Micro and macro data integration for international immigrations statistics

National Statistical Institute (INE)
Department of Foreign Affairs and Migration (DEM)

July 2024.
Graph 1. Percentage of the foreign or foreign-born population (1854-2017)

Source: Population and housing censuses 1854-2017

Notes: From 1854 to 1952 the indicator corresponds to the total number of people of foreign nationality over the total number of people registered. From 1960 to 1982 it corresponds to the total number of people born abroad out of the total number of people registered. From 1992 to 2017 corresponds to the total number of people born abroad and habitual residents (excluding passers-by) out of the total of the habitual resident population.
1. Significant increase of international immigration after census
2. Different official figures for the stock of immigrants in Chile in 2018
3. Need for information to generate public policies
There are at least four institutions involved in the management of the migratory flow, which has led to databases with varying concepts, objectives, and quality.

SERMIG
- National Migration Service (Ministry of Interior)
- Visa registry in Chile
- Expiration of tourism extension permit

MINREL
- Ministry of Foreign Affairs
- Consular visa registry

PDI
- Investigations police
- Border control registry
- Police reports
- Place of residence**

SRCeI
- Service of Civil Registration and Identification
- Registry of vital events (births, deaths, marriages)
- Identification number: RUN

** for cases of visas requested outside the country
Methodology for estimating foreigners residing in Chile (2022)

**Macro integration**

- **Census**
  - 2017-04-19
  - "Stock"

- **Administrative records**
  - 2017-04-20 to 2022-31-12
  - "Flow"

= **Estimate of foreigners resident in Chile**
  - 2022-31-12

**Micro integration**

- **Residence permits**
- **Tourism overstay Police and security records**
- **border crossing control**
- **Deaths**

- **Aging**
- **Census undercoverage**
Methodology for estimating foreigners residing in Chile (2018-2021)

- **Census 2017-04-19**: "Stock" Estimate of foreigners resident in Chile
- **Administrative records 2017-04-20 to 2022-31-12**: "Flow" Estimate of foreigners resident in Chile

2022 Estimates results

- **782,495** (regular) + **107,223** (irregular) = **1,625,074** (100%)

**735,356** (regular) + **107,223** (irregular) = **842,579** (100%)
The Unique National Role, also known by the acronym RUN, is the unique and unrepeatable identification number that every Chilean has, whether resident or not in Chile, and every foreigner who remains, temporarily or permanently, with a visa other than the tourist visa in Chile.

For foreigners without RUN, these are key variables for identification and deduplication:
This section consolidates the data on consular visas granted between 2017 and 2022 from Minrel, and the base of temporary residences granted outside Chile from Sermig administrative records.

<table>
<thead>
<tr>
<th>Deterministic combinations (14 total)</th>
<th>Filtered rows</th>
<th>Total weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Full name, passport and date of birth</td>
<td>7,711</td>
<td>87.6%</td>
</tr>
<tr>
<td>2 Full name, country of origin and date of birth</td>
<td>764</td>
<td>8.7%</td>
</tr>
<tr>
<td>3 Surname, passport and date of birth</td>
<td>123</td>
<td>1.4%</td>
</tr>
<tr>
<td>4 Names, passport and date of birth</td>
<td>59</td>
<td>0.7%</td>
</tr>
<tr>
<td>5 Full name, passport, day of birth and month of birth</td>
<td>37</td>
<td>0.4%</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Total rows filtered:</td>
<td>8,801</td>
<td></td>
</tr>
</tbody>
</table>
This part of the estimate of foreigners in Chile is carried out with the data that comes exclusively from Sermig, corresponding to the administrative records that maintain their first act after the 2017 Census. In total, this database contains 10,062,876 administrative records, which must be transformed into persons, after subtracting the 8,939,550 duplicate files, a base of **1,123,326 foreign people**.

<table>
<thead>
<tr>
<th>Filter criteria</th>
<th>Filtered rows</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Foreign code (ID-Sermig)</td>
<td>7,499,336</td>
</tr>
<tr>
<td>2 Full name, date of birth and nationality</td>
<td>12,756</td>
</tr>
<tr>
<td>3 filter the date of the first act until December 31. And after the 2017 Census</td>
<td>1,427,458</td>
</tr>
<tr>
<td>Total filtered rows:</td>
<td>8,939,550</td>
</tr>
</tbody>
</table>
With the purpose of consolidating the databases of foreign people who may be within Chilean territory, the Minrel base and the Sermig base are united in a large database of people. Finally, this base of 1,193,221 people is the final record that is sent to the PDI to know their last border movement.
The main objective of this stage is to attach to each of the migrants the last border movement registered in the PDI information system.

-Stage that corresponds to an internal process that carries PDI, and is carried out with the SQL programming language.

### Percentages of linking Sermig and Minrel administrative data with PDI border crossing data, by estimation year

<table>
<thead>
<tr>
<th>Year</th>
<th>Sermig</th>
<th>Minrel</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2022</td>
<td>93.3%</td>
<td>80.6%</td>
<td>92.4%</td>
</tr>
<tr>
<td>2021</td>
<td>93.7%</td>
<td>72.2%</td>
<td>91.1%</td>
</tr>
<tr>
<td>2020</td>
<td>94.0%</td>
<td>75.6%</td>
<td>91.7%</td>
</tr>
<tr>
<td>2019</td>
<td>93.7%</td>
<td>47.7%</td>
<td>88.5%</td>
</tr>
<tr>
<td>2018</td>
<td>92.6%</td>
<td>78.9%</td>
<td>91.8%</td>
</tr>
</tbody>
</table>
Final thoughts and challenges

- **Unpacking the methodology:** Need to discuss the procedures that are currently in use in the construction of official Statistics, considering the quality of the administrative records and the possibilities of use of their key variables (unique identifiers, names, surnames, dates of birth, etc.). The estimation of foreigners residing in Chile is an example and opportunity to do so.

- **Database integration:** the integration of administrative records, to deduplicate people, must be done within a methodological framework that establishes its scope based on the result that is intended to be achieved (deterministic methodology), as well as the limitations (programming languages, infrastructure processing computing).

- **Challenges:** the methodological framework of the estimation serves as an example to show the integration of administrative records, and its potential to integrate other administrative records that allow estimating in the most adequate way possible the current migratory reality in Chile (e.g.: migratory irregularity), being sensitive to the inclusion of different population groups (migrant children and adolescents).