CAMBRIDGE
UNIVERSITY PRESS

TRANSLATIONAL ARTICLE

Migration data collection and management in a changing Latin American landscape

María E. Cervantes-Macías* (iD)

Latin American Open Data Initiative, Montevideo, Uruguay
Department of Geography, University of British Columbia, Vancouver, Canada
*Corresponding author. E-mail: mariaesther@idatosabiertos.org

Abbreviations: COMAR, Mexican Commission for Assistance for Refugees; IDB, International Development Bank; INEGI, The National Institute of Statistics and Geography; INM, National Migration Institute; OKFN, Open Knowledge Foundation; PEP, Special Residency Permit; TMF, UN, Border Mobility Card; UNHCR, United Nations High Commissioner for Refugees

Abstract

While many Latin American countries have a tradition of receiving migrants, including the countries selected as case studies, there are no institutionalized mechanisms for the integration and settlement of migrants. The objective of this article is to explore how to improve migration data collection and management in a region that does not have many migration integration policies in place. I assess the state of migration data collection and management in three case studies: the city of Cucuta in Colombia, the North Huetar Region in Costa Rica, and the city of Monterrey in Mexico. The three countries publish data exclusively at the national level, rather than the local or municipal. Despite all case studies having a variety of administrative data, mainly in the form of entries and exits by nationality, these data are not enough to properly identify the sociodemographic characteristics of migrant populations in a country, and much less in specific cities. I make recommendations divided into three main themes to improve migration data in Latin America.

Policy Significance Statement

Latin America is going through a process of regional transformation as new migration flows are intensifying in the region. Despite a tradition of welcoming migrants and refugees in Latin America reflected in the signing of documents such as the Cartagena Declaration, there are no formal integration policies or settlement institutions in Latin American countries. This research argues for the importance of creating and publishing high-quality, open migration data beyond entries and exits to a country to facilitate the creation of group-targeted integration policies. This article evaluates the state of migration data governance in three case studies in the region and presents a series of recommendations for improvement.

1. Introduction

Migration is a major challenge for Latin America and should be approached from a multifactorial perspective. The region, and in particular countries with high levels of immigration and emigration, face many challenges, among them: managing increased migrant flows, increased asylum petitions, and

CrossMark

returns from their migrant populations. Several countries in Latin America are traditional places of passage for migrant communities aiming to reach the United States through Mexico in the north of the continent and Argentina and Brazil in the south. Other countries have long traditions of migration among neighbors such as Nicaraguans in Costa Rica and Venezuelans in Colombia.

Despite such migratory traditions in the region, there are currently no public policy frameworks for the integration of migrants in many of the countries. The objective of this article is to explore how to improve migration data collection and management in Latin America. I do this with three case studies: the city of Cucuta in Colombia, the North Huetar Region in Costa Rica, and the city of Monterrey in Mexico. The rationale for choosing these case studies to provide a variety of border and quasi-border cities, with diversity in terms of historical international migration patterns and new international migration patterns, and because they are each industrial and agro-industrial hubs in their respective countries, making them attractive destinations to seek employment and settlement for migrants. Having sociodemographic data can provide better access to services for migrants, regardless of temporary or permanent settlements in the border cities.

While many Latin American countries have a tradition of receiving migrants, including the countries selected as case studies, there are very few institutionalized mechanisms for the integration and settlement of migrants. In this article, I assess the state of migration data collection and management in each case study and explore some of the best practices for migration data collection and management from countries with a long history of integration policies. I argue that having data that tells us how many immigrants are in the country or specific cities, where they come from, their sociodemographic profiles, who they travel with, and their objective from migration, can help change traditional narratives about migration in the region.

Migration data governance can improve through better practices of data accessibility and sharing. Better data management through the use and publication of open and structured migration data can contribute to improving migration integration policies, which are necessary as international migration increases in the region. Integration policy can be defined in many ways, but for this article, I follow Banting's (2014) multicultural conception of integration, in which the expectation is that immigrants will choose to visibly express their ethnic identity and accept an obligation on the part of public institutions to accommodate their distinctiveness. Having high-quality, accessible, comparable, and interoperable data is the first step in understanding who are the migrants entering the region to create integration policies for those who decide to settle in Latin American countries. However, it is important to consider that despite the possible benefits of having more data about migrants, there are some risks associated with migrants' data collection, particularly those related to the rights and privacy of migrants and vulnerable groups, which I also explore further in this article.

To understand how to improve migration data collection and management, I must first define what I mean by migration data. I follow the International Organization for Migration's (IOM) migration data categorization. The IOM uses four types of migration data in its data portal: census data, innovative data sources, survey data, and administrative data sources.

For most countries, administrative data sources might be the principal source of migration data. According to the IOM, "statistics derived from data in these sources usually refer to administrative records rather than people" (Migration Data Portal, 2021). Some examples of sources include administrative registers, border data collection systems, visas, residence permits, and/or work permits. While certainly the most common type of migration data, it is oftentimes limited and does not include sociodemographic data that could be useful for the creation of integration policies.

Survey data are defined by the IOM as household surveys that produce statistics on aspects, such as the drivers and the impact of migration, internal migration, socioeconomic characteristics, emigration, migrant stocks and flows of immigrants and emigrants, as well as labor force surveys that produce statistics on migrant stocks in the labor market. This is data collected by governments through national or local surveys. Survey data can also be collected by academic institutions or civil society/religious organizations, in either research studies or as part of their service delivery. However, these data are much

less likely to be accessible, since these organizations are often dispersed and lack an online presence. These organizations are usually understaffed and under-resourced and have to prioritize those resources to provide humanitarian assistance, rather than to improve their documentation processes (Gabarrot and Cervantes, 2021). Furthermore, organizations and migrants distrust governmental authorities and external researchers, since they represent a risk to their clients and/or their operations. Here, there is a lack of clear mechanisms for data-sharing practices between civil society organizations and the government. According to the Global Data Barometer (2022), the majority of participating countries have partial or no data-sharing mechanisms between the government and other sectors. This demonstrates a fundamental imbalance in the power dynamics of data sharing between civil society organizations and the government.

The third category of migrant data considered by the IOM is census data. A national census might have some migration data included in their questionnaires, depending on the country. The IOM argues that population censuses collect migration-related data which can be used to produce official statistics on migrant stocks, socioeconomic characteristics, migrant flows (limited), and some emigration figures. Census data can be extremely useful as it provides information on everyone, at the same time, from the whole country and is repeated regularly (Baffour et al., 2013). However, a major limitation is that, given the size of a census, the information is usually limited to 5 or 10 year periods, and does not capture rapidly changing trends such as short, rapid increases in temporary migration.

Lastly, the IOM considers innovative data sources as a new source of migration data. The IOM provides some examples such as big data sources like mobile phones, online tools, and platforms, including social media and online payment services, as well as other data sources such as public opinion on migration data. These data are more consistent across platforms and can be less accessible than others, especially big data sources that are owned by private entities. Because migration is an activity that encompasses many different aspects of an individual's life, oftentimes data that might be useful for migrant's integration, will not be classified as migration data. These other forms of data can be considered innovative data sources. Some additional examples of these data are school enrollments, access to health services data, housing availability, among others. If this data exists, it may be owned by civil society organizations or local government offices that do not make the data public. The variety of stakeholders involved in the collection and publication of migration data makes the study of migration data collection and management even more critical.

1.1. Migration data collection and management

In this article, I argue that to improve migration data collection and management for the creation of migrant integration policies in Latin America, governments should publish structured, interoperable, and comparable migration data in the region, commonly known as open data. The Open Knowledge Foundation (OKFN) defines open data as:

> *Open data can be freely accessed, used, modified, and shared by anyone for any purpose—subject only, at most, to requirements to provide attribution and/or share-alike. Specifically, open data is defined by the Open Definition and requires that the data be A. Legally open: that is, available under an open (data) license that permits anyone freely to access, reuse and redistribute B. Technically open: that is, that the data be available for no more than the cost of reproduction and in machine-readable and bulk form (OKFN, 2020).*

There is also a rating system proposed by Berners-Lee (OKFN, 2020), which establishes a five-star mechanism to evaluate the openness of data. These stars can be obtained by fulfilling the following criteria: (a) data must be available on the Web under an open license, (b) be in the form of structured data, (c) be in a nonproprietary file format, (d) use URIs as its identifiers, and (e) include links to other data sources. While the rating system is a set of guidelines, following the criteria makes the data more accessible and reliable.

Another improvement to migration data collection and management could be the use of data standards. The OKFN defines data standards as:

A published specification for, e.g., the structure of a particular file format, recommended nomenclature to use in a particular domain, a common set of metadata fields, etc. Conforming to relevant standards greatly increases the value of published data by improving machine readability and easing data integration (OKFN, 2020).

Improving data collection and management through the use of open data and data standards can contribute by allowing policymakers and scholars with statistical information to create and promote immigration integration and settlement policies in Latin America. While having better and more accessible data on migrants is a necessity for creating better public policies for their reception and integration, it is important to remember that a good data collection and management of migrant data has to prioritize the safe keeping of the information of these vulnerable populations. Reflecting on the state of migration data collection and management in three cities in Latin America provides a starting point to identify the challenges and make recommendations for improvement by assessing the state of their migration data collection and management structures at the national and local levels.

2. A Changing Regional Landscape

According to the Inter-American Development Bank (IADB), more than 40 million people from Latin America and the Caribbean currently reside outside their countries of origin. Intraregional migration in Latin America has increased significantly in recent years due to government crises in countries such as Nicaragua and Venezuela, and the increasingly strict immigration restrictions in the United States of America. According to the South American Migration Report (2017), in 2015 there were 5 million immigrants in South America, although these data were collected before the Venezuelan migration crisis of 2016. The Inter-agency Coordination Platform for Refugees and Migrants from Venezuela (2022) reports that around 6 million Venezuelans have been displaced, and close to 5 million are located in Latin America. The combination of the Venezuelan crisis with traditional migration patterns across the region necessitates a distinction between voluntary and forced migration.

There are specific legal responsibilities receiving states have to people in conditions of forced migration. However, voluntary migration has its challenges, particularly in terms of integration and settlement. Most of the foundational literature on migrant integration, which is primarily produced in traditional destination countries such as the United States or Canada, focuses on a debate of integration versus assimilation (Portes and Zhou, 1993; Putnam, 2007; Vertovec, 2007). This debate is not necessarily applicable in a Latin American context, since migration is usually intraregional. However, without easily accessible, up-to-date migration data, it is hard to identify if these migration patterns have changed. Data is essential to understanding and responding to any changes in traditional migration patterns. One way in which this research contributes to the literature on migration in Latin America is by identifying the state of migration data collection and management in various border cities and regions, to then make recommendations for improvement, in the hopes of encouraging the development of migration data collection and management structures to facilitate migrants integration in the long term.

Despite recent changes made at the beginning of the decade that extended protections to migrants and particularly refugees in the region, in recent years, the goodwill expressed through these regulatory changes has been confronted by an unprecedented increase in migration in the region. For example, Colombia and Peru, have received more than 1.8 million Venezuelan immigrants and refugees, thus implementing massive regularization campaigns to grant them temporary legal status (Seele, 2019). In a report by the Migration Policy Institute, Selee et al (2019) argue that countries such as Colombia, Peru, Ecuador, Argentina, Chile, Brazil, Panamá, and Mexico developed creative solutions to face the Venezuelan migratory crisis. Some examples are new legal permits (i.e., Special Stay Permit in Colombia, Temporary Stay Permit in Perú, and the Democratic Responsibility Visa in Chile), regularization

programs (i.e., Administrative Registry of Venezuelan Migrants in Colombia), and expanding the use of existing visas (i.e., UNASUR visa in Ecuador). Some of these solutions were designed for the short term, and others offer pathways toward permanent residence for Venezuelan nationals.

Many Latin American countries have traditionally welcomed refugees in the region, including through processes outlined in the Cartagena Declaration. Although the Cartagena Declaration is nonbinding, countries must decide whether they will assume the responsibilities conferred by this Convention, or make their immigration policies more restrictive. An example of this is Mexico's use of the Cartagena Declaration for Venezuelans but not for Central Americans (Sanchez and Freier, 2021). Making this decision is increasingly complicated due to the economic crisis and border restrictions generated by the ongoing COVID-19 pandemic. Challenges faced by the general population have exacerbated negative reactions on the part of local populations toward migrants. An Oxfam (2019) study carried out in Ecuador, Peru, and Colombia in 2019, shows that 70% of those surveyed support greater restrictions on border crossings with Venezuela, while 70% perceive newcomers as a threat to job stability.

Although much of the recent literature on migration in Latin America concentrates on the Venezuelan diaspora, some researchers are identifying Caribbean (Audebert, 2017) and extra-continental (Yates, 2019) migratory flows in Latin America that have different sociodemographic characteristics than Venezuelan migration, and thus have diverse integration needs. Changing sociodemographic characteristics of migrants, in combination with the ongoing effects of the COVID-19 pandemic, heighten the importance of having good quality, easy-to-access migration data to understand who migrants are and to create public policies that contribute to their integration and well-being in the societies that receive them.

Sociodemographic data, in particular, is essential to create group-targeted integration policies, by identifying the specific needs of particular populations, as immigrant populations cannot be treated as a homogenous group. In an extensive review of integration policy outcomes, Bilgili et al. (2015) find that outcomes differ significantly for different immigrant populations depending on countries of origin, ages, genders, or reasons for migration. While some of the literature disagrees between the success of group-targeted and mainstreaming integration policies (Benton et al., 2015), these discussions are happening in Europe and other regions in the Global North, which have a long history of implementing integration policies. Such discussions are very recent in Latin America, and the contexts are widely different than those in Europe or the United States and Canada.

In the next section, I explore three case studies in different Latin American countries to assess their current migration data collection and management. I will evaluate the state of migration data collection and management using Berners-Lee's five-star rating system for open data.

Data must be available on the Web under an open license
Be in the form of structured data
Be in a nonproprietary file format
Use URIs as its identifiers
Include links to other data sources

2.1. Scope of the analysis

The case studies presented in the next section were crafted by performing a review of the data published by a variety of national migration portals, as well as municipal websites, Supplementary Material presents a breakdown of the websites which were revised. This review was done by downloading each individual file and identifying the different types of data published, as well as the formats in which the data was published. In some cases, open data portals at the national and state levels were also available. This review was performed during the summer of 2020.

While the focus of the case studies is on cities, cities are not administrative units, which resulted in a lack of city-specific information. Local governments are essential to migrants as they "have broad powers

to establish their own local regulations and projects, as long as these do not contravene national legislation" (FAO, 2022). In the case of immigration policy, local governments have been essential in providing access to services and protections to migrants in countries such as the United States. By settling temporarily or permanently, migrants become part of the community and transform it. The case studies focus on cities as the municipal or local governments are the ones providing migrants with the services they need, having to rely on the state, province, or national government to provide data and limited resources. Future research could explore the ways in which city governments respond to increases in migration flows in regards to data collection and capacity building for local authorities.

3. Case Studies

3.1. Cúcuta, Colombia

Cucuta, Colombia is the capital city of the Norte de Santander department, in the northeast of the country and close to the border with Venezuela. It has a population of about 1 million inhabitants in its metropolitan area. The 2007 free trade agreement between Colombia and the United States contributed to the industrial development of this city.

Haddad et al. (2018) explain that historically the Colombian–Venezuelan border has been recognized for its closeness and binational commerce. According to Mudarra and Ortiz (2019), of the total migrant population that enters through land border posts, 94% enter through the Simón Bolívar bridge in Cucuta, which makes it the busiest crossing point in Colombian territory. Since the onset of the Venezuelan economic crisis, about 200,000 Venezuelan nationals have immigrated to the area, according to the UN (2018). Starting in 2005 due to unemployment in the Venezuelan oil sector and with peaks in 2017 and 2018 due to the political instability of the country (Haddad et al., 2018, p. 5) as well as an increase in human rights violations (Amnesty International, 2021). Horta and Rossiasco (2019) explain that although there is a higher concentration of Venezuelan national migrants in Bogotá, in relative terms, the municipalities along the border, such as Cucuta, are receiving a greater number of Venezuelans compared to other regions.

3.1.1. Migration data collection and management in Cucuta

In the case of Cucuta, Migración Colombia is the federal entity charged with collecting and publishing migration data. In a document outlining how they create their statistics, Migración Colombia states that their objective is to "produce official statistical information on the entry and exit of Colombians and foreigners to the country through the 44 Immigration Control Posts enabled by Migración Colombia" (2019). They also have the specific objective of ensuring the precision, accuracy, interoperability, and comparability of administrative records. The only mention of sociodemographic data in this document is in regards to surveys conducted by the institute. This documentation demonstrates that sociodemographic data are not a priority for Migración Colombia, which is reflected in the type of data published by the institute.

Taking their objectives into account, I focus on the quality of the data that they do publish. Most of the data can be found in a public Tableau called "Migratory Flows." Tableau is a tool that facilitates the creation of interactive dashboards with large quantities of data, and also allows for the structured data to be downloaded in a nonproprietary format. Migración Colombia's data fits 3 out of 5 stars in the open data five-star rating system. Additionally, the data is updated monthly. This Tableau is a great example of structured, accessible data, and it presents data on: month of entry, entry transportation, nationality, age range, and authorized activity in the country.

Migración Colombia also publishes annual Statistical Bulletins as reports in PDF format. In these reports, the document data is in four main categories: general information about migratory flows, entries, and exits of foreigners, entries, and exits of Colombians, and data about Venezuelans in Colombia specifically. These Statistical Bulletins are not open data, but some of these data can be found in the Public Tableaus published by Migración Colombia.

Despite the natural focus on Venezuelan migration, there is not enough data to know what their specific needs are. However, there are some specialized survey data. For example, in 2019, Migración Colombia published survey results about Venezuelans' experience with the Special Residency Permit (PEP), where there is data on PEP benefits, types of visas, education level, current job, and recent mobility of this migrant population.

Despite the numbers of migrants living in Cucuta, the only available data that is specific, is data on entries and exits from Migración Colombia. The city of Cucuta does not have a website or an open data portal where they could publish their data. Overall, Colombia's use of dashboards is an innovative way to showcase migration data. Additionally, as most migrants in Colombia are from Venezuela, having a focus on Venezuelan migration data is a good first step to create integration policies geared specifically at this population.

3.2. Huetar Norte Region, Costa Rica

Costa Rica and Nicaragua share 312 km of common border, with a large migrant population and international transit. The North Huetar Region is administratively composed of the municipalities of San Carlos, Upala, Guatuso, and Los Chiles, as well as the districts of Sarapiquí that belongs to Heredia, Río Cuarto of the Canton of Grecia and Peñas Blancas of the Canton of San Ramón. Morales et al. (2010) explain that in terms of demographic dynamics on the border, two trends stand out: a growing population that concentrates on urban agglomerations and the intensification of cross-border flows, mainly of migrant labor (Morales et al., 2010, p. 5). Morales et al. (2010) explain that although the majority of the migrant population concentrates in the Central Region of Costa Rica, in relative terms the greatest concentration can be observed in the North Huetar Region, with 14% of inhabitants being of Nicaraguan origin.

3.2.1. Migration data collection and management in the North Huetar Region

The General Directorate of Immigration and Foreign Nationals is the organization in charge of collecting and publishing migration data. They publish three main categories of data and an annual report. These categories are migratory flows, foreign nationals, and professional migration police. The migratory flows category consists of monthly PDF reports on entries and departures of Costa Ricans and foreigners according to the port of entry and nationality.

The foreign nationals category is published as an Excel file, titled "Migratory Categories" which contains data on entries, exits, requests for permanent residence, denials of permanent residence, requests for refuge, denials, approvals of refuge, and transit requests. The professional migration police category presents an annual report on deportations by nationality and month, arrests and expulsions, and entries to migrant detention centers administered by the police. Lastly, the annual report is published as an Excel file compiling the data from all three categories, in addition to passports expedited by Costa Rica's consular offices, the number of "special" residents, and permanent residents by nationality in the year.

In addition to the data published by the General Directorate of Immigration and Foreign Nationals, some migration data can also be found in the National Census and other specialized surveys, which is the responsibility of the National Institute of Statistics and Census of Costa Rica. Most data on migration is part of the National Household Survey. Given that the North Huetar Region has no administrative organization, the only data available for this region is that of the municipalities that are part of the region. Some of these municipalities have open data portals, such as the cities of San Carlos and Guatuso. However, these local jurisdictions do not publish any migration data.

Overall, the General Directorate of Immigration and Foreign Nationals data is administrative data that is kept relatively up to date. However, it does not fit any of the criteria established by the five-star open data rating system. The Census data is published in Excel and does not fit any criteria in the five-star open data rating either. The data is semi-structured, and while some of the categories might be comparable between themselves, the data would require cleaning and reordering to make any comparisons and

identify trends. There is no sociodemographic data published by these institutions, except for the nationality of entries and departures.

3.3. Monterrey, Mexico

Monterrey is the capital of the state of Nuevo León, and one of the largest cities in Mexico with more than 4.6 million inhabitants in its metropolitan area according to the 2020 Population Census. Monterrey concentrates on the largest industrial and service activity in the country, along with the Valley of Mexico and is located 250 km from the United States border.

The increase in international migration flows is a recent phenomenon in the city. This increase is related to the binational response by Mexico and the United States, mainly with the U.S.'s implementation of the Migrant Protection Protocols (MPP) agreement beginning in March 2020 at the start of the COVID-19 pandemic (Sedas et al., 2020). The Department of Homeland Security defines the MPP as:

> a U.S. Government action whereby certain foreign individuals entering or seeking admission to the U.S. from Mexico—illegally or without proper documentation—may be returned to Mexico and wait outside of the U.S. for the duration of their immigration proceedings where Mexico will provide them with all appropriate humanitarian protections for the duration of their stay (Department of Homeland Security, 2019).

The MPP has an effect on Monterrey, as Monterrey is an attractive destination to find employment, whether formal or informal, due to its reputation as the industrial capital. Many migrants choose or are made to wait in Monterrey while waiting for admission to the United States. Previously, Monterrey was characterized by internal Mexican migration patterns, which were usually permanent (Ybáñez Zepeda and Lara, 2017). Monterrey has an established network, albeit with limited resources, of migrant shelters and civil society organizations that provide support to migrants. In 2018, the United Nations High Commissioner for Refugees (UNHCR) and the Mexican Commission for Assistance for Refugees (COMAR) opened offices in the city. The presence of migration shelters and institutional offices in the city demonstrates that there is an increased demand for their services, as well as an increased need for access to basic services such as housing, health, and education. However, it is difficult to find up-to-date migration data to properly estimate how many migrants are in the city and need access to these services.

3.3.1. Migration data collection and management in Monterrey

The Mexican federal government and the different entities involved in migration data collection and management in the country publish some data about migrants in Monterrey. The National Institute of Statistics and Geography (INEGI, 2020) publishes migration data at the national and state/provincial level, but not disaggregated by city.

INEGI publishes the following information about international migration: population born in another country by sex and nationality, resident population born in another country, migrant population by country, and returned migrant population. INEGI also publishes Census data, including reasons for migration, number of immigrants by nationality, and cause of migration. The Mexican government also publishes migration data on the Open Data portal of the Federal Government, as a responsibility of the National Migration Institute (INM). The data published in this open data portal is the following: conditional stay authorizations, assistance to foreign victims of crime, the exit of minors, percentage of trained public servants, immigration reviews, verification visits, requests of help from the Paisano Program, complaints received by the Paisano Program, migratory processes, the assistance provided by the Paisano Program, migrants assisted by Migrant Protection Beta Group and immigration documents. However, these data are not frequently updated, the last update was in 2018 for these resources.

The INM also publishes a monthly Statistical Bulletin, providing data on documentation and conditions of stay in Mexico, entry registry, aliens presented and returned, migrant protection groups,

and repatriation of Mexicans and Mexicans returned by Canada, with many subcategories. The INM's monthly Statistical Bulletin has the highest rate of consistency and frequency in their publication, in Mexico and out of the three case studies as well. However, the only publishing entity that fits two of the five starts of open data is the INM, with some data published with an open data license and links to other sources. All the information is published in monthly PDFs and an annual PDF at the end of the year, and some data included in the Statistical Bulletin is disaggregated and published in Excel format. Additionally, the INM publishes a PDF report which compiles all of the data published by the INM each year. While useful, the data presented in the Statistical Bulletin is not in open data format, which makes it difficult to analyze.

In terms of local migration data collection and management, despite the steady increase of migrants, the government of the city of Monterrey does not publish any data on migrants. The state of Nuevo León, of which Monterrey is the capital, does not publish data on migrants either. The open data portal of the state of Nuevo León, managed by the Commission for Transparency and Access to Information of the State of Nuevo León (COTAI) does not have migration datasets. The COTAI has a chatbot to provide rapid assistance and using this function they confirmed that INM is the organization in charge of these data, while the Mexican government's data published by the INM or INEGI, particularly as part of the monthly statistical bulletins, is comprehensive and up to date, it is not enough to build sociodemographic profiles at the local level that could be useful to create integration policies. Additionally, the data is not structured, which restricts its interoperability and compatibility.

The next table presents a summary of how each case study fits into the five-star rating system for open data:

Colombia, Monterrey	Data must be available on the Web under an open license
Colombia	Be in the form of structured data
Colombia	Be in a nonproprietary file format
None	Use URIs as its identifiers
Monterrey	Include links to other data sources

4. Discussion

4.1. *Migration data collection and management in Latin America*

In this section, I build on the findings of each case study to comment on the state of migration data collection and management in these cities and regions. I return to the case studies to make comparisons between the state of migration data collection and management in the case studies with particular attention to the interactions between local and national data collection systems. Additionally, I look at migration data collection and management in other countries to identify best practices on data collection and use that could be replicated in the region.

The three countries publish data almost exclusively at the national level, rather than the local or municipal, and most data are about entries and exits from each country. In terms of demographic data, most data center on nationality and, in some cases, sex. However, sociodemographic data on migrants is extremely limited, which complicates the ability of local governments to create targeted policies that facilitate the integration of specific immigrant populations into the receiving society.

While I will not talk in length about the debates between group-targeted and mainstreaming integration policies, it is important to highlight that Latin American countries have an additional challenge when considering widespread inequality and challenges in access to basic services for their citizens. Any changes in traditional migration patterns will add demand for access to basic services, stretching already underfunded services. In some cases, federal governments offload service provision responsibilities to local governments, without providing additional resources despite migration being a national government responsibility. As an example from one of the case studies, the mayor of Monterrey, Luis Donaldo Colosio, recently mentioned that the city is trying to work on a migrant data registry:

(Create a migrant data registry) to have a record of all the people who are going to be arriving in our municipality, which is supposed to be much fewer people than those originally said. All they want is to get here to be able to work and move in a way not necessarily as refugees, but simply they can wait here while they are given a visa to be able to go to work in the United States (Maldonado, 2022).

Statements like this showcase that having up-to-date information on migrant populations at the local level is essential for the creation and coordination of group-targeted integration policies. The case studies showcase how current migration data collection and management structures stem almost exclusively from national institutions, as migration is considered a national policy. Demonstrating this need, the National Administrative Department of Statistics (DANE) from Colombia, started publishing a Geoportal with International Migration Statistics in June 2022. The data encompasses a wide range of subjects related to migration and can be shown in several geographic scales, from national, to provincial and even municipal. This initiative has the objective of having an Integrated Statistical Registry of Migrant Population to provide information and timely statistics of the migrant population and give.

Migration data can be used by a wide variety of stakeholders, including national governments, local governments, international organizations, national civil society organizations, local civil society organizations, religious organizations, researchers, and migrant organizations. These stakeholders have different data needs and widely differ in access to economic, logistic, technical, and human resources. In the case of local stakeholders, the data are oftentimes used to understand the immediate needs of recent arrivals and to prioritize resources into the most needed areas. In their study of municipal immigrant integration, Rodriguez et al. (2018)) find that state and local governments in the United States have increased their efforts to address immigration issues primarily as a response to the lack of movement at the federal level to address communities' need for comprehensive immigration reform.

However, increased data collection as a response to increased migration flows with no increased resources for training, data security protocols, safe data storage, and so forth, could potentially result in the misuse of the data collected. Another concern, raised by Molnar (2020) is the role of the private sector in the collection, use, and storage of migration data. Molnar (2020) uses Facebook and Twitter as an example, private companies that might be called by the government to share tools or know-how to conduct social media intelligence for immigration enforcement purposes. While I could not identify this happening in the case studies, it is important to keep in mind that without proper training, state and local governments creating last-minute solutions, for example, a Google Form to collect private data to provide services, to their immigrant integration challenges could become risky for the individuals whose data are being collected.

In terms of migration data, the focus of the case studies has been mainly on national governments' migration data as these are the entities that have the most resources to collect their migration data, and primarily administrative data. These administrative data can be found mainly in the form of entries and exits by nationality; this data is not enough to properly identify the sociodemographic characteristics of migrant populations in a country, and much less in specific cities. Administrative data usually measure events, such as entries and exits, and not people (González Arias, 2017; Torre-Cantalapiedra and Yee-Quintero, 2018) and in many cases, these collection processes reinforce the assumption that all migrants in transit through a country will leave the country at some point rather than settling and integrating into the locale.

In terms of institutional actors involved in the migration data collection and management in each case study, we can identify that the principal actors in each case study are the INMs, specifically: Migración Colombia in Cucuta, the General Directorate of Immigration and Foreign Nationals in Costa Rica, and the National Institute of Migration in Mexico. The second category of actors involved in migration data collection and management is the National Statistical Institutes of each country: the National Administrative Department of Statistics (DANE) in Colombia, the National Institute of Statistics and Census (INEC) in Costa Rica, and INEGI in Mexico. While useful, census data reflect long-term trends, rather than the current state of affairs. Lastly, some of the local governments had their open data portals, however, none of these open data portals publish migration, or at least, not public migration data.

The current migration data collection and management structure creates additional challenges for local governments struggling to provide migrants with access to basic services, which results in an overreliance on civil society organizations and international organizations with limited resources. For example, Venezuelan organizations such as the Venezuelan Foundation and the Nueva Ilusión organization provide support for their compatriots in Cucuta. In Costa Rica, academia produces survey data in collaboration with civil society, with entities such as the Central American Borders Research Unit monitoring migration patterns. In Monterrey, organizations such as the Documentation Network of Organizations for the Defense of Migrants (REDODEM) make an effort to assist migrants as they conduct surveys and intake interviews to better monitor the sociodemographic features of people in situations of mobility, their migratory flows, security challenges, and human rights violations in their immigration process.

The different actors involved in migration data collection and management have different mandates and interests which reflect how they collect and use such data. While this article advocates for the importance of obtaining sociodemographic data on migrant populations to create context-specific public policies, I also recognize that publishing sociodemographic data on vulnerable populations might not be desirable, as it increases the risk of identification if proper data protection protocols are not in place. However, sociodemographic data is necessary to understand what are the specific needs of particular migrant populations, so there has to be a balance between addressing the data needs and enforcing mechanisms to keep these data safe.

5. From National to Local: Recommendations for Local Migration Data Collection and Management

In this section, I present a series of recommendations to improve migration data collection and management in Latin America based on the case studies discussed above. The five-star open data rating system that I used in the case studies is useful to evaluate all types of open data. The recommendations are divided into three main themes: structured data, group-targeted data, and capacity building.

5.1. Structured data

Structured data is an essential component of open data. However, in regards to migration data collection and management, it can also contribute to protecting the privacy of migrants' identity and therefore risk of sharing such information with authorities. The lack of structured data can present a risk to migrants as it could be easy to identify a person based on the data provided in some of these datasets. For example, in the Migratory Processes dataset, from Mexico's INM, the data shows age range, type of migration process, nationality, and state (by gender) in a single line. This result in the column showing a single person that fits that criteria per row, instead of structuring data to prioritize the anonymization and privacy of individuals.

An example of a best practice is New Zealand's Migration Data Explorer. The Migration Data Explorer is a single interface for migration statistics that provides access to migration data. The Ministry of Business, Innovation, and Employment of New Zealand describe it as "a new conceptual framework for analyzing migration trends based on determining migration's impact on the population and labor supply, which moves away from existing administrative metrics such as visa approval numbers" (Ministry of Business, Innovation and Employment, 2022). All the data in the Migration Data Explorer is structured in a way that the data explorer can use different datasets to sort and filter according to what the user is looking for. Users can choose a dataset, time period, up to four variables (such as demographic characteristics) and add filters to disaggregate by category.

5.2. Group-targeted data

Given that most migration data published in the case studies center on entries, exits, and other administrative data, it is important to start collecting data that, if published, could provide more information on the conditions and profiles of migrants. Some data fields can be nationality, age, education

level, primary language, secondary language, gender identity, current employment, migration with family, type of current housing, access to potable water, access to storm drainage, and access to electricity. Obtaining data on the characteristics and living conditions of migrants can contribute to creating group-targeted integration policies, which have worked in other countries. In contrast to entries and exit data, sociodemographic data requires more effort and resources across time and could potentially be incorporated into local governments providing immigrants with access to services.

For example, Canada has a Longitudinal Immigration Database (IMDB), which Statistics Canada defines as "an annual Canadian dataset that allows users to study the characteristics of immigrants to Canada at the time of admission and their economic outcomes and regional (inter-provincial) mobility over a time span of more than 35 years" (IRCC, 2021). Additionally, The IMDB includes Immigration, Refugees and Citizenship Canada (IRCC) administrative records with information about immigrants who were admitted to Canada since 1952.

These extensive data collection exercises require an investment in economic and human resources and having more data about migrant needs will allow for the allocation of resources for targeted interventions. However, having these data allows countries such as Canada to examine and improve the performance and impact of immigration programs across time, and the quality of life of immigrants in Canada.

Migration open data should also be updated frequently, particularly given the transitory condition of many migrants in Latin America. While entries and exit data are updated monthly in all of the case studies, besides these data, there is no up-to-date data on migrants' characteristics, access to services, or even the number of migrants in a particular city. Innovative data sources, such as big data, could potentially contribute to increasing the amount of information available. Tjaden (2021) argues that where administrative data sources, surveys, and censuses are not available, not accessible, or too slow, the "three V's" of big data: volume, velocity, and variety can be useful. However, this would require economic and technical resources that might be hard to obtain for local or national governments in Latin America.

5.3. Capacity building

In Latin America, many civil society organizations are the de facto providers of assistance and humanitarian relief to migrants and refugees. While these groups have a wealth of information and expertise, they lack the resources to prioritize long-term investments in capacity building, specifically in regards to improving migration data collection and management processes (Gabarrot and Cervantes, 2021). This is particularly true as civil society groups frequently have to prioritize the use of their resources for immediate humanitarian assistance rather than long-term training or collection efforts. In the case of migration data collection and management, I follow Mason and Fiocco's (2017) arguments that the combination of a high-need, vulnerable target population with two strict-regulatory environments—migration and data sharing—dictates the need for specialized capacity building in regards to migration data collection and management, not only for civil society organizations but for local and state governments as well. There are some good practices in countries such as Canada, where the CIC Settlement Program funds the delivery of settlement programming across the country and outside Canada and provides support for initiatives that contribute to the capacity enhancement of recipient organizations (IRCC, 2021). The countries explored in the case studies above could draw on this model to create their own.

6. Conclusions

The case studies in this article showcase some key issues in migration data collection and management faced in the region. Increased migrant flows and lack of resources to deal with these increases have resulted in a shift from responsibility from national governments to local governments who have different migration data needs and less resources to work with. Additionally, depending exclusively on national governments to publish data on migration and migrants in their cities, with little information about the

specific characteristics or needs of migrants in their localities, is placing municipalities in an impossible migrant integration situation.

Given the state of migration data collection and management in the region, it is important to consider alternative methods that allow for more actors to participate in the migration data collection process. Through this work, and particularly through the case studies, I observe that civil society organizations and individual relationships are currently providing the majority of transnational social protections (Levitt et al., 2017) to migrants, without an institutionalized relationship with the State or other service providers.

This article contributes to visualizing the state of migration data collection and management processes in three Latin American cities or regions. I conclude that the production and management of migration data in the region are concentrated on administrative data, which makes it difficult to create public integration or assistance policies and shows a short-sighted approach to migration data collection in the region. In the face of regional migration transformations in recent years, we need to better map how different sectors in Latin American countries are managing the increase in migrant populations in their territories. This article contributes as a first step to lay the groundwork to later develop more robust settlement services for migrants in the region.

References

Amnesty International. (2021). Venezuela.

Audebert C (2017) The recent geodynamics of Haitian migration in the Americas: Refugees or economic migrants? *Revista Brasileira de Estudos de População 34*(1), 55–71.

Baffour B, **King T and Valente P** (2013) The modern census: Evolution, examples and evaluation. *International Statistical Review 81*(3), 407–425.

Banting K (2014) Transatlantic convergence? The archaeology of immigrant integration in Canada and Europe. *International Journal 69*(1), 66–84.

Benton M, **McCarthy H and Collett E** (2015) *Into the Mainstream: Rethinking Public Services for Diverse and Mobile Populations*. Brussels: Migration Policy Institute Europe.

Bilgili Ö, **Huddleston T and Joki A-L** (2015) *The Dynamics Between Integration Policies and Outcomes: A Synthesis of the Literature. The Migration Policy Group.*

Carballo O, **Ángel L**, **Voorend K**, **Rivera FR and Bermúdez KV** (2013) Nicaragüenses en el Norte: Condiciones laborales y prácticas de contratación de hombres y mujeres inmigrantes en la Región Huetar Norte.

Department of Homeland Security (2019) Migrant protection protocols.

FAO (2022) Herramienta para formular, monitorear y evaluar proyectos de administración de tierras en América Latina.

Gabarrot M and Cervantes M (2021) Recolección de datos de las personas migrantes para la integración social. ILDA Next Generation Research Presentation. Available at https://www.youtube.com/watch?v=JUw7I_nQmwc&t=386s. Date accessed: November 4, 2022.

Global Data Barometer (2022) Global Data Barometer Report. First Edition Report – Global Data Barometer. IL-DA. DOI: https://doi.org/10.5281/zenodo.6488349

González Arias A (2017) El estudio de la migración de tránsito: definición y proceso de análisis. In González Arias A, Aikin Araluce O, Acosta García RG, Hernández López RA, Martínez Ortiz E, Ruiz Marrujo OT, Vega Villaseñor H, and Woo Morales O (eds), *Procesos Migratorios en El Occidente de México*. Mexico: ITESO.

Haddad P, **Eduardo M**, **Sánchez MRT and Cardona CAP** (2018) Impacto de la crisis migratoria venezolana en los colegios públicos del municipio de Cúcuta. *Revista Academia & Derecho 4*(7), 1–16.

Horta IAM and Rossiasco P (2019) *How Venezuelan Migration is Changing Urban Expansion in Cucuta*. Colombia: World Bank Blogs.

INEGI (2020) Presentación de resultados, Nuevo León.

Inter-agency Coordination Platform for Refugees and Migrants from Venezuela (2022) R4V. Available at https://www.r4v.info/en. Date accessed: November 4, 2022

IRCC (2021) Settlement program terms and conditions.

Levitt, P., **Viterna, J.**, **Mueller, A.**, & **Lloyd, C** (2017). Transnational social protection: Setting the agenda. *Oxford Development Studies*, *45*(1), 2–19.

Maldonado O (2022) Planteará Colosio registro de migrantes en Monterrey. *Milenio*, 25 January 2022.

Mason DP and Fiocco E (2017) Crisis on the border: Specialized capacity building in nonprofit immigration organizations. *Voluntas: International Journal of Voluntary and Nonprofit Organizations 28*(3), 916–934.

Migración Colombia (2019) *Así se construyen nuestros datos estadísticos: Flujos migratorios.*

Migration Data Portal (2021) Migration Data Sources. *International Organization for Migration*. Available at:https://www.migrationdataportal.org/ar/themes/migrationsdatenquellen. Date accessed: November 4, 2022

Ministry of Business, Innovation and Employment (2022) *Migration data explorer.*

Molnar P (2020) *Technological Testing Grounds: Migration Management Experiments and Reflections from the Ground Up*. EDRI, Refugee Law Lab. Toronto, Canada.

Morales Y (2020) México se convierte en el mayor receptor de migrantes. *El Economista.*

Morales A, **Wing-Ching KL and Acuña G** (2010) *Migración y Salud en Zonas Fronterizas: Nicaragua y Costa Rica*. Chile: CEPAL.

Mudarra F and Ortiz O (2019) *Cúcuta, el epicentro de la emergencia migratoria de Venezuela. El País.*

Open Knowledge Foundation (OKFN) (2020) *Open Data Handbook.*

Organización Internacional para las Migraciones (2017) *Tendencias Migratorias en América del Sur.*

Oxfam (2019) *Yes, but not here: Perceptions of xenophobia and discrimination towards Venezuelan migrants in Colombia, Ecuador, and Peru.*

Portes A and Zhou M (1993) The new second generation: Segmented assimilation and its variants. *The Annals of the American Academy of Political and Social Science 530*(1), 74–96.

Putnam RD (2007) E pluribus unum: Diversity and community in the twenty-first century, the 2006 Johan Skytte prize lecture. *Scandinavian Political Studies 30*(2), 137–174.

Rodriguez DX, **McDaniel PN and Ahebee M-D** (2018) Welcoming America: A case study of municipal immigrant integration, receptivity, and community practice. *Journal of Community Practice 26*(3), 348–357.

Sanchez Najera F and Freier LF (2022) The Cartagena refugee definition and nationality-based discrimination in Mexican refugee status determination. *International Migration 60*, 37–56.

Sedas AC, **Aguerrebere M**, **Martínez LA**, **Zavala-de Alba LE**, **Eguiluz I and Bhabha J** (2020) Reporte situacional: migración de tránsito en México durante la pandemia de COVID-19.*Lancet Migration*. May 2020, 1-11.

Seele A (2019) *Latin America's migration lesson for the world. America's Quarterly.*

Selee A, **Bolter J**, **Muñoz-Pogossian B and Hazán M** (2019) *Creativity Amid Crisis: Legal Pathways for Venezuelan Migrants in Latin America*. Washington, DC: Migration Policy Institute and Organization of American States.

South American Migration Report (2017) Migration Data in South America. Available at: https://www.migrationdataportal.org/es/regional-data-overview/datos-migratorios-en-america-del-sur. Date accessed: November 4, 2022

Tjaden J (2021) Measuring migration 2.0: A review of digital data sources. *Comparative Migration Studies 9*(1), 1–20.

Torre-Cantalapiedra E and Yee-Quintero JC (2018) México¿ una frontera vertical? Políticas de control del tránsito migratorio irregular y sus resultados, 2007–2016. *LiminaR 16*(2), 87–104.

Vertovec S (2007) Super-diversity and its implications. *Ethnic and Racial Studies 30*(6), 1024–1054.

Yates C (2019) As more migrants from Africa and Asia arrive in Latin America, governments seek orderly and controlled pathways. *Migration Information Source.*

Ybáñez Zepeda E and Lara CB (2017) Trayectorias recientes de la migración interna en la Zona Metropolitana de Monterrey: características, orígenes y destinos a nivel municipal, 2010. *Estudios Demográficos y Urbanos 32*(2), 245–281.