispensable for the accurate assessment of return migration.

6 Conclusion

The EU and individual Member States including the Netherlands have taken various measures to increase the return of irregularly staying TCNs, which include different types of bilateral and EU-wide return frameworks. To evaluate the effects of the European return policies, and to obtain more scientific insight in the determinants of enforced return, it is crucial that researchers can accurately measure rates of enforced return, which are comparable across EU Member States. Our methodological assessment of the current Eurostat data, which include data on return decisions and numbers of enforced return, indicates considerable methodological limitations in the measurement of enforced return. A significant portion of the methodological problems arises because of differences between Member States in the implementation of the Return Directive (e.g., issuing return decisions in different situations) and because of differences between Member States in the definition and registration of return decisions and enforced returns. Because of these inconsistencies, differences between EU Member States in measured return rates will therefore be clouded by international differences in the use and registration of return decisions and returns.

We have described several statistical techniques to increase the validity and reliability of measured rates of enforced return, and ways to explore the effects of bilateral and EU-wide frameworks given the data limitations that are identified. These methods include the deviation method; using data on rejected asylum applications to measure the population at risk (Leerkes & Van Houte, 2020); smoothing methods; calculating of rates for longer time periods; limiting the analysis to relatively comparable European countries; and using fixed effect methods to estimate the effects of (new) bilateral and EU-wide return frameworks, which we decided to use in part 2 of the study (see Leerkes, Maliepaard & Van der Meer, 2022). While these techniques mitigate the methodological issues, they cannot overcome them completely. It is therefore critical that researchers and governmental organisations that publish data on enforced return mention these limitations in some detail and reflect thoroughly on the implications for their findings and claims. Where possible, mixed-method designs are advised. Our overview of make-do techniques is not meant to be exhaustive; demographers have, for example, developed advanced statistical techniques to estimate true migration flows given inconsistent data on flows (cf. Raymer, 2017). These techniques are outside of the scope of the present project, but future research could examine whether such techniques can fruitfully be applied to the European data on enforced return.

Improving the methodological quality of the European data on enforced return is important. We have proposed several ways forward. Many of these rely on (increased) sharing of information and documentation between different Member States, for instance on the domestic use and registration of returns decisions – ideally also *retrospectively* so that better longitudinal data become available. Information sharing may also be a way to obtain more standardisation in the data that are provided to Eurostat. With a view to improving the documentation and standardisation of enforced return data, it might be worthwhile to connect with existing know-how and protocols of GOFAIR initiatives that aim to promote 'FAIR' (findable, accessible, interoperable, reusable) data (cf. Wilkinson et al., 2016). It may also be a good idea to involve national statistical offices with a view to improving the documentation and standardisation of data on enforced returns. Governments could also facilitate research in the field of migrant return, and learn more about the effects of their policies, by

making (anonymised) individual-level data and cohort data on returns available (cf. Leerkes et al., 2014, 2017), as well as data on preparatory bureaucratic procedures in the field of migrant return (e.g., on the outcomes of laissez passer requests or detentions with a view to expulsion). Finally, EU Member States should consider making more data available on the sustainability and human rights outcomes of enforced returns. The Return Directive (2008/115/EC) aims to 'provide for clear, transparent and fair treatment of irregularly staying migrants, while fully respecting the human rights and fundamental freedoms of the persons concerned', but there is still a dearth of data on the re-integration trajectories and human rights outcomes for migrants who are legally required to leave the EU.

References

Allison, P. D. (2009). *Fixed effects regression models*. SAGE publications.

Belmonte, M., Sermi, F., Tarchi, D. (2021). *How to measure the effectiveness of return? Problem definition and alternative definitions of the return and readmission rates (EUR 30491 EN).* Luxembourg: Publications Office of the European Union.

Carrera, S. (2016). *Implementation of EU readmission agreements: Identity determination dilemma and the blurring of rights.* New York: Springer.

Cassarino, J.P. (2007). Informalising Readmission Agreements in the EU Neighbourhood. *The International Spectator, 42*(2), 179-196.

Cleton, L., & Chauvin, S. (2020). Performing freedom in the Dutch deportation regime: Bureaucratic persuasion and the enforcement of 'voluntary return'. *Journal of Ethnic and Migration Studies*, *46*(1), 297-313.

Directive 2008/115/EC of the European Parliament and of the Council of 16 December 2008 on common standards and procedures in Member States for returning illegally staying third-country nationals. Retrieved from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:348:0098:0107:en:PDF. Accessed 12 May 2021.

Eisele, K. (2020). *The Return Directive 2008/115/EC, European Implementation Assessment.* Brussels: European Parliamentary Research Service.

Eurostat (2011). *European Statistics Code of Practice: For the national and community statistical authorities.* Luxembourg: Eurostat. Retrieved from https://ec.europa.eu/eurostat/documents/3859598/5921861/KS-32-11-955-EN.PDF/5fa1ebc6-90bb-43fa-888f-dde032471e15. Accessed 12 May 2021.

Eurostat (2015). *General compliance with definitions from EIL technical guidelines (unit of measure: persons).* Retrieved from https://ec.europa.eu/eurostat/cache/metadata/Annexes/migr_eil_esms_an1.pdf. Accessed 12 May 2021.

Eurostat (2021). *Eurostat metadata EIL: Enforcement of Immigration Legislation (migr_eil)*. Retrieved from https://ec.europa.eu/eurostat/cache/metadata/en/migr_eil_esms.htm#annex1610370771048. Accessed 12 May 2021.

European Commission (EC) (n.d.) *New Pact on Migration and Asylum: A fresh start on migration in Europe.* Retrieved from https://ec.europa.eu/info/strategy/priorities-2019-2024/promoting-our-european-way-life/new-pact-migration-and-asylum_en. Accessed 12 May 2021.

European Commission (EC) (2017). *Fifth Progress Report on the Partnership Framework with third countries under the European Agenda on Migration.* (COM(2017) 471 final). Retrieved from https://ec.europa.eu/home-affairs/sites/default/files/what-we-do/policies/european-agenda-migration/20170906_fifth_progress_report_on_the_partnership_framework_with_third_countries_under_the_eam_en.pdf. Accessed 12 May 2021.

European Commission (EC) (2020). *Migration.* Retrieved from https://ec.europa.eu/info/sites/default/files/new-pact-on-migration-and-asylum-package_1.pdf. Accessed 12 May 2021.

European Commission (EC) (2021). *Communication from the Commission to the European Parliament and the Council: Enhancing cooperation on return and readmission as part of a fair, effective and comprehensive EU migration policy.* (COM(2021) 56 final). Retrieved from https://ec.europa.eu/transparency/regdoc/rep/1/2021/EN/COM-2021-56-F1-EN-MAIN-PART-1.PDF. Accessed 12 May 2021.

European Court of Auditors (2021). *EU readmission cooperation with third countries: relevant actions yielded limited results.* Luxembourg: Publications Office of the European Union. Special Report No. 17/2021.
European Migration Network (EMN) (2017). *The effectiveness of return in EU Member States: Challenges and good practices linked to EU rules and standards – Synthesis Report*. Brussels: European Migration Network.
Houte, M. van, & Leerkes, A. (2019). *Dealing with (non-) deportability: A comparative policy analysis of the post-entry migration enforcement regimes of Western European countries: A Report commissioned by the 'Van Zwol committtee'.* Maastricht: UNU-Merit.
International Organization for Migration (IOM) (2019). *Glossary on Migration.* Geneva: IOM. Retrieved from https://publications.iom.int/system/files/pdf/iml_34_glossary.pdf. Accessed 12 May 2021.
Jurje, F., & Lavenex, S. (2014). Trade Agreements as Venues for 'Market Power Europe'? The Case of Immigration Policy. *Journal of Common Market Studies, 52*(2), 320-336.
Leerkes, A., Boersema, E., Galloway, M., Van Os, R., & Van Londen, M. (2014). *Afgewezen en uit Nederland vertrokken? Een onderzoek naar de achtergronden van variatie in zelfstandige terugkeer onder uitgeprocedeerde asielzoekers*. The Hague: WODC. Cahier 2014-3.
Leerkes, A., & Houte, M. van (2020). Beyond the deportation regime: Differential state interests and capacities in dealing with (non-) deportability in Europe. *Citizenship Studies*, *24*(3), 319-338.
Leerkes, A., Maliepaard, M., & Van der Meer, M. (2022). *Intergovernmental relations and return: Part 2: From paper to practice? EU-wide and bilateral return frameworks between EU+ and non-EU+ countries and their effects on enforced return*. The Hague: WODC. Memorandum 2022-2.
Leerkes, A., Os, R. van, & Boersema, E. (2017). What Drives 'Soft Deportation'? Understanding the Rise in Assisted Voluntary Return among Rejected Asylum Seekers in the Netherlands. *Population, Space and Place, 23*(8), 1-11.
Leerkes, A., Van der Meer, M., Paasche, E. & Brekke, J (2022) *Intergovernmental relations and return: Part 3: Beyond return frameworks: An exploration of Dutch and Norwegian intergovernmental strategies to implement enforced return to Afghanistan, Iran, and Iraq*. The Hague/Oslo: WODC/ISF. Memorandum 2022-3.
Pluim, M., & Hofmann, M. (2016). *Rethinking Post-Arrival Assistance: The Road Towards a European Return System* (ICMPD Working Paper 10). Vienna: ICMPD.
Raymer, J. (2017). *Measuring flows of international migration*. IZA World of Labor 2017: 354.
Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection and repealing Council Regulation (EEC). Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32007R0862&from=EN. Accessed 12 May 2021.
Regulation (EU) 2020/851 of the European Parliament and of the Council of 18 June 2020 amending Regulation (EC) No 862/2007 on Community statistics on migration and international protection. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32020R0851&from=EN. Accessed 12 May 2021.
Scheel, S., & Ustek-Spalda, F. (2019). The politics of expertise and ignorance in the field of migration management. *Environment and Planning D: Society and Space, 37*(4), 663-681.
Vara, J. S. (2019). Soft international agreements on migration cooperation with third countries: a challenge to democratic and judicial controls in the EU. In

S. Carrera, J.S. Vara & T. Strik (Eds), *Constitutionalising the External Dimensions of EU Migration Policies in Times of Crisis* (pp. 21-38). Cheltenham: Edward Elgar Publishing.
Vaus, D. de (2001). *Research Design in Social Research.* Thousand Oaks: SAGE Publishing.
Webber, F. (2011). How Voluntary are Voluntary Returns? *Race & Class, 52*(4), 98-107.
Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data, 3*(160018), 1-9.

The WODC (Research and Documentation Centre) is the knowledge centre in the field of the Dutch Ministry of Justice and Security. The WODC carries out independent scientific research for policy and implementation purposes; it does so both by itself, and on the WODC's commission.

More information:

www.wodc.nl