Instituto Nacional de Estadísticas • Chile

National Statistics Institute

Methodological Document National Employment Survey (ENE)

2020

NATIONAL STATISTICS INSTITUTE
Technical Subdirectorate
Department of Labour Statistics
Department of Methodology and Statistical Innovation

METHODOLOGICAL DOCUMENT NATIONAL EMPLOYMENT SURVEY (ENE)

February / 2020
Version 3.0 (updated 3 March 2020)

Address: Morandé N° 801, piso 22, Santiago, Chile
Postal code: 8340148
Telephone: 562 3246 1010
Website: www.ine.cl
E-mail: ine@ine.cl
Facebook: @ChileINE
Twitter: @INE_Chile

Table of Contents

1 Introduction....7
2. Methodological context....10
2.1 Background of the National Employment Survey (ENE)....10
2.2 Objectives of the ENE....11
2.2.1 General objective....11
2.2.2 Specific objectives....11
3 Reference framework....12
3.1 Conceptual framework....12
3.1.1 Classification of the working-age population....14
3.1.2 Population of the labour force....15
3.1.3 Population outside the labour force....21
3.1.4 Characterization of the working-age population....23
3.1.5 Principal rates and indicators of the labour market....30
4 Characteristics of the sampling design....34
4.1 Target population....34
4.2 Geographical coverage and domains of the survey....34
4.3 Reference period and periodicity....34
4.4 Master sampling frame....35
4.4.1 Geographical coverage of the sampling frame....36
4.4.2 Geographical stratification of the sampling frame....36
4.4.3 Structure of primary sampling units....37
4.4.4 Socioeconomic stratification of the sampling frame....41
4.4.5 Special areas....42
4.4.6 Definitive structure of the master sampling frame....43
4.5 Sampling frame of the National Employment Survey (ENE)....45
4.5.1 Geographical stratification....45
4.5.2 Structure of the sampling frame....47
4.5.3 Stratification of the ENE sampling frame....50
4.6. Statistical units....51
4.6.1 Sampling unit....51
4.6.2 Unit of information and analysis....51
5 Sample size....52
5.1 Sampling strategy....52
5.1.1 Considerations about the sample size....52
5.2 Calculations of the sample size....53

5.3 Distribution of the sample size 55
5.4 Calculation of the oversample 56
5.5 Sample size of dwellings by primary sampling unit (PSU) 58
5.6 Selection of sampling units 66
5.6.1 Selection of units of the first stage 66
5.6.2 Selection of units of the second stage 67
5.7 Rotation of sampling units 69
5.8 Incorporation of the sample from the 2017 sampling frame of dwellings (SFD) 71
5.8.1 Incorporation of the new ENE sampling by increments 73
5.8.2 Selection of dwellings to be removed each month 73
6 Development of the expansion factor 75
6.1 Expansion factor for the selection of clusters 76
6.2 Expansion factor for the selection of dwellings 76
6.2.1 Eligibility adjustment 77
6.2.2 Adjustment for non-response 79
6.2.3 Adjustment of population groups (calibration) 80
7 Estimators 84
7.1 Estimators of the total for variable Y 84
7.2 Ratio estimators 85
8 Precision measurements 87
8.1 Estimator variance 87
8.2 Absolute and relative errors 88
8.2.1 Absolute error and confidence interval 88
8.2.2 Relative error and coefficient of variation 89
8.3 Representativeness of estimates 90
8.3.1 Aspects of sample size 90
8.3.2 Aspects of estimate precision 90
9 Scope of reporting of results 92
9.1 Dissemination of results 92
9.2 Data accessibility and legal framework 93
10 Annexes 94
Annex 1 Number of PSUs, dwellings, and working-age population (WAP) by geographical stratum of the ENE 94
Annex 2 Sample stratification of the sampling frame of the ENE 98
Annex 3 Visitation codes of the ENE 103
Annex 5 Population estimates for the National Employment Survey (ENE) 107
Annex 6 List of auxiliary variables from provinces 110

Annex 7 Calculation of expansion factors in the transition period..... 110
Annex 8 Development of tools for the calibration of the foreign population in expansion factors of the ENE..... 116
11 Bibliography..... 127

List of figures

Figure 1 ENE classification of the population within and outside the labour force 15
Figure 2 Criteria for classification as employed.....17
Figure 3 Exceptions to the criteria for classifying employment.....19
Figure 4 Criteria for classification as unemployed.....20
Figure 5 Classification of the population according to relationship with the labour market (extended version).....23
Figure 6 Geographical stratification of the 2017 sampling frame of dwellings (SFD).....37
Figure 7 Stratification of urban entities.....38
Figure 8 Enumeration by the meandering method 39
Figure 9 Stratification of rural entities.....40
Figure 10 Rotation scheme of the sample of dwellings 70
Figure 11 Incorporation of the 2017 sampling frame of dwellings (SFD) 72
Figure 12 Diagram of frame and sample configuration..... 111
Figure 13 Updated diagram of frame and sample configuration.....112
Figure 14 Results of the nonlinear logistic model of four parameters 124

List of tables

Table 1 Examples of employment work (measured by the ENE) and other forms of work (productive activities not considered).....16
Table 2 PSUs with size outside the range, by region 43
Table 3 Socioeconomic stratification of the 2017 sampling frame of dwellings (SFD), by region 44
Table 4 Special PSUs, by region.....45
Table 5 Number of geographical strata by region and area.....46
Table 6 Strata that do not meet the minimum number of PSUs.....47
Table 7 Special PSUs and PSUs excluded from the sampling frame of the ENE, by region....49
Table 8 Number of sampling strata by region and area.....50
Table 9 Characteristics of parameters of interest 53
Table 10 Design effects on selected combinations of the clustered sample size and intraclass correlation.....60
Table 11 Design effects by domains of the survey, for varying average numbers of dwellings per PSU.....61
Table 12 Sampling errors of the unemployment rate for urban domains, according to average number of dwellings per PSU 62
Table 13 Sampling errors of the unemployment rate for rural domains, according to average number of dwellings per PSU 63
Table 14 Distribution of the sample and results of the simulation 64
Table 15. Distribution of the sample and results of the simulation 65

Table 16 Estimate of the total foreign population (as of 30 June each year) from migration balance of the population projections of base year 2017 118

List of graphs

Graph 1 Gradual increase of theoretical sample during ENE transition period 73
Graph 2 Annual estimate of the foreign population based on the migration balance of the population projections of base year 2017..... 119
Graph 3 Annual estimate of foreign population based on the migration balance of population projections of base year 2017 and on the estimates of selected models..... 120
Graph 4 Monthly estimate of the foreign population according to the logistic model and annual estimate based on the migration balance..... 121
Graph 5 Estimate of foreign population according to the 2017 Census logistic model and the estimate of resident foreign population as of 31 December 2018..... 122
Graph 6 Assumptions of the nonlinear regression model: correct mean function 124
Graph 7 Assumptions of the nonlinear regression model: homoscedasticity 125
Graph 8 Assumptions of the nonlinear regression model: normality..... 125
Graph 9 Assumptions of the nonlinear regression model: independence..... 126

1 Introduction

The purpose of the National Employment Survey (ENE) is to provide fundamental statistical data on the labour market in order to guide decision-making and analysis of economic policies and employment work at a national level. The National Statistics Institute (INE), which produces the ENE, has undertaken this survey of the labour market since 1966. Continuous data has been published since the December–February 1986 quarter.

In the last decade, the National Employment Survey has been subjected to a series of modifications that more thoroughly capture the changes in the labour market and the fluctuations in the global market. These modifications have led to the rise of new forms of employment and work and thus to a new understanding of the phenomenon of work. The objective of these modifications to the survey is to harmonize the measurement of labour statistics with standards such as those developed by the International Labour Organization (ILO) and Organisation of Economic Cooperation and Development (OECD). These modifications have facilitated comparisons with other countries.

The most profound change in the survey occurred in 2010, when important conceptual and methodological changes led to the division of the survey into its older version (used between the December–February 1986 quarter and the same quarter in 2010) and the newer version, which has been used since the January–March 2010 quarter. Because the changes in methodology created a break in the survey, data from these two versions is not comparable. The introduction of new and more rigorous concepts that redefine the target population of the survey has altered the conceptual comparability of groups of employed persons, unemployed persons, persons outside the labour force, and inactive workers who are potentially active. The guidelines used in these changes are from the 19th International Conference of Labour Statisticians of the ILO. This update enabled INE to lead the development of labour force surveys in various national statistical offices in the region by updating the ENE with the latest international recommendations, which are in line with the development and progress of the labour market.

Beginning in 2020, INE will incorporate new modifications to the current measurement. With these modifications, the innovations in the labour market of the country can be captured. These changes included updates to the sampling frame of the survey, to the statistical design, and to the calibration method for expansion factors. The last of these updates incorporates population projections based on the 2017 Census.

The update of the **sampling frame of the survey** was based on the 2016 Precensus. The first phase of the update of the sampling frame was from November 2018 to November 2019. During this period, the number of dwellings was updated from information from the 2016 Precensus. The second phase, which began in January 2020, includes the use of new sampling clusters

based on information from the 2017 Census. Among the benefits of these clusters is the homogeneity in the size of their dwellings.

The modernization of the **sampling design** is based both on the frame of dwellings, which was constructed from 2017 Census data, and on new population projections from the year 2017. The new sampling design uses a confidence level of 95% to calculate estimation errors, an improvement over the previous design, which had a confidence level of 90%. The new design will be implemented throughout the first half of 2020.

The change in the calibration **method of the expansion factors** incorporates population projections based on the 2017 Census while considering international good practices.

As part of the recommendations of the ILO for the measurement of statistics on labour relations, new questions on the use of digital platforms, the ownership of the means of production, and other topics will be tested in the questionnaire in 2020. These questions will be monitored, and a publication plan will be determined from 2021 onwards. This methodological document is the result of intensive work to strengthen the National Employment Survey with the assistance of the International Labour Organization (ILO) and the Survey and the Employment Specialist Group, which was formed in 2019 by INE. We are grateful for the contributions of this group of specialists, which includes of Carolina Casas-Cordero Valencia[1], Susana Eyheramendy Duerr[2], Javiera Barrera Martínez[3], Ricardo González[4], and Andrés Gutiérrez Rojas[5]. Their participation was fundamental in driving the process of methodological change. The periodic review and strengthening of statistical operations constitute international good practices in the field of

[1] Ph.D. in Survey Methodology, University of Maryland, USA. Carolina Casas-Cordero Valencia is a researcher associated with the Centro MIDE UC and the Centro UC de Encuestas y Estudios Longitudinales of the Pontificia Universidad Católica de Chile. She is currently the head of the Master's Program in Social Survey Design and Analysis. She teaches the course Fundamentals of Survey and Data Science at the University of Maryland.

[2] Ph.D. in Statistics, Rutgers University, USA. Susana Eyheramendy Duerr is an associate professor in the department of Statistics at the Pontificia Universidad Católica de Chile and in the Faculty of Engineering and Sciences at Universidad Adolfo Ibáñez.

[3] Ph.D. in Applied Mathematics from L'Université Paris-Descartes and in Engineering Sciences in Mathematical Modeling from la Universidad de Chile. Currently, Javiera Barrera Martínez is associate professor and subdirector of the graduate program at the Faculty of Engineering and Sciences at Universidad Adolfo Ibáñez (UAI).

[4] Business Engineer and Master of Economics from la Pontificia Universidad Católica de Chile. Ricardo González has also done advanced studies in methodology of surveys at the GESIS-Leibniz Institute for the Social Sciences in Cologne, Germany. Since September 2013, he has been the coordinator in the Area of Public Opinion at the Centro de Estudios Públicos (CEP) and the coordinator of International Social Survey Programme (ISSP) in Chile.

[5] Ph.D. in Statistics and Master of Statistical Sciences from Universidad Nacional de Colombia. Andrés Gutiérrez Rojas currently works as a regional expert in Social Statistics for the United Nations Economic Commission for Latin America and the Caribbean (ECLAC).

statistical quality[6]. Thus, the methodologies, procedures, and assumptions used in the surveys must be reviewed regularly in consideration of the changing social, demographic, and economic phenomena of different countries.

This document is divided into nine chapters that address the updating of the methodology of the National Employment Survey in the year 2020.

- The second and third chapters set out the methodological context of the survey, describing the background and objectives of the statistical product, in addition to the conceptual framework underpinning the measurement of employment.
- The fourth, fifth, and sixth chapters explain the new statistical design of the survey and the characteristics of the sampling frame.
- The seventh and eighth chapters present the estimators and precision measurements used in the new measurement.
- The ninth chapter addresses the scope of dissemination of results.

[6] Recommendation 8 of the "Recommendation of the OECD Council on Good Statistical Practice" (OECD, 2015) indicates that statistical products need regular and comprehensive reviews that include the participation of external experts.

2. Methodological context

2.1 Background of the National Employment Survey (ENE)

Labour-market statistics have long been gathered by INE. In 1966, INE created the first version of the National Employment Survey, one of the oldest in Latin America (Cecchini, 2005). Since then, many modifications have been made to the survey, including changes to sampling frames (BCCh, 2001). Thus, since the 1960s there have been three series of data.

In 2010, the incorporation of international standards of measurement in the conceptual and methodological design of the ENE led to the updating of the concepts of the populations of employed, unemployed, and those outside the labour force (or inactive). Furthermore, “traditional” and “non-traditional” forms of employment were distinguished, as were the various reasons for inactivity.

Maintaining this structure, the conceptual updates were accompanied by a description of the main indicators calculated from the categories of labour activity of the population, including a summary of the main analytical indicators, which consist of a set of complementary indicators that facilitate the analysis and monitoring of the labour market.

More specifically, the modifications made to the survey in 2010 included batteries of questions for updating information on the working-age population (WAP) and an expansion in the coverage of indicators for the analysis of employment and unemployment. In addition, INE developed complementary indicators for the characterization of employment and of persons outside the labour force. The new topics relate to the quality of employment of dependent workers and to measurement of the underutilization of labour, such as involuntary part-time work and inactive workers who are potentially active.

In 2017, questions were incorporated to measure labor informality through the “informal employment rate”, the “employment rate in the informal sector” and the number of “informal employed persons”, among others.

In 2020, measures have been incorporated into the ENE to improve the capture of the new forms of work (for example, through digital applications) and of the labour flexibility currently experienced by dependent and independent workers. These phenomena are explained in detail in Resolution I of the 20th ICLS of 2018. These latter adjustments are intended to advance the adoption of the main international recommendations in this area.

2.2 Objectives of the ENE

2.2.1 General objective

The objective of the National Employment Survey is to classify and characterize the working-age population (aged 15 and over) habitually resident in Chile, according to their relationship with the labour market. This complies with international standards for measuring labour statistics, which provide indicators that enable international analysis and comparison of the characteristics of employment in the country in a timely and structured manner.

2.2.2 Specific objectives

1. To quantify the population aged 15 years and over by its classification into three mutually exclusive groups: employed and unemployed (both constituents of the labour force) and those outside the labour force, in accordance with current international standards.

2. To determine characteristics of employment conditions and modalities of labour insertion of workers, whether they are dependent or independent, formal or informal, and full-time or part-time while measuring the underutilization of the labour force, among other aspects. Thus, indicators complementary to the traditional ones can be constructed in order to monitor and analyze the labour market.

3. To capture the development of traditional indicators (such as the unemployment rate, participation rate, and employment rate) at the national level and at the other levels of estimation defined in the sampling design.

4. To characterize the sociodemographic nature of the Chile's resident population (including nationals and foreigners) with special attention to the population aged 15 and over, which is defined as the working-age population.

3 Reference framework

This chapter introduces the concepts that form the basis from which the ENE classifies and characterizes the working-age population according to its relationship with the labor market. Through the questions that make up the questionnaire, these concepts are measured in a specific reference period, which is the definite and limited period in which respondents are consulted about their activities and experiences. Specifically, the ENE uses the reference week.

3.1 Conceptual framework

As conceptual foundation, the ENE includes the most updated criteria and good practices of classification from the resolutions of the International Conferences of Labour Statisticians (ICLS) of the International Labour Organization (ILO)[7]. The ICLS makes recommendations on selected topics of labour statistics in the form of resolutions and guidelines, which form part of international standards for concepts, definitions, classifications, and other methodological procedures that are agreed to be representative.

In particular, the conceptual framework of the ENE includes the resolution on statistics of the economically active population, employment, unemployment, and underemployment, adopted by the 13th International Conference of Labour Statisticians (1982); the resolution on statistics of labour, employment, and labour underutilization, of the 19th ICLS (2013)[8]; and the resolution on statistics of labour relations of 2018. These criteria are the same as those adopted by OECD and EUROSTAT for all their member countries.

The 19th ICLS (2013) adopted five resolutions on labour statistics in employment work and underutilization of the labour force, forced labour, cooperatives, and labour migration, in addition to resolutions on the functioning of the conference. Specifically, the resolution on labour statistics, employment, and underutilization of the labour force identifies five mutually exclusive forms of work, which are distinguished according to the intended purpose of production (for own final use or for consumption by third parties) and the nature of the transaction (monetary or non-monetary transactions and transfers).

The five forms of work are the following:

[7] The ICLS is held every five years, and it includes the participation of government experts (mostly from ministries responsible for labour and from national statistical offices) in addition to employer and worker organizations.

[8] The 19th ICLS meeting was held in 2013, but preparatory meetings began in 2010. Thus, the new version of the National Employment Survey was able to incorporate the latest ILO recommendations.

1. Own-use production work
2. Employment work
3. Unpaid trainee work
4. Volunteer work
5. Other work activities

Employment work establishes the scope of reference activities for labour-force statistics (i.e., the labour supply available for the production of goods and services in exchange for remuneration or profit).

Regarding the expression “in exchange for remuneration or profit”, the ICLS establishes that it “...refers to work done as part of a transaction in exchange for remuneration payable in the form of wages or salaries for time worked or work done, or in the form of profits derived from the goods and services produced” (19° CIET, 2013). This exchange involves market transactions, With respect to the remuneration or profit of employment, the same resolution states the following:

a) It includes remuneration in cash or in kind, whether actually received or not, and may comprise additional components of cash or in-kind income.
b) The remuneration may be payable directly to the person performing the work or indirectly to a household or family member (19° CIET, 2013).

For measurement purposes, the 19th ICLS also introduced the criterion of having worked at least one hour in the reference period, enabling precarious occupations to be included. The criterion also included the coverage of all productive activities (including part-time, temporary, occasional, and sporadic activities), thus ensuring a comprehensive measurement of all production-related labour inputs. At the same time, that criterion maintains consistency with the concepts of production boundary and economic activity of the System of National Accounts according to which the aggregate value of the economy and, therefore, the gross domestic product (GDP) are calculated.

The 20th ICLS (2018) adopted four resolutions concerning statistics on labour relations, child labour, and indicators of the Sustainable Development Goals (SDG) for labour rights and youth. Labour-relations statistics address (a) the relationships of authority between workers and the economic units in which they work or for which they perform work and (b) the economic risks arising from the contractual situation or other conditions under which the work is performed.

Since its 2010 version, the ENE has adopted the concepts that, as of 2013, are part of the guidelines from the resolution on labour statistics, employment, and underutilization of the labour

force. Specifically, the survey measures employment, which includes work performed for third parties in exchange for remuneration or profit.

The incorporation of international standards in the ENE has resulted in the following:

- Using an updated concept of the employed population by incorporating the criteria of having worked at least one hour in the reference week and of receiving remuneration or profit for activity performed
- Distinguishing between “traditional” and “non-traditional” employed persons by considering whether the interviewee recognizes or identifies their activity as work
- Updating the conceptualization of the unemployed population by incorporating the criteria of having sought employment in the last four weeks and of being available to work in the two weeks following the reference period
- Classifying people outside the labor force according to reasons for inactivity and to their potential for participation in the labor market
- Establishing the relationships of authority and economic risk between workers and the economic units for which they work

Based on this conceptual framework, the ENE has met its objective. Through the use of the ENE, it is possible to classify persons of working age who are habitual residents in private occupied dwellings in the country as employed or unemployed, or as persons outside the labour force (mutually exclusive categories) and to characterize them according to a range of variables.

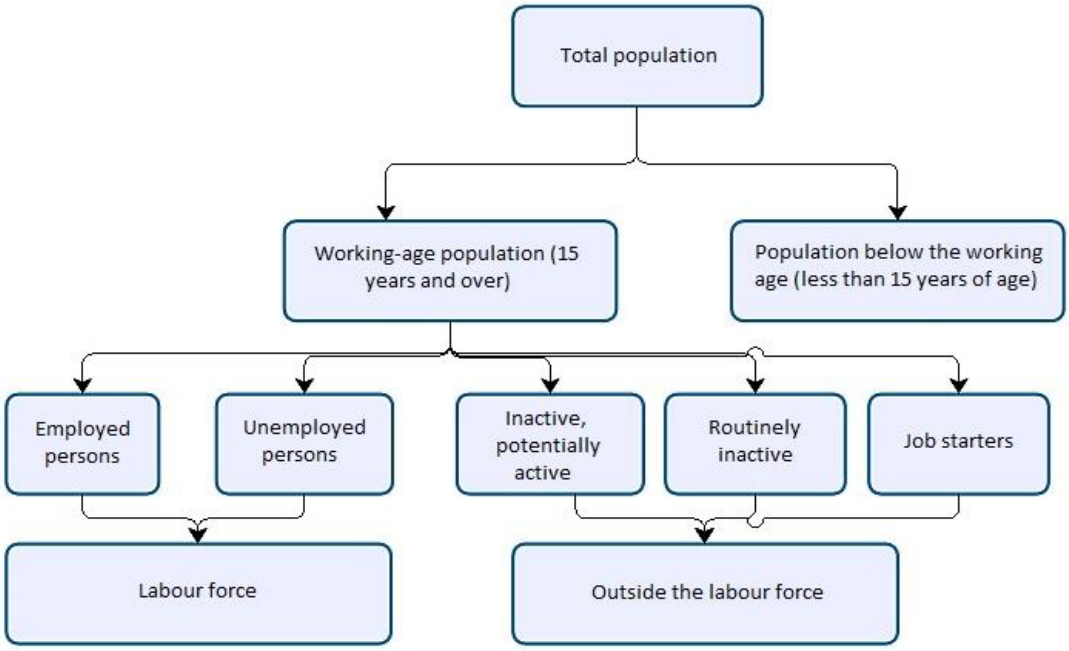
3.1.1 Classification of the working-age population

The working-age population (WAP) consists of persons who are able to have a job (i.e., engage in employment work). In Chile, this has been established as the population aged fifteen years and over.

The WAP can be classified according to the situation in the labour force by applying a set of priority rules, which are based on the international approach to this measurement through household surveys.

To implement the classification, a distinction is first made between those who are part of the labour force and those who are not, as shown in Figure 1.

Figure 1 ENE classification of the population within and outside the labour force



Source: Produced by the National Statistics Institute

3.1.2 Population of the labour force

The population within the labour force, previously known as the economically active population, includes all those persons aged 15 or over who, during the reference week, meet the requirements for inclusion in the category of employed or unemployed persons (i.e., persons who have some relationship within the labour market).

To establish the situation of the population within the labour force (employed and unemployed) or outside the labour force, two concepts must be used. The first is whether the person receives remuneration or profit for the good produced or the service rendered, and the second is the criterion of one hour.

In the first concept, activities included within employment work are the set of actions contributing to the production of goods and services (whether legal or not) that occur in a framework of market transactions, which assumes a consent between the parties. Likewise, other forms of work (such as own-use production and unpaid household and care work) are not included in this group, nor are unpaid voluntary work and unpaid trainee work. These productive activities are included in the four other forms of work apart from employment work.

Table 1 shows examples of employment work captured by the ENE and other forms of work or productive activities not included in the survey.

Table 1 Examples of employment work (measured by the ENE) and other forms of work (productive activities not considered)

Employment work	Other forms of work
• Paid domestic work	• Performing household chores, whether or not in own household, without receiving payment
• Manufacture of products in the household for sale on demand	• Manufacture of products in the household for self-consumption
• Apprentice or personnel in training related to the productive activity of a company, for which remuneration is received	• Training not related to the productive activities of a company, even if a financial subsidy is received
• Work in foundations or nonprofit organizations for remuneration (administrators, researchers, coordinators)	• Unpaid volunteer work in foundations or nonprofit organizations
• Caring for children, the elderly, the sick, or disabled people, whether or not they are family members, for remuneration	• Care of children, the elderly, the sick or the disabled, whether or not they are family members, without receiving remuneration
• Selling products in the street or in the household, whether purchased or produced in the household	--
• Preparation or promotion of products in the street for their subsequent sale	--
• Hunting and fishing activities for selling in the market	• Hunting and fishing activities for recreation and leisure purposes
• Collection of waste material (cans, cardboard, glass, etc.) for sale	--
• Working on one's own or another person's land to obtain a product for sale in the market	--
• Work in exchange, where work is exchanged between family members for productive purposes, such as farming on another farmer's land (*minga*)	--
• Own-account worker who spends time planning, improving the business, or looking for customers	• Own-account worker who is in the process of starting a business or activity
• Management of one's own company or farm, even if one does not work in production	• Investment in a business without contributing to its management or operation, such as holding shares in a company
• Work without pay in a family business, but with a share in the profits	• Work without remuneration for a member of the household who is not the owner of a company or estate, but is, in turn, the recipient of a salary

--

- Repair of working tools for future use (for example, when a fisherman repairs his net for future fishing for market exchange purposes)
- Work without remuneration for a member of the household who is not a family member
- Repairing the dwelling where one resides

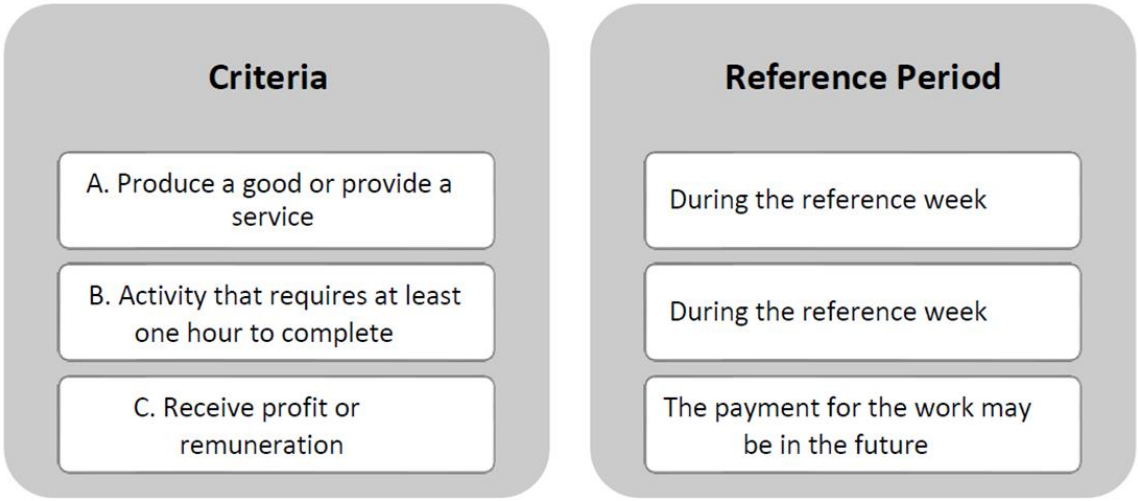
Source: Produced by the National Statistics Institute

The second concept, the one-hour criterion, which was established by the ILO, is complementary to the criterion of payment in money or in kind. It refers to the minimum amount of time spent in a market-oriented economic activity during the reference week. These two concepts make it possible to define persons as employed or not employed. The latter group can then be divided into those who are unemployed and those who are outside the labour force.

3.1.2.1 Employed persons

Employed persons are all working-age persons who, during the reference period, meet the following criteria:

Figure 2 Criteria for classification as employed



Source: Produced by the National Statistics Institute

Among the employed population, “traditional” and “non-traditional” employed persons can be distinguished according to each person's conception of the work they perform. Thus, those who indicate in the first question of the survey that they worked at a job and that they will be paid for it in money or in kind are classified, according to the specified criteria, as a “traditional employed

persons”. On the other hand, those who do not identify their activity as a job at the outset are classified as “non-traditional employed persons” by further investigation of their activity.

The greater precision in the classification of employed persons makes it possible to include persons who work at jobs that, because of their nature (in hours worked, the frequency with which they perform it, or place of work) would otherwise not have been identified, in principle, as employment work. An example is the work of those who sell products by catalogue. Sometimes these workers, in addition to this activity, carry out tasks related to care or household chores. Another example is students who work as assistants or private tutors, an activity for which they receive payment.

However, among the employed population, two exceptions to the previously established criteria must be considered: unpaid family workers and absent employed persons.

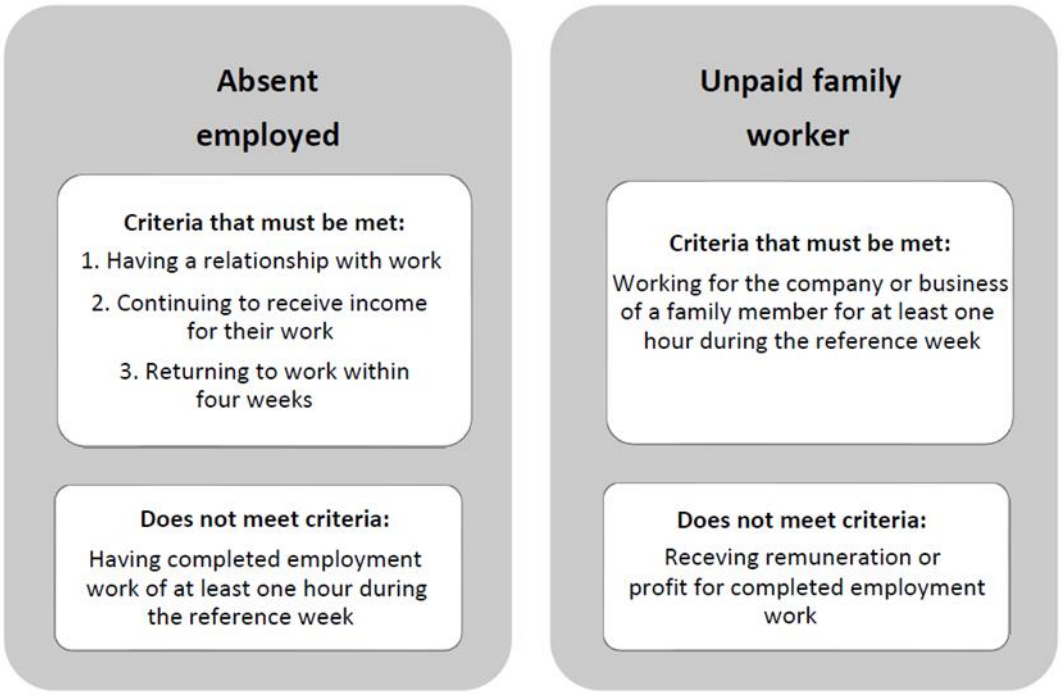
(a) Unpaid family workers are those who performed some employment work for a family member for at least one hour during the reference week without receiving remuneration or benefits, if the family member is the owner of the enterprise or business and is not a dependent worker. An example of an unpaid family worker can be found in families who work at a stall in a street market and have no established system of payment for their members.

(b) Absent employed persons are persons who, having a current occupation, were absent from it during the reference week. To be considered as such, they must maintain a relationship with the work, whether formal or informal. This is determined according to the respondent's certainty of whether they will return to work within a given time and whether they will receive remuneration. In the case of dependent work, the strong relationship criterion is used in which the person is considered to have a relationship with their job resulting from a contractual relationship or verbal agreement. That is, their absence was a result of vacation, medical leave, or training. Within this group can be distinguished the following:

- With a relationship of support payment: employed persons who are absent from their job but who maintain wages or profit
- With a relationship of prompt return: employed persons absent from their jobs who will return within four weeks

The summary of the above is shown in Figure 3.

Figure 3 Exceptions to the criteria for classifying employment



Source: Produced by the National Statistics Institute

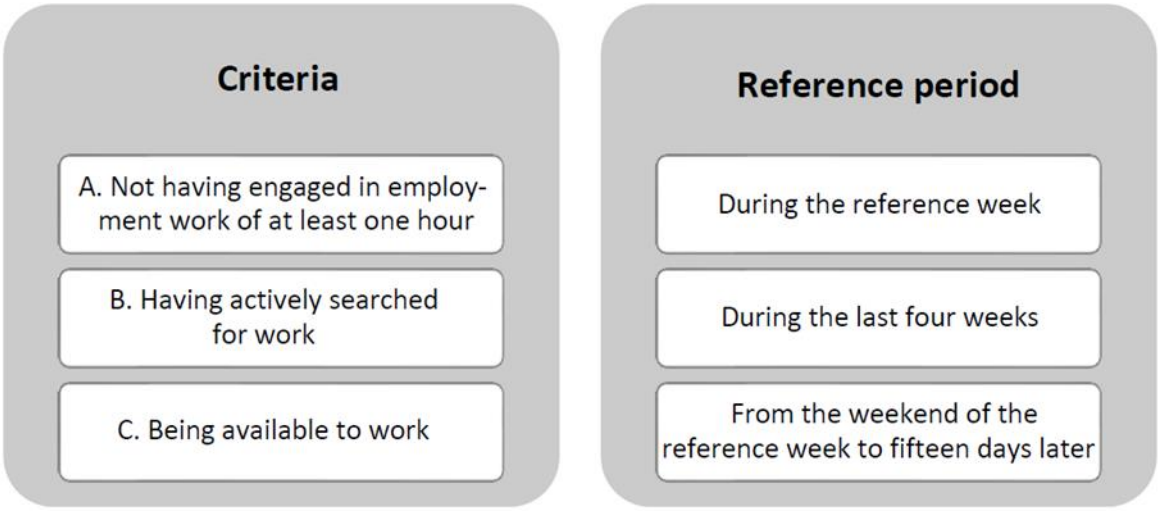
3.1.2.2 Unemployed persons

Persons aged fifteen and over are classified as unemployed during the reference period if they meet the following criteria:

a) They were without employment work. That is, they did not have wage-earning job or an independent job in the reference week, as defined in the previous point.
b) They are available to work. That is, they are available to work in a wage-earning job or in an independent job during the reference period (the two weeks following the reference week).
c) They are seeking employment work. That is, they have taken concrete steps to search for a wage-earning job or to begin an independent job in the previous four weeks.

These criteria are summarized in the following figure.

Figure 4 Criteria for classification as unemployed



Source: Produced by the National Statistics Institute

The criterion of seeking work involves having taken specific steps to search for work in a recent period (four weeks). The search covers dependent, independent, full-time, part-time, temporary, seasonal, or occasional work. In general, it covers any type of employment work, as previously defined.

The concrete steps to search for employment work refer to an active search. These steps include the following: talking directly with employers; reviewing or responding to advertisements (in newspapers, internet); placing advertisements (in newspapers, internet); sending CVs to companies or institutions; participating in tests, tenders, or interviews for recruitment; consulting employment agencies; asking acquaintances or relatives for recommendations or advice on getting a job; registering with the Municipal Employment Information Office (OMIL); making arrangements to set up business; and seeking clients or orders, among other activities.

In addition, the concrete measures to find employment work must have taken place within a specified recent period, established as the last four weeks. This ensures that real pressure on the labour market will be captured.

A person is considered available to work if the person is willing to work whenever the opportunity arises in a period that does not exceed fifteen days or two weeks following the reference week of the survey.

If a person satisfies the conditions of this definition of unemployment, they will be considered unemployed even if their situation is compatible with other categories (e.g., if they are a student or they perform housework).

Unemployed persons can be subdivided into the following groups:

a) **The redundant:** All unemployed persons who have previously held employment work for at least one month.
b) **First-time job seekers:** persons who have met the criteria for classification as unemployed but who have not held employment work of at least one month.

There are exceptions to the rule in specific cases. For example, persons will be considered unemployed if they meet the following criteria: they did not work during the reference week; they were available to work, but did not look for work, because they had already found a job with which they have a formal relationship (contract or arrangements made); and they will be incorporated into the labour force in more than four weeks from the reference week. In this case, the criterion of effective job search is not required.

In addition, persons absent from work who are not considered employed (i.e., those who receive an income from work equal to or less than 50% of what they usually receive and who are going to return to work after thirty days) are categorized as unemployed if they sought work and are available to perform it.

3.1.3 Population outside the labour force

The population outside the labour force includes all persons 15 years or older who did not have a job and did not seek one and/or were not available for employment work. In other words, they are classified as neither employed nor unemployed during the reference week. These persons are therefore considered to be outside the labour force.

Within this group, three subgroups can be distinguished according to their availability or to the potential with which they can move within or outside the labour force (either as employed or unemployed persons). The subgroups are the following:

a) **Inactive workers who are potentially active:** persons who sought employment work but who are not available to work, or persons who did not work but are available to work. These cases are evaluated regardless of the reason why they occur.

   The potential activity of persons outside the labour force refers to a presumption of the feasibility that a person who is not pressuring the labour market will actually do so at some point.

b) **Routinely inactive:** persons who did not seek employment work and who are not available to start work.

c) **Job starters:** persons who have not sought employment work in the previous four weeks, because they will soon begin work. However, these persons may be available to work in the two weeks following the reference week.

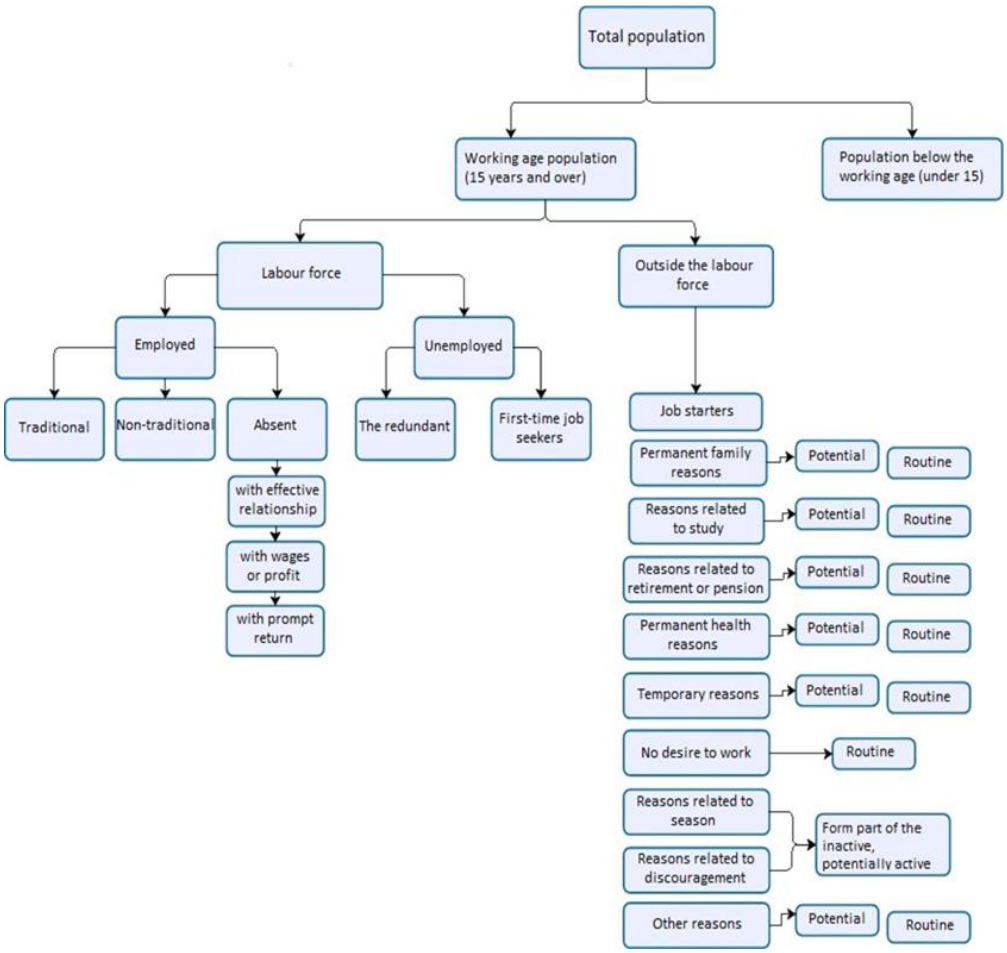
If job starters report that they are available, both the ILO and member countries of the Organization for Economic Co-operation and Development (OECD) recommend that they be included among the unemployed. Therefore, they should be included in the calculation of the unemployment rate. Otherwise, they are considered outside the labour force because they did not seek work and were not available for work. In general, availability supposes a greater short-term participation in the labour force because, in essence, being a job starter (available or not) is an intermediate stage between being outside the labour force and being part of it. According to the gradation of people outside the labour force according to their potential to be part of it, job starters are the subset with the greatest potential.

For analytical and comparative purposes at the international level, a monthly publication of the ENE shows the unemployment rate with available job starters, which is used as a complementary measure to the traditional unemployment rate and as part of the indicators of underutilization of the labour force.

Persons outside the labour force give various reasons for their inactivity, which can be categorized according to their potential or regularity. These reasons include the following: seasonal downturn, discouragement, temporary personal reasons, permanent family reasons, study, pension, retirement, permanent health reasons, no desire to work, and other reasons.

Everything presented so far regarding the classification of the working-age population is summarized in Figure 5.

Figure 5 Classification of the population according to relationship with the labour market (extended version)



Source: Produced by the National Statistics Institute

3.1.4 Characterization of the working-age population

Once the working-age population has been identified and classified according to relationship to the labour market, it is characterized according to a series of sociodemographic and economic variables.

The capture of the answers to sociodemographic questions makes it possible to characterize the population based on certain basic variables, segmenting it into homogeneous groups. These variables are collected in the sociodemographic questionnaires or household registration cards, and they are associated with persons regardless of whether they participate in the interview. For the ENE, the answers to sociodemographic questions are collected from all members of the household, regardless of whether they meet the requirement of being part of the target sample of the survey (i.e., persons aged fifteen years and over).

In the survey, the sociodemographic characterization includes questions on the identification of the household and its members, including sex, age, date of birth, marital status, educational level, kinship, main provider, nationality, migration, and indigenous identity, among others. The International Standard Classification of Education (ISCED)[9], developed by the United Nations Educational, Scientific and Cultural Organization (UNESCO) in 1997, is used to characterize the population according to its educational level. This instrument is used for the compilation and presentation of education statistics in different countries, thus enabling international comparability. This classification covers two aspects: levels and fields of education.

3.1.4.1 Characterization of the employed population

A first general characterization of the employed population is made according to the employment situation, whose conceptual framework comes from the ILO's International Classification of the Status in Employment (ICSE)[10] of 1993.

The ICSE-93 classifies jobs according to the type of explicit or implicit contract of employment of the person with other persons or organizations. The basic criteria used to define the groups in the classification are the type of economic risk and the type of authority:

a) The type of economic risk is determined by the strength of the attachment between the person and the job.
b) The type of authority is determined by the authority the person has over establishments and other workers.

For the first criterion, the employment situation may be dependent or independent. In a dependent situation, the worker has a contractual relationship or work agreement that establishes their subordination to an employer, and they may (or may not) have workers under their responsibility. In an independent situation, the worker does not have a contractual link or work agreement, but they may (or may not) have workers under their responsibility. An independent worker assumes the economic risk inherent in the exercise of his activity.

This initial characterization can be supplemented according to the specific needs of countries, who can divide and adapt the general characterization to the national context. Thus, the national adaptation of the ICSE-93, classifies employed persons as follows:

[9] For further details, see the ISCED 2011 International Standard Classification of Education (UNESCO, 2006).

[10] For further details, see the International Classification of Status in Employment (ICSE), (OIT, 1993).

a) Employers: independent workers who, as part of their work, have workers under their responsibility
b) Own-account workers: independent workers who engage in an economic activity without an employer and who work alone or with a partner (i.e., without workers under their responsibility)
c) Wage earner of the private sector: dependent worker with a contractual relationship with a private economic entity
d) Wage earner of the public sector: dependent worker with a contractual relationship with a public economic unit, that is, a worker in State institutions or companies, whether in public administration or in other economic sectors (health, education, etc.)
e) Domestic service workers: dependent workers who engage in domestic chores, whether they live in or outside their place of work
f) Unpaid family worker: a worker who, as previously stated, does not fulfil the condition of receiving a payment in money or in kind for their work, but, because their work is carried out for a family member who owns the economic unit, they participate indirectly in the economic benefits produced by this unit

One of the features of this national adaptation is that it allows for the disaggregation of domestic service workers (those who perform domestic work for remuneration or profit). These workers were given a separate category because they were included within the group of salaried workers in the original ICSE. In addition, salaried workers in the public and private sectors have been distinguished, whereas in the original ICSE they form a single category. This adaptation does not include the category “members of producers' cooperatives”.

In addition to the employment situation, the employed population is characterized according to the International Standard Classification of Occupations (ISCO)[11], an instrument for organizing occupations into a series of groups defined according to the tasks performed.

ISCO-08 defines a job as “a set of tasks and duties performed, or meant to be performed, by one person, including for an employer or in self-employment”. Occupation is defined as “a set of jobs whose main tasks and duties are characterized by a high degree of similarity” (INE, 2018a).

Jobs are classified by occupation according to the type of tasks that the worker performs, or is meant to perform. The basic criteria used to define the system of major, sub-major, minor, and unit groups are the “skill level” and “skill specialization” required to perform the tasks and duties of the occupations competently.

[11] For more information, see *Clasificador Chileno de Ocupaciones* (INE, 2018a).

At the beginning, the ENE used the ISCO-88 version, which was divided into ten major groups. In the year 2018, this classification was adapted to the situation in Chile, which resulted in the CIUO-08.CL, a national adaptation that maintained the basic principles and structure of the classification, including the following major groups:

1. Managers
2. Professionals
3. Technicians and associate professionals
4. Clerical support workers
5. Service and sales workers
6. Agricultural and fishery workers
7. Craft and related trades workers
8. Plant and machine operators and assemblers
9. Elementary occupations
10. Unidentified

With the conceptual framework of the International Standard Industrial Classification (ISIC), INE can characterize the economic sector where the job or occupation takes place, according to the branch of economic activity. Within this framework, it is necessary to understand the concept of economic unit, which refers to entities (institutions, companies, businesses, or persons) engaged in the production of goods, purchase or sale of merchandise, or the provision of public and private services. Thus, the ISIC makes it possible to establish the sector of the economic unit. Currently (and officially since 2016), the ENE uses the National Classification of Economic Activities from Sociodemographic Surveys (CAENES)[12], an adaptation of the CIIU4.cl 2012, which was developed by INE following international recommendations. This adaptation arose from the need to collect updated information on the productive sectors in household surveys and to fulfill the commitment that INE made with the Organisation for Economic Co-operation and Development (OECD).

Between 2010 and 2016, the survey used CIIU rev. 3 (the Spanish version of ISIC rev. 3), which has seventeen sections for the classification of the branches of economic activity.

The CAENES classifies employed persons into twenty-one sections as follows:

[12] For more information, see *"Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas"* (INE, 2016a) and the technical supplement regarding its adaptation (INE, 2016b).

A) Agriculture, forestry, and fishing
B) Mining
C) Manufacturing
D) Electricity, gas, steam, and air conditioning supply
E) Water supply; sewerage, waste management, and remediation activities
F) Construction
G) Wholesale and retail trade; repair of motor vehicles and motorcycles
H) Transportation and storage
I) Accommodation and food service activities
J) Information and communication
K) Financial and insurance activities
L) Real estate activities
M) Professional, scientific, and technical activities
N) Administrative and support activities
O) Public administration and defense; compulsory social security
P) Education
Q) Human health and social work activities
R) Arts, entertainment, and recreation
S) Other service activities
T) Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
U) Activities of extraterritorial organizations and bodies

For classification, it is important to consider for whom employed persons work and with whom the contractual or work agreement has been established. If the economic unit for which employed persons work is not the same as the economic unit with which they have a work agreement, workers are outsourced.

Outsourcing is characterized by the limited communication channels between workers and those who exercise positions of authority over them. In addition, there may be a situation of “double authority”, where workers must answer for their actions both to the authority of the unit where they work and to the unit with which they have a work agreement.

The ENE identifies different situations of outsourcing in the national labour market:

- with a contractor or subcontractor of goods or services
- with a temporary services and staffing company
- with an agricultural contractor

The International Labour Organization (ILO) points out that, for a complete monitoring of the labour market, it is not enough to look only at traditional indicators such as the rate of unemployment, employment, and participation. The analysis needs to be complemented by indicators that can draw attention to other population groups that are in an unfavorable situation in terms of their relationship with and participation in the labour market. Therefore, the investigation of labor informality[13] and other issues was proposed.

At the 15th International Conference of Labour Statisticians in 1993, the ILO made a standardized proposal on the concept of the informal sector and the scope of labour informality from the perspective of the economic unit. In 1997, the United Nations Department of Statistics established the Delhi Group to strengthen the statistical standards developed by ILO. Thus, the concept of informal employment was adopted at the 17th ICLS in 2003, which introduced the labour perspective based on the characteristics and conditions of the job a person holds. Both concepts established the definition of the informal economy as a general and inclusive framework for the measurement of statistics of labour informality.

In 2013, the ILO developed the Statistical Manual on the Informal Sector and Informal Employment, which is the main conceptual framework for measuring labor informality of the ENE.

The following definition of the informal sector and occupation has been adopted in Chile[14]:

a) **Informal Sector:** all economic units in the market that are not registered with the Internal Revenue Service (SII) and that do not have complete or simplified accounting that allows them to effectively separate their expenses (for own-account workers or employers), nor do they have an accounting office nor the services of an accountant (for wage earners). Therefore, they cannot be classified as quasi-corporations.
b) **Informal employment:** all employees or domestic workers who, because of their relationship with an employer, do not pay contributions to health insurance (Isapre or Fonasa) or to social security (AFP). Also, included in this definition are all unpaid family workers of the household, own-account workers who work in the informal sector, and employers who own an economic unit in the informal sector.

The development of official indicators on informal employment and employment in the informal sector was made possible through the implementation of a set of new questions in the ENE

[13] For more information on the conceptual framework of the measurement of informality, see the document *Estadísticas de informalidad laboral: Marco conceptual y manual metodológico* (INE, 2018b).

[14] For more information, see https://www.ine.cl/estadisticas/sociales/mercado-laboral/informalidad-y-condiciones-laborales.

questionnaire. These questions focus on a range of aspects in order to identify and characterize the situation of formality[15]. In July 2017, the official observation of new questions on labour informality began, and the first data in the series was from the July–September 2017 moving quarter[16].

In addition, other variables of interest for categorizing the employed population include labour benefits, social security, hours worked (distinguishing between contracted, actual, and normal hours).

3.1.4.2 Unemployed persons

In addition to the variables of sociodemographic characterization of the unemployed population, other economic variables can be used to classify one of its subgroups: the redundant (i.e., persons who have previously had a job the lasted for at least one month).

This subgroup can be classified according to the last employment activity. Thus, a set of variables similar to employed persons is used to obtain details such as the branch of economic activity in which they worked and their occupational category.

3.1.4.3 Outside the labour force[17]

As mentioned above, this group can be categorized into three subgroups according to their availability or the potential with which they can move within or outside the labour force (either as employed or unemployed persons). These three subgroups are inactive workers who are potentially active, the routinely inactive, and job starters.

In addition to the sociodemographic characterization, the population outside the labour force can be classified according to reasons for their situation, which include economic, seasonal, and personal reasons for not entering the labour market. Thus, inactivity can be understood not as a permanent condition but as a condition that can change quickly.

The stated reasons for being outside the labour force are classified as follows:

---

[15] The official collection of data from the new labour informality questions began in July 2017. Thus, the first data published in the series was for the July–September 2017 moving quarter. The first publication was on 31 January 2018.

[16] The first official publication of data on labor informality was on 31 January 2018.

[17] In the previous conceptual framework, this group was identified as “economically inactive persons”.

- **Reasons related to season:** The population classified as outside the labour force for related to season consists of persons who state that they “are waiting for the season of greatest activity”. At the same time, they indicate that they are available to work. This is the group most likely to enter the labour force within a short period or whenever conditions in the labour market change. Although not a very large population, this group can influence measurements of labour force pressure.
- **Reasons related to discouragement:** Those who state this reason did not look for employment work in the last four weeks. They became tired of looking for a job because they think they will not find one because of their age, because of the excessive arrangements needed to start an own-account activity, or because of other circumstances. They also state that they are available to start a job in the two weeks following the reference week. Like those who give seasonal reasons, this population generally joins those who are marginally connected to the labour market. It is the largest group according to reasons stated.
- **Other reasons:** This is the category of persons outside the labour force whose reasons for inactivity cannot be classified into any other category.
- **Temporary personal reasons:** This category is of those whose reasons for inactivity are related to health, pregnancy, or short-term family responsibilities. If they state that they are available to work, they are classified as inactive workers who are potentially active.
- **Permanent family reasons:** This category includes all persons who whose reasons for inactivity are related to care and/or domestic chores.
- **Reasons related to study:** Like the previous category, it is also an important category among the inactive. This category is associated with students.
- **Reasons related to pension or survivor's benefits:** These persons have permanent income from their previous active working life or from being widowed. The potential activity of this group is very low. Persons of this category usually remain in this condition.
- **Reasons related to retirement:** This is a population with very similar characteristics to the previous one. They are people who receive income from previous work.
- **Permanent health reasons:** Persons whose reasons for inactivity are related to their health.
- **No desire to work:** Persons who state that they do not want to work. Thus, their potential for participation is zero.

3.1.5 Principal rates and indicators of the labour market

The main indicators for measurement are derived from the conceptual framework of the ENE.

1. *Unemployment rate:* expresses the percentage of unemployed people within the labour force.

$$Unemployment\ rate = \frac{unemployed\ persons}{labour\ force} * 100$$

2. *Employment rate:* expresses the percentage of employed persons within the working-age population.

$$Employment\ rate = \frac{employed\ persons}{working\text{-}age\ population} * 100$$

3. *Participation rate:* is the percentage of persons who participate in the labour market or make up the labour force over the total number of persons of working age.

$$Participation\ rate = \frac{employed\ persons + unemployed\ persons}{working\text{-}age\ population} * 100$$

4. *Unemployment rate with available job starters (UU1)*: is the unemployment rate added to available job starters in both the numerator and the denominator. It expresses an expanded unemployment rate by adding the population that, although not yet in the labour force or inactive, has expressed their availability to work and has declared that they will start a labour activity in the near future.

$$UU1 = \frac{unemployed\ persons + available\ job\ starters}{labour\ force + available\ job\ starters} * 100$$

5. *Combined Rate of Unemployment and Involuntary Part-Time Work (UU2):* is the rate of unemployment with available starters, including available job starters in both parts of the fraction. To highlight those who are not employed full time, it also includes involuntary part-time workers in the numerator.

$$UU2 = \frac{unemployed\ persons + available\ job\ starters + involuntary\ part\text{-}time\ workers}{labour\ force + available\ job\ starters} * 100$$

6. *Combined Rates of Unemployment and the Potential Labour Force (UU3):* is the rate of unemployment with the expanded labour force, which consists of the available labour force, available job starters, and the potential labour force (inactive workers who are potentially active). This indicator accounts for the unemployment of those who participate in the labour market and those who could potentially enter it in the short term.

$$UU3 = \frac{unemployed\ persons + available\ job\ starters + inactive, potentially\ active}{labour\ force + available\ job\ starters + inactive, potentially\ active} * 100$$

7. *Labour force pressure rate:* is the ratio of unemployed persons, employed persons seeking work, and available job starters over the labour force and available job starters. Because it includes persons who declared their availability and the possibility of their starting a job soon, this indicator shows the percentage of people who pressure the labour market within the labour force.

$$Labour\ force\ pressure\ rate = \frac{unemployed + available\ job\ starters + employed\ seeking\ work}{labour\ force + available\ job\ starters} * 100$$

8. *Seasonally adjusted unemployment rate:* is an unemployment rate that eliminates exogenous seasonal factors of a non-economic nature.

9. *Informal employment rate:* expresses the percentage of persons who are employed in informal activities as a proportion of the total number of employed persons.

$$Informal\ employment\ rate = \frac{informal\ employed\ persons}{total\ employed\ persons} * 100$$

10. *Employment rate in the informal sector:* the percentage of persons working in informal economic units as a proportion of the total number of employed persons.

$$Employment\ rate\ in\ the\ informal\ sector = \frac{employed\ persons\ within\ informal\ economic\ units}{total\ employed\ persons} * 100$$

As a complement to the main employment indicators, we suggest the inclusion of other types of indicators whose variation is relevant to the analysis of relationship between work and the economy. These indicators also enable international comparisons to be made.

11. *Branch of economic activity:* Distribution of the impacts of the different branches of economic activity (according to CAENES) in the employed population.

12. *Occupational category:* Distribution of the impacts of occupational categories in the employed population. The categories included are those defined by ICSE-93.

13. *Occupational group:* Distribution of the impacts of occupational groups in the employed population, according to CIUO 08.CL.

14. *Normal and actual working time:* Distribution of the volume of employment work, measured in actual hours worked by the employed population. The separation of voluntary and involuntary part-time work is recommended in order to highlight the underutilization of the labour force.

4 Characteristics of the sampling design

The objective of the ENE's new sampling design, whose implementation began in January 2020, is to obtain more precise estimates of the main indicators[18] of the population and of the relationship of the population to the labour market in defined domains of the survey. By implementing mechanisms that regularly update the sampling frame and the sample (if necessary), non-sampling errors associated with an out-of-date sampling frame can be reduced.

4.1 Target population

The target population is all persons aged 15 years and over who habitually reside in private occupied dwellings within the national territory.

4.2 Geographical coverage and domains of the survey

The ENE seeks to be representative of the entire national territory, covering 97% of communes, including almost the entire master sampling frame[19].

The new sampling design of the ENE, which has been used since January 2020, guarantees representativeness and acceptable sampling errors[20] for the domains to be surveyed. These domains are the national, national urban, national rural, regional, the urban area of the sixteen regions, and the rural area of the metropolitan region and of the regions of O'Higgins, El Maule, La Araucanía, Los Lagos, Los Ríos, and Ñuble.

4.3 Reference period and periodicity

For the reference period of the information collected, the week (Monday to Sunday) prior to the date of the interview has been adopted as the rule, which is in line with ILO recommendations (19° CIET, 2013). For questions related to the measurement of income—which are included once per year in the last calendar quarter—the reference period is normally the previous month.

[18] These are the unemployment rate, the employment rate, and the participation rate.

[19] Because of exclusions made for operational reasons, the ENE does not cover 0.57% (35,287 dwellings) of the total number of dwellings at the national level. The exclusion process is explained in more detail in section 4.5.2.

[20] For more information, see the document "*Lineamientos para el uso de medidas de precisión en encuestas de hogares*" (INE, 2020).

The ENE is a continuous survey in which the results, based on moving quarters, are published each month.

4.4 Master sampling frame

A sampling frame is defined as the list or the procedures that enable all elements of a target population to be identified (Groves, 2004, p, 68). Frames are fundamental to the design of sample surveys. Therefore, it is of the utmost importance that the methodology for their construction be properly documented. Among the desirable properties for a sampling frame are completeness, accuracy, and timeliness (Naciones Unidas, 2009).

The property of completeness means that the sampling frame effectively covers the entire population of interest. The property of accuracy refers to using reliable information without duplication of records, etc. The property of timeliness refers to the maintenance of recent and timely information in the sampling frame. Although perfect timeliness is impossible, it is important that a sampling frame have procedures in place to minimize non-sampling errors arising from obsolescence.

The sampling frame most frequently used in social surveys consists of a list of geographical units[21] whose characterization and updating are based on census information. From this list are chosen the dwellings of the sampling frame that subsequently become part of the sample for each survey. Thus, a frame must cover the population of interest comprehensively, it must be duly updated, and it must have the required information for the construction of adequate sampling strata.

The current sample of the ENE, which became effective in January 2020, was selected from the sampling frame of dwellings (hereinafter, 2017 SFD)[22]. The 2017 SFD, in turn, was produced from the results of the Population and Housing Census of 2017 and from the lists of dwellings associated with the 2016 Precensus. The SFD is continuously updated between censuses from administrative records and, if necessary, from a process of enumeration of prioritized geographic units.

[21] According to the United Nations (2009), this type of frame is called an area frame. It consists of the geographical units of a country in a hierarchical order.

[22] Until December 2019, the sample of the ENE was selected from a sampling frame of two lists of units. The first list, the sampling frame of blocks (2016 SFB), was updated from the lists of the 2016 Precensus. The 2016 SFB was used for the city (CD) stratum, which consists of cities and large urban centers. The second list, the sampling frame of sections (2016 SFS), was also partially updated from the lists of the 2016 Precensus. The 2016 SFS was used for the RAU (Rest of Urban Area) and rural strata.

The 2017 SFD is made up of primary sampling units (PSU), corresponding to geographical areas that are homogeneous in terms of the number of “frame dwellings”. Frame dwellings are defined as private dwellings with present or absent inhabitants[23] that were recorded in the 2017 Population and Housing Census. Urban and rural areas are distinguished. The average size of urban areas is 200 dwellings with an interval of 20%. That is, the number of dwellings per PSU is in the range of 160 to 240. On the other hand, the rural area has an average size of 90 dwellings and a range of 20%. Thus, the number of dwellings per PSU is in the range of 70 to 110.

The attributes of the new frame include the generation of PSUs that are homogeneous in size, the socioeconomic stratification of the frame, and the identification of areas for special treatment. These attributes contribute to the precision of estimates derived from surveys that use this frame.

4.4.1 Geographical coverage of the sampling frame

The new sample framework covers 345 of the 346 communes that make up the national, continental, and insular territory. The commune excluded is Chilean Antarctica, part of the Magallanes region.

4.4.2 Geographical stratification of the sampling frame

The 2017 SFD contains variables that identify the country's political and administrative divisions[24] and the rural and urban areas of each commune of the country.

The definition of urban and rural is given in the specifications of the Census of Population and Housing of 2017. Thus, urban and rural entities were defined in the 2017 Census (INE, 2019a).

- Urban entity: a human settlement with continuity and concentration of constructions in a regular grid pattern with a population greater than 2,000 inhabitants, or between 1,001 and 2,000 inhabitants, where less than 50% of the population state that they work in primary activities. For example, the commune of Coquimbo has four urban entities: Coquimbo, Tongoy, El Peñón, and Guanaqueros. Together these entities make up the urban area of the commune of Coquimbo.
- Rural entity: concentrated or dispersed human settlement with 1,000 or fewer inhabitants, or between 1,001 and 2,000 inhabitants, where more than 50% of the

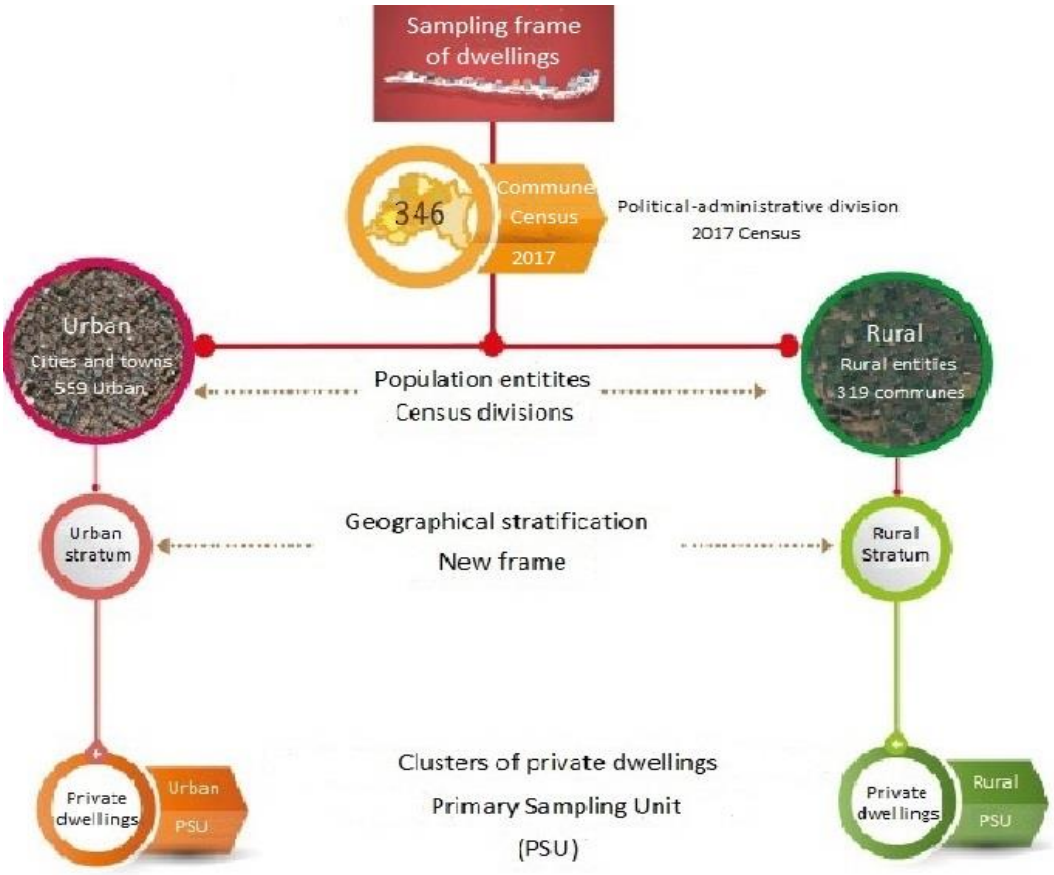
[23] Seasonal and collective housing is excluded.

[24] These are region, province, and commune.

population state that they work in primary activities. For example, the commune of Talca has three rural entities: Villa Illinois, El Porvenir, and El Oriente.

Of the 345 communes, 319 communes have an urban area. Of these, 27 communes contain the urban stratum exclusively. For its part, the rural stratum is found in 319 communes, of which 27 contain it exclusively.

Figure 6 Geographical stratification of the 2017 sampling frame of dwellings (SFD)



Source: Produced by the National Statistics Institute

4.4.3 Structure of primary sampling units

One of the main characteristics of the 2017 SFD is the definition of PSUs of homogeneous size in terms of private dwellings in both rural and urban areas. PSUs are constructed from information from lists of the 2017 Census and the 2016 Precensus. This makes them more stable units over time, decreasing the loss of units in the survey and reducing the risk of non-sampling errors. Likewise, the use of homogeneous PSUs also has benefits from the point of

view of the sampling design because selecting a similar number of dwellings in each PSU reduces the variability of the value of the expansion factors and decreases possible outliers. Therefore, it improves the precision of the estimates.

4.4.3.1 Criteria for urban PSUs

In the creation of urban PSUs, the 559 urban entities defined by the 2017 Census were taken as a starting point. These entities are a political-administrative division at the level of commune and consist of entities such as cities and towns.

Figure 7 Stratification of urban entities

Source: Produced by the National Statistics Institute

The average size of PSUs is 200 dwellings, and their range is between 160 and 240 dwellings, with certain exceptions. In the enumeration, only “frame dwellings” are included. As reported in the 2017 Census, these are defined as occupied private dwellings with present or absent inhabitants, while unoccupied dwellings, seasonal and collective dwellings, and dwellings for other uses are excluded.

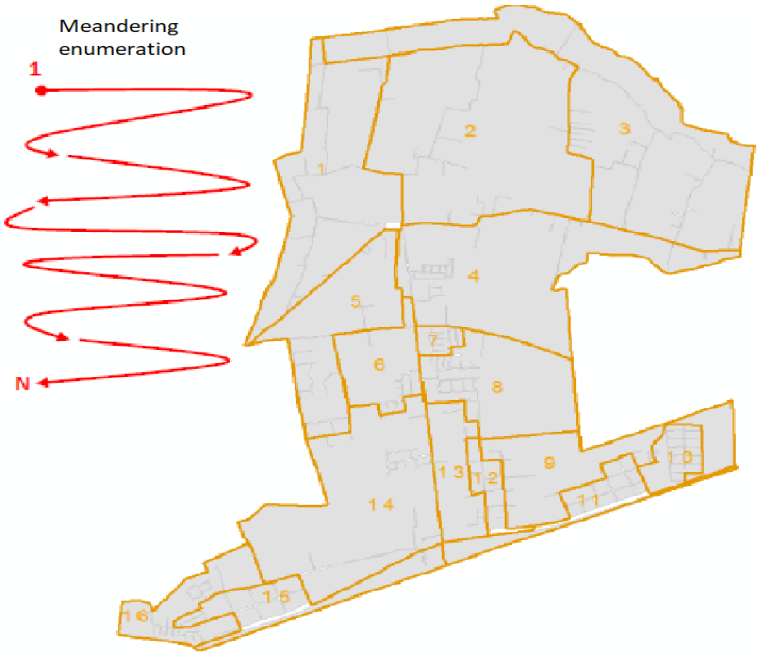
The fundamental criterion in the construction of PSUs is the maintenance of compact and homogeneous units for an efficient two-stage sampling and for an optimal functioning of the sample. For this reason, exceptions to the range in size of PSUs are made in order to preserve their compactness. Exceptions include PSUs of blocks that have between 240 and 300 dwellings, which exceeds the upper range. These PSUs are included because dividing such blocks would produce units below the established range and would compromise the principle of homogeneity. For blocks with complex or irregular shapes (e.g., when one block surrounds

another), the previous criterion is made more flexible to include up to 360 dwellings before they are divided.

For exceptions below the lower end of the range, it should be noted that there must be at least two PSUs within each urban division (cities, towns, etc.). For this reason, when it is necessary to divide a territory in two, the first option of division is that both PSUs have as similar number of dwellings as possible, even if they are out of the established range. In these cases, one PSU should not be within the range while the other is outside the range, although this depends on the geography of the blocks and their relationship with the environment[25].

The enumeration method of the PSUs is a meandering method that goes from north to south, starting from the top left, and "meandering" right to left towards the bottom. In these cases, the level is communal and enumeration is by strata (see figure 8).

Figure 8 Enumeration by the meandering method



Source: Produced by the National Statistics Institute

[25] For further background on the criteria for the formation of urban PSUs, see the methodological document "*Unidades primarias de muestreo, Nuevo marco muestral de viviendas*" (INE, 2019b, pp. 36-48).

4.4.3.2 Criteria for urban PSUs

Rural PSUs are constructed from rural entities surveyed in the 2017 Census, which total 319 communes nationwide.

Figure 9 Stratification of rural entities

Source: Produced by the National Statistics Institute

The average size of PSUs is 90 dwellings, and their range is between 70 and 110 dwellings, with certain exceptions. The objective is to maintain homogeneous units.

Sparse rural areas and village entities are treated distinctly. Therefore, the enumeration and grouping of dwellings for the formation of PSUs is carried out at the level of the census sector. At the same time, the connectivity between PSUs is considered in order to make the survey more efficient.

If boundaries of rural PSUs do not contravene the principles of homogeneity and connectivity, they should be supported by existing elements of mapping (e.g., road axes, watercourses, political-administrative boundaries, and boundaries) for census purposes. As in the case of urban PSUs, the enumeration of PSU uses the meandering method, moving from north to south, starting from the top left.

Normally, when the number of dwellings in rural PSUs is outside the established range, it is due to the particularities of the country's geography. For example, undersized PSUs are permitted when dealing with areas enclosed between the communal boundary and the urban area. Similarly, communes may be permitted if they have only one rural PSU that exceeds the established range. This upholds the principle of homogeneity better than dividing the PSU in two with each size below the range. These special cases include PSUs with island entities, PSUs

where connectivity is outside the commune or even the region, and cases where the route of the PSU crosses an urban area[26].

4.4.4 Socioeconomic stratification of the sampling frame

Among the new features incorporated into the 2017 SFD is an update of the socioeconomic stratification at the level of the PSU. The update is based on an analysis of a range of variables collected during the 2017 Census. The result is a stratification variable with three categories, called socioeconomic levels (SEL). The process for obtaining this variable is detailed below[27].

For the determination of the socioeconomic strata, various methods and levels of aggregation of information were considered. The methods were various multivariate classification algorithms (such as Principal Component Analysis (PCA)), k-means clustering, and Principal Component Analysis by means of Alternating Least Squares (PRINCALS). Algorithms based on the level of persons (which are then grouped into PSUs) were considered, as was working directly at the level of PSU from the beginning. Other points to consider were the number of strata to be constructed and the variables to be used.

From the various combinations of techniques, groupings, quantities of strata, and variables, 180 stratification vectors were obtained, each of which had between three and five strata. To evaluate their performance, the design effect of twenty-four indicators was calculated from which the stratification with the least variance could be chosen. These indicators were the following:

- Employment rate according to sex
- Unemployment rate according to sex
- Inactivity rate according to sex
- Percentage of foreigners
- Percentage of persons in the primary labour force within economically active population (EAP), according to sex
- Percentage of single-person households
- Percentage of the population of employed women according to branch
- Percentage of the population of employed men according to branch
- Percentage of persons by educational level, according to sex

[26] For further background on the criteria for the formation of urban PSUs, see the methodological document "*Unidades primarias de muestreo, nuevo marco muestral de viviendas*" (INE, 2019b, pp. 99-119).

[27] Further background is available in the document "*Estratificación socioeconómica del marco muestral de viviendas 2017 (MMV 2017)*", (INE, 2019c).

After the evaluation, the method used for the stratification of the 2017 SFD was the optimum stratification of the first principal component of the PCA, which includes the following indicators: the percentage of persons with higher education, the employment rate, the percentage of dwellings with a high index of structural condition, and the total number of live births. The choice of the number of strata (three or four) was based on the analysis of PSU distributions according to the intersection of commune and area. Because the low prevalence of PSUs in remote regions (particularly in rural areas), it was decided to use three strata.

These results are for 35,090 PSUs, while the 2017 SFD has 35,149 PSUs. This disparity can mainly be attributed to the lack of information on dwellings in the census database. The 2017 SFD includes dwellings that were not occupied at the time of the census.

As a result, fifty-nine PSUs were left without socioeconomic classification because the required census information was not available for use in the methodology of socioeconomic stratification. These PSUs were assigned the socioeconomic status (SES) of neighboring PSUs. If adjacent PSUs had varying SES, the unclassified PSUs were assigned the preponderant level among the neighboring PSUs. This process was carried out through cartographic observation of the PSUs.

4.4.5 Special areas

In the construction of a sample frame, it is necessary to identify special areas susceptible to exclusion. The difficulty in gathering data on these areas is often due to factors such as climate, transportation, cost of access, and restricted access. For the 2017 SFD, special areas (SA) were defined.

The 2017 SFD classifies SA into eight types according to the complexities in carrying out operational work. The eight types are as follows:

1) Travel time
2) Climate
3) Altitude
4) Special transportation
5) Insularity
6) Pedestrian access
7) Road conditions
8) Restricted access

This enables the calculation, at the level of PSU, of the percentage of dwellings that are difficult to access. In addition, the specific reasons why they were defined as special areas can be identified. Exclusion can be made at the level of PSU rather at the level of commune. Therefore, the coverage of the 2017 SFD is greater than the coverage of its predecessor.

At the level of PSU, assessment of exclusion can be made in conjunction with regional teams. Thus, the detailed information gathered during fieldwork can be coordinated and evaluated for further operations.

4.4.6 Definitive structure of the master sampling frame

Enumerators measured 35,149 separate PSUs from the urban and rural strata with 6,145,493 frame dwellings. The urban stratum has 26,195 PSUs in 5,336,492 frame dwellings, and the rural stratum has 8,954 PSUs in 809,001 frame dwellings.

Although limits were established to the number of dwellings per PSU, exceptions had to be made in some cases. The urban area has 1,417 PSUs outside the range of 160 to 240 dwellings, and the rural area has 198 PSUs outside the range of 70 to 110 dwellings.

In socioeconomic stratification, 35,090 PSUs were classified into low (1), medium (2) and high (3) strata. At the national the level, this resulted in 12,728 PSUs of the low stratum, 15,518 PSUs of the medium stratum, and 6,844 PSUs of the high stratum.

Table 2 PSUs with size outside the range, by region

	Urban			Rural		
Region	**Total PSUs**	**PSU < 160 dwellings**	**PSU > 240 dwellings**	**Total PSUs**	**PSU < 70 dwellings**	**PSU > 110 dwellings**
Total	**26,195**	**363**	**1,054**	**8,954**	**65**	**133**
Arica y Parinacota	331	2	5	74	-	-
Tarapacá	504	10	15	66	1	1
Antofagasta	895	7	19	70	3	1
Atacama	462	1	8	145	1	1
Coquimbo	1,092	20	31	646	2	13
Valparaíso	3,004	46	115	683	13	20
Metropolitan	11,135	188	580	919	9	19
O'Higgins	1,196	25	56	954	13	20
Maule	1,374	15	38	1,171	7	27
Ñuble	591	4	16	634	-	5
Biobío	2,372	22	99	754	8	10

	Urban			Rural		
Region	Total PSUs	PSU < 160 dwellings	PSU > 240 dwellings	Total PSUs	PSU < 70 dwellings	PSU > 110 dwellings
Araucanía	1,209	9	27	1,189	-	5
Los Ríos	495	2	10	482	1	2
Los Lagos	1,100	10	33	996	2	5
Aysén	153	-	-	119	2	-
Magallanes	282	2	2	52	3	4

Source: Produced by the National Statistics Institute

Table 3 Socioeconomic stratification of the 2017 sampling frame of dwellings (SFD), by region

	Socioeconomic stratification											
	Unclassified			Low			Middle			High		
	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total	Urban	Rural	Total
National	**48**	**11**	**59**	**5,846**	**6,882**	**12,728**	**13,611**	**1,907**	**15,518**	**6,690**	**154**	**6,844**
Arica y Parinacota	-	-	-	108	70	178	178	4	182	45	-	45
Tarapacá	2	1	3	133	60	193	248	4	252	121	1	122
Antofagasta	1	-	1	203	51	254	481	17	498	210	2	212
Atacama	-	-	-	226	124	350	161	19	180	75	2	77
Coquimbo	-	-	-	356	590	946	568	54	622	168	2	170
Valparaíso	4	3	7	671	404	1,075	1,780	268	2,048	549	8	557
Metropolitan	35	4	39	1,159	338	1,497	5,726	524	6,250	4,215	53	4,268
O'Higgins	-	-	-	356	746	1,102	696	205	901	144	3	147
Maule	1	-	1	459	1,023	1,482	735	145	880	179	3	182
Ñuble	-	-	-	225	579	804	295	50	345	71	5	76
Biobío	3	2	5	887	681	1,568	1,053	66	1,119	429	5	434
La Araucanía	-	1	1	476	1,066	1,542	558	118	676	175	4	179
Los Ríos	-	-	-	191	416	607	236	60	296	68	6	74
Los Lagos	1	-	1	368	683	1,051	563	286	849	168	27	195
Aysén	-	-	-	17	44	61	121	58	179	15	17	32
Magallanes	1	-	1	11	7	18	212	29	241	58	16	74

Source: Produced by the National Statistics Institute

From information on special features identified in each PSU of the 2017 SFD, 973 PSUs with at least one such feature were identified as special PSUs. This number represents 2.77% of the total number of PSUs in the 2017 SFD. These PSUs include 87,831 dwellings, or approximately 1.43% of the total number of dwellings in the 2017 SFD.

Table 4 Special PSUs, by region

Region	Total PSUs of the frame	Special PSU total	Special PSU %	Frame dwellings total	Special frame dwellings total	Special frame dwellings %
National	**35,149**	**973**	**2.77%**	**6,145,493**	**87,831**	**1.43%**
Arica y Parinacota	405	9	2.22%	72,602	828	1.14%
Tarapacá	570	22	3.86%	107,473	2,046	1.90%
Antofagasta	965	10	1.04%	186,917	878	0.47%
Atacama	607	11	1.81%	105,554	924	0.88%
Coquimbo	1,738	45	2.59%	278,060	4,030	1.45%
Valparaíso	3,687	38	1.03%	673,541	3,413	0.51%
Metropolitan	12,054	17	0.14%	2,352,958	1,543	0.07%
O'Higgins	2,150	71	3.30%	331,870	6,331	1.91%
Maule	2,545	171	6.72%	386,081	15,778	4.09%
Ñuble	1,225	17	1.39%	179,827	1,526	0.85%
Biobío	3,126	75	2.40%	553,538	6,709	1.21%
La Araucanía	2,398	162	6.76%	355,092	14,679	4.13%
Los Ríos	977	77	7.88%	143,005	6,937	4.85%
Los Lagos	2,096	187	8.92%	316,294	16,847	5.33%
Aysén	272	44	16.18%	41,682	3,996	9.59%
Magallanes	334	17	5.09%	60,999	1,366	2.24%

Source: Produced by the National Statistics Institute

4.5 Sampling frame of the National Employment Survey (ENE)

The sampling frame of the National Employment Survey is constructed from the master sampling frame according to the objectives of the survey, the established levels of estimation, and special areas to be excluded.

4.5.1 Geographical stratification

Geographical stratification begins with the political-administrative division and the identification of areas in urban and rural sectors of the master sampling frame. The aim of stratification is to create a geographical division that meets the minimum requirements of PSUs for survey rotations and that reflects to the spatial clustering of the labour market.

In order to have enough PSUs for updating the sample in the periods between censuses, each stratum must have a minimum of 36 PSUs. In addition, urban and rural areas were stratified separately so that the rural segment of a commune can share a stratum with other communes, while an urban segment constitutes a stratum by itself.

To account the clustering of the population, the process of building strata began with groups of communes that share conurbations and that meet the criterion of containing at least 36 PSUs. In regions where these conurbations do not coincide with regional capitals, a group of communes is also considered a stratum if they meet the minimum number of PSUs.

The rest of the communes belonging to the same province were then grouped into strata. In these cases, an entire province might constitute one stratum, or it might contain several if its population was deemed excessive.

A Principal Component Analysis (PCA) was used when the population was deemed excessive. Thus, strata such as the conurbation of Greater Concepción, the conurbation of Greater Santiago, and the province of Linares were created.

At the end of this process, 163 geographical strata were obtained. The strata were distributed throughout the national territory, except for the exclusion zones detailed in the following section. Because eleven of the 163 strata have no other communes with which to be grouped, they do not meet the minimum number of PSUs. These eleven strata are the exceptions to the rule. The regional distribution of the strata is shown in Table 5, and the exceptional strata (with fewer than 36 PSUs) are shown in Table 6. The complete list of geographical strata can be found in Annex 1.

Table 5 Number of geographical strata by region and area

	Urban areas	Rural areas	Total
National	**88**	**75**	**163**
Arica y Parinacota	1	1	2
Tarapacá	2	2	4
Antofagasta	3	2	5
Atacama	3	3	6
Coquimbo	5	5	10
Valparaíso	12	9	21
Metropolitan	15	9	24
O'Higgins	6	6	12
Maule	9	10	19
Ñuble	4	4	8
Biobío	7	5	12
Araucanía	6	4	10
Los Ríos	5	5	10
Los Lagos	6	6	12
Aysén	2	2	4
Magallanes	2	2	4

Source: Produced by the National Statistics Institute

Table 6 Strata that do not meet the minimum number of PSUs

Region	Area	Geographical stratum	PSUs	Reason
Tarapacá	Urban	Province of Del Tamarugal	27	No other urban areas in the region, except the regional capital.
Tarapacá	Rural	Conurbation of Iquique - Alto Hospicio	13	To maintain the regional capital and to ensure its presence in the sample of the stratum.
Antofagasta	Rural	Province of Antofagasta (Antofagasta, Mejillones, Sierra Gorda, Tal Tal)	20	No other rural areas in the province.
Atacama	Rural	Conurbation of Copiapó (Copiapó, Tierra Amarilla)	33	To maintain the regional capital and to ensure its presence in the sample of the stratum.
Atacama	Rural	R.R. Atacama (Province of Chañaral and Commune of Caldera)	30	Caldera was added to a stratum outside its province to maintain the conurbation of the regional capital and obtain two strata of similar size.
Valparaíso	Urban	R.P. Valparaíso (Casablanca)	30	No other adjacent urban areas, except for the regional capital
Valparaíso	Rural	Conurbation of Valparaíso (Valparaíso, Concón, Quilpué, Villa Alemana)	33	To maintain the regional capital and to ensure its presence in the sample of the stratum.
Valparaíso	Rural	R.P. Valparaíso (Puchuncaví, Quintero)	34	No other rural areas in the province.
Metropolitan	Urban	R.P. Cordillera	35	All adjacent urban areas form part of conurbations.
Magallanes	Rural	Punta Arenas + Laguna Blanca + Río Verde + San Gregorio	28	To maintain the province of the regional capital and to ensure its presence in the stratum.
Magallanes	Rural	R.R. Magallanes	23	Rest of the communes in the Magallanes region.

Source: Produced by the National Statistics Institute

4.5.2 Structure of the sampling frame

With the inclusion of the geographical strata, the 2017 SFD already has all the necessary information to be used in the ENE according to the defined sampling design. However, the PSUs to be excluded from the study must be determined. PSUs susceptible to exclusion have been identified as special areas of the 2017 SFD and areas excluded in the previous survey. To decide which of these PSUs to exclude, INE's technical team coordinated with the regional survey teams to consider the capacities of existing coverage[28].

The result was the exclusion of 355 PSUs, which represent 1% of the PSUs at the national level. Ten of the 345 communes covered by the master sampling frame are totally excluded, a

[28] A list of PSUs subject to exclusion was sent to each regional team so that they could determine whether to exclude the PSU from the survey based on their knowledge of the area.

substantial concentration of these exclusions from region of Los Lagos, where the province of Palena is totally excluded. At the national level, 0.57% of the dwellings in the frame were excluded.

Table 7 shows the regional distribution of PSUs and frame dwellings, special and excluded.

Table 7 Special PSUs and PSUs excluded from the sampling frame of the ENE, by region

Region	Total PSUs of the frame	Special PSUs total	Special PSUs %	Total PSUs excluded	Excluded PSUs, %	Frame dwellings total	Special frame dwellings total	Special frame dwellings %	Excluded frame dwellings total	Excluded frame dwellings %
National	**35,149**	**973**	**2.77%**	**355**	**1.01%**	**6,145,493**	**87,831**	**1.43%**	**35,287**	**0.57%**
Arica y Parinacota	405	9	2.22%	8	1.98%	72,602	828	1.14%	736	1.01%
Tarapacá	570	22	3.86%	13	2.28%	107,473	2,046	1.90%	1,195	1.11%
Antofagasta	965	10	1.04%	4	0.41%	186,917	878	0.47%	359	0.19%
Atacama	607	11	1.81%	6	0.99%	105,554	924	0.88%	511	0.48%
Coquimbo	1,738	45	2.59%	0	0.00%	278,060	4,030	1.45%	0	0.00%
Valparaíso	3,687	38	1.03%	28	0.76%	673,541	3,413	0.51%	3,921	0.58%
Metropolitan	12,054	17	0.14%	0	0.00%	2,352,958	1,543	0.07%	0	0.00%
O'Higgins	2,150	71	3.30%	6	0.28%	331,870	6,331	1.91%	524	0.16%
Maule	2,545	171	6.72%	5	0.20%	386,081	15,778	4.09%	457	0.12%
Ñuble	1,225	17	1.39%	8	0.65%	179,827	1,526	0.85%	726	0.40%
Biobío	3,126	75	2.40%	24	0.77%	553,538	6,709	1.21%	2,064	0.37%
La Araucanía	2,398	162	6.76%	49	2.04%	355,092	14,679	4.13%	4,494	1.27%
Los Ríos	977	77	7.88%	23	2.35%	143,005	6,937	4.85%	2,023	1.41%
Los Lagos	2,096	187	8.92%	167	7.97%	316,294	16,847	5.33%	16,895	5.34%
Aysén	272	44	16.18%	11	4.04%	41,682	3,996	9.59%	1,177	2.82%
Magallanes	334	17	5.09%	3	0.90%	60,999	1,366	2.24%	205	0.34%

Source: Produced by the National Statistics Institute

5.3 Stratification of the ENE sampling frame

The sampling design of the ENE includes information from both the geographical and socioeconomic stratification of the 2017 SFD. Sampling strata are formed f the intersection between geographical and socioeconomic stratification, if the geographic strata meet certain conditions. To disaggregate a geographical stratum by its socioeconomic level (SEL), it must meet the following requirements:

- Have more than 150 PSUs
- Be of sufficient socioeconomic diversity as follows:
  - The geographical stratum must not contain more than 80% of the PSUs in a single SEL, or
  - it must have approximately 40% of its PSUs in one SEL and 60% in another SEL
- Ensure that all resulting strata have at least 36 PSUs

Under these criteria, 232 sampling strata are obtained. Of these, 122 strata come directly from geographical stratification, and 110 strata use information from socioeconomic stratification. The regional distribution of the strata is shown in Table 8.

Table 8 Number of sampling strata by region and area

	Urban areas	Rural areas	Total
National	**153**	**79**	**232**
Arica y Parinacota	3	1	4
Tarapacá	4	2	6
Antofagasta	5	2	7
Atacama	5	3	8
Coquimbo	8	5	13
Valparaíso	22	9	31
Metropolitan	33	10	43
O'Higgins	8	7	15
Maule	13	10	23
Ñuble	6	4	10
Biobío	15	5	20
Araucanía	8	4	12
Los Ríos	7	5	12
Los Lagos	11	8	19
Aysén	2	2	4
Magallanes	3	2	5

Source: Produced by the National Statistics Institute

The complete list of sampling strata can be found in Annex 2.

4.6. Statistical units

4.6.1 Sampling unit

There are two sample units.

a) The primary sampling unit (PSU) is defined as a homogeneous cluster in terms of the number of private dwellings[29] of which it consists.
b) Secondary Sampling Unit (SSU) is defined as a private occupied dwelling within the selected PSUs.

As mentioned in section 4.4 above, the average size of PSUs in the urban area is 200 private dwellings with a range of 160 to 240, excluding seasonal dwellings. The average size of rural PSUs is 90 private dwellings with a range of 70 to 110, excluding seasonal dwellings.

4.6.2 Unit of information and analysis

In order to meet the objective of the survey, the units of information and analysis are defined as those persons who form part of the target population (i.e., persons aged 15 years and over who habitually reside in households located in private occupied dwellings within the national territory). However, the sociodemographic characteristics of all persons residing in the selected dwelling are recorded by means of the questionnaire constructed for this purpose.

[29] This excludes seasonal housing.

5 Sample size

One objective established for the new sampling design of the ENE is to reach a greater level of precision in its estimates. Therefore, the calculation of sample sizes for the domains of the survey was based on the new stratification of the sample frame.

5.1 Sampling strategy

The sampling design of the ENE 2020 has been defined as probabilistic, stratified, and two-stage.

The sampling strata are a combination of the “geographical stratum and the socioeconomic stratum” wherever such a combination can be made (see section 4.5.3).

5.1.1 Considerations about the sample size

The calculation of the sample size uses information from the 2017 Census of Population and Housing and the sample frame of the ENE, from which simulations can be constructed for the following purposes:

- Estimating design effect ($deff$) that arises from the use of complex sampling instead of simple random sampling
- Studying the behavior of the sample
- Assessing the relevance of different levels of estimation
- Making various decisions regarding sample size

The main parameter of interest for obtaining the sample size is the unemployment rate, which is calculated as the number of unemployed persons over the number of persons participating in the labour force. It is estimated by means of a ratio estimator.

Table 9 Characteristics of parameters of interest

Parameter	Description
Parameter to be estimated	Unemployment rate
Design variable	$X_k = \begin{cases} 0, \text{if person } k \text{ belongs to the labour force and is not unemployed} \\ 1, \text{if person } k \text{ belongs to the labour force and is unemployed} \end{cases}$
Associated estimator	$Ratio\ estimator\ p = \frac{Number\ of\ unemployed\ persons}{Number\ of\ persons\ in\ the\ labour\ force}$

The sample-size simulations consist of a series of scenarios in which PSUs are selected from the ENE sampling frame. Next, dwellings belonging to these PSUs are selected using the 2017 Census as the sample frame. Based on these simulations, it was decided to survey ten dwellings per PSU, except in the urban areas of the regions of Atacama, Aysén and Magallanes, where eight dwelllings per PSU were surveyed[30]. Likewise, the design effect at the region-area level is estimated and used in the calculation of the final size.

5.2 Calculations of the sample size

As mentioned, the sampling design of the ENE is based on a complex, stratified sampling with two stages of selection. Thus, the calculation of the sample size is obtained by the following stages:

1. Obtaining census parameters. From the 2017 Population and Housing Census[31], the variance of the variable of interest was calculated at the level of the intersection of region and area crossing (S_{ra}^2)

$$S_{ra}^2 = P_{ra} \cdot (1 - P_{ra}) \qquad (1)$$

Where P_{ra} is the proportion of persons who are unemployed[32] at the level of region-area.

2. Initial sample size calculation. The first sample size of persons is obtained through a simple random sampling formula, which includes an expected absolute error (e_{ra}). In

[30] Section 5.5 includes a discussion on the methodology for determining the optimum number of dwellings per PSU.

[31] It should be noted that the 2017 Population and Housing Census and the ENE differ in their manner of determining unemployment. These differences are in the reference period and the number of questions asked to determine the condition of unemployment. Nevertheless, this estimate can be taken as an approximation.

[32] $P_{ra} = \frac{Number\ of\ unemployed}{Number\ in\ the\ labour\ force}$

addition, the confidence level is established at 95%, and it is associated with the $t_{1-a/2,v-1}$ percentile of the t-student distribution with v degrees of freedom. However, because the degrees of freedom vary for each domain of the survey, a fixed value of two has been chosen as a measure of simplicity and robustness when making inferences from the survey.

$$\widetilde{m}_{0,ra} = \frac{2^2 \cdot s_{ra}^2}{e_{ra}^2} \quad (2)$$

Where,

$\widetilde{m}_{0,ra}$: Sample size of people for the domain of region-area, under simple random sampling.

3. Incorporation of design effect. The use of a complex multi-stage design requires the consideration of the design effect, which leads to the loss of efficiency in complex sampling when compared to simple random sampling. To account for the design effect, the initial sample size increases according to equation (3).

$$\widetilde{m}_{1,ra} = \widetilde{m}_{0,ra} \cdot deff_{ra} \quad (3)$$

Where,

$\widetilde{m}_{1,ra}$: Sample size of persons for the domain of region-area, corrected for design effect.

$deff_{ra}$: Design effect at the level of dwelling for the domain of region-area. The estimate was based on sample simulations[33] from the selection framework and the base of persons from the 2017 Census.

4. Size at the level of dwelling: Once the sample size from equation (3) with information at the number of persons in the labour force[34] the size in terms of number of dwellings and PSU must be determined. This is done by dividing the sample size at the level of person by the average number of persons in the dwellings who participate in the labour force ($\overline{FT}ra$). This data is obtained from the 2017 Census.

$$m_{1,ra} = \frac{\widetilde{m}_{1,ra}}{\overline{FT}ra} \quad (4)$$

[33] For more information, see the report *Simulaciones de tamaños muestrales de ENE 2020*.

[34] Economically active population (i.e., the labour force) is defined as persons of working age who, during the reference week, meet the requirements to be included in the category of employed or unemployed persons.

Where,
$m_{1,ra}$: Sample size of dwellings for the domain of region-area

5. Finite population adjustment: With the size at the level of dwelling, information from the sampling frame can be used to implement the sampling adjustment for finite populations. In this case, the number of dwellings in the framework of the ENE (M_{ra}) is used.

$$m_{2,ra} = m_{1,ra} \cdot \left(\frac{1}{1+\frac{m_{1,ra}}{M_{ra}}} \right) \quad (5)$$

Where,

$m_{2,ra}$: Sample size of dwellings for the domain of region-area, adjusted for finite populations

6. Operational adjustments $(m'_{2,ra})$: the sizes obtained are rounded up to the nearest multiple of 10, in order to ensure that ten dwellings are, in theory, measured in all PSUs. The sizes obtained are then compared with the sizes of the previous version of the survey and their operational feasibility is verified. The sizes are then distributed to the sampling strata according to the following criteria:

a) All strata must have a sample of at least six[35] PSUs.
b) No stratum may have a sampling fraction greater than 50%.

The application of these criteria can modify the sample sizes obtained up to the preceding point. However, the aim is to ensure that there is no significant impact on the expected levels of error.

5.3 Distribution of the sample size

The sample size is allocated with the aim of identifying the number of units that should be selected in each stratum. The resulting size of this process is reported in the last column of table 14. The method used is Neyman's optimal allocation, which assigns a greater sample to larger strata and to strata with a greater standard deviation in the parameter of interest, as follows:

$$m_h = m'_{2,ra} \cdot \frac{s_h \cdot M_h}{\sum_h s_h \cdot M_h} \quad (6)$$

[35] This is to ensure that there are at least of two PSUs per month and that the stratum contributes to the calculation of variance.

Where,

m_h: The quantity of the sample attached to stratum h

$m'_{2,ra}$: The sample size by region r and area a. (Because this value has already been adjusted for operational reasons, it does not necessarily match the values reported in column $m_{2,ra}$ of table 14.)

S_h: Standard deviation of the parameter of interest in stratum h, calculated with data from the 2017 Census

M_h: Number of occupied private dwellings in the stratum h as reported in the ENE frame

The size of dwellings obtained at the level of stratum is divided by eight or 10, as appropriate (see section 5.5), rounded up, and expressed in terms of PSUs. This result is modified to ensure compliance with the criteria mentioned in the previous section that all the strata have a sample of at least six PSUs, none of which may have a sampling fraction greater than 50%.

Because the sampling fraction criterion is implemented through a truncation of sample size in the stratum, the use of these criteria results in a distortion in the sample allocation and a loss of the required sample at certain levels of estimation. This may happen, for example, in strata that show a high level of variability of the phenomenon. These strata are given a large sample size compared to other strata of the same level of estimation, but they end up with a fraction of over 50%[36] of the sample. Thus, the complete level of estimation may lose a significant proportion of the sample. To avoid this, the lost proportion of the sample is redistributed among other strata and it is subtracted from levels of estimation that have more sample than necessary as a result of the criterion for a minimum number of PSUs.

Finally, the sample size per stratum at the level of PSU is obtained from the allocation process.

5.4 Calculation of the oversample

One of the changes proposed in the ENE 2020 sampling design is the calculation of an oversample as an alternative to the replacement mechanism that was previously used. The purpose of the oversample is to counteract the effects of non-response and loss of sampling data in successive rounds of the survey.

Panel surveys, such as the ENE, face considerable non-response problems, which are reflected in reductions in the sample size from period to period. They are due to factors such as mobility, death, and withdrawal of cooperation from informants.

[36] These cases occurred in two urban strata in the region of Aysén.

To estimate the oversample, non-attainment rates of the ENE are calculated[37] to account for sample loss in each rotation. The condition of attainment or non-attainment is determined by the contact registration codes[38] indicated on the Household Registration Card (HRC)[39], which was used in the previous version of the survey. With this classification, the situation of the dwellings can be categorized into thirteen different codes and four groups.

The attainment rate for visited dwelling[40] (TL_{gest}) is calculated using the following expression:

$$TL_{gest} = \frac{dwellings\ attained}{dwellings\ attained + (type\ A\ dwellings) + (type\ B\ dwellings) + (type\ C\ dwellings)} \qquad (7)$$

The rate of non-attainment (TNL_{gest}) is calculated by the following formula:

$$TNL_{gest} = 1 - TL_{gest} \qquad (8)$$

This form of calculation is used to determine the required oversample for obtaining a sample size that compensates for non-response, regardless of the cause.

Lists of dwellings visited between January 2018 and December 2018 were used in the calculation of non-attainment rates. Considering that the ENE sample is distributed across three months, the various moving quarters were reconstructed to fall within the year 2018.

In order to calculate the indicators of attainment at the highest possible level of disaggregation[41] in accordance with the new sampling design, it was necessary to locate the dwellings in the samples of the year 2018 within the new geographical strata because the ENE 2020 sample uses a different stratification than that of the previous sampling design.

[37] The non-attainment rate approximates the non-response rate. The difference between the two is in the denominator: the non-attainment rate includes the total number of units in the denominator, while the response rate excludes those units that are not eligible (seasonal dwellings, dwellings for other uses, etc.,) from the denominator.

[38] For more details on the contact registration codes, see Annex 3, Visitation codes of the ENE.

[39] Since January 2020, improvements have been implemented in the instrument previously included in the Household Registration Card (HRC). These improvements are part of the strategy to strengthen and update the sampling design of the ENE. They focus on classifying dwellings according to final disposition codes, complying with international standards, and capturing complementary information during the collection process.

[40] Visited dwellings are defined as all dwellings to which enumerators were sent to collect information. Therefore, such dwellings include the original dwellings or their replacement, regardless of whether it is as classified as "attained").

[41] For the ENE, the level at which oversampling rates could be calculated was at the geographic stratum. The data used for the estimate came from the previous sampling design.

The failure rates for the 163 geographical strata of the new design by round[42] and mobile quarter were calculated. Quarterly values were then averaged to obtain a single indicator per geographic stratum and round. Finally, a single value was calculated for the non-attainment rate at the level geographical stratum by averaging the non-attainment rates for each round[43] within each stratum. Because of budgetary constraints, the upper limit of non-attainment rates was set at 30%.

The values obtained for non-attainment rates at the level of geographical strata are listed in Annex 4.

The formula for calculating the oversample is the following:

$$(n * \bar{m})_{h\,sm} = \frac{(n*\bar{m})_h}{1-TNL_{hgest}} \qquad (9)$$

Where,

$(n * \bar{m})_{h\,sm}$: Sample size with oversampling in stratum h

$(n * \bar{m})_h$: Target sample size in stratum h

TNL_{hgest}: Non-attainment rate for dwellings visited in stratum h

5.5 Sample size of dwellings by primary sampling unit (PSU)

Because the sampling design of the ENE 2020 is probabilistic, stratified, and two-stage, calculations are needed for the average size of the secondary sampling units (dwellings) per PSU and for the sample size of the PSUs.

Generally, dwellings in the same cluster or PSU are similar in parameters of interest (e.g., income, education, occupation, etc.). This similarity may be due to their having similar levels of income, sharing the same attitudes towards the current problems, and being exposed to the same environmental conditions (climate, infectious diseases, natural disasters, etc.). As units belonging to the same cluster are often similar to one another, intraclass correlation is nearly always positive (Naciones Unidas, 2007).

[42] "Round" refers to the consecutive number of interviews of the same units of analysis.

[43] Although the previous sampling design includes six rounds at the CD (city) stratum, nine for the RAU (rest of the urban area), and twelve for the rural stratum, the new design includes six rounds in the urban and rural strata (see section 5.7). Therefore, only the first six rounds were analyzed.

The advantage of using clusters is that they considerably reduce the costs associated with conducting surveys. On the other hand, the greater the correlation between the elements of the cluster, the greater is the increase in variance. Because of the multiplier effect $(\tilde{n} - 1)$, even a small positive correlation can have a great impact. Thus, when determining the number of target sample units within each PSU, it is essential to achieve the maximum balance between minimizing costs and maximizing precision.

To determine how the clustering component in the design effect can be minimized or controlled, the following mathematical definition should be used (Naciones Unidas, 2009)[44]:

$$deff = 1 + (\tilde{n} - 1)\rho \qquad (10)$$

Where,

$deff$: Design effect

$\tilde{n}$: Average number of units from the target population[45] in the sample of the secondary sampling unit within the cluster

ρ: Value of the coefficient of homogeneity, or intraclass correlation

The above expression is a multiplicative function of intraclass correlation (ρ) and cluster size ($\tilde{n}$). The intraclass correlation of the variables under analysis cannot be controlled; however, the sample size can be adjusted to control for the design effect.

Table 10 shows how the average size of the target sample per PSU and the intraclass correlation influence the design effect.

[44] This formula does not consider stratification nor the other factor involved when clusters are not uniform in size. Therefore, the formula is not, strictly speaking, the $Deff$ formula. Nevertheless, because the clustering component is the predominant factor in $Deff$, the formula can be used as an approximation of how clustering affects the sampling design and can provide guidance on how to control it.

[45] For the ENE, the value of $\tilde{n}$ is equal to the number of dwellings in the cluster multiplied by the average number of persons in the labour force per dwelling.

Table 10 Design effects on selected combinations of the clustered sample size and intraclass correlation

$\tilde{n}$	ρ													
	0.01	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**	**0.1**	**0.15**	**0.2**	**0.35**	**0.5**
5	**1.0**	**1.1**	**1.1**	**1.2**	**1.2**	**1.2**	**1.3**	**1.3**	**1.4**	**1.4**	**1.6**	**1.8**	**2.4**	3.0
10	**1.1**	**1.2**	**1.3**	**1.4**	**1.5**	**1.5**	**1.6**	**1.7**	**1.8**	**1.9**	**2.4**	**2.8**	4.2	5.5
15	**1.1**	**1.3**	**1.4**	**1.6**	**1.7**	**1.8**	**2.0**	**2.1**	**2.3**	**2.4**	3.1	3.8	5.9	8.0
20	**1.2**	**1.4**	**1.6**	**1.8**	**2.0**	**2.1**	**2.3**	**2.5**	**2.7**	**2.9**	3.9	4.8	7.7	10.5
30	**1.3**	**1.6**	**1.9**	**2.2**	**2.5**	**2.7**	3.0	3.3	3.6	3.9	5.4	6.8	11.2	15.5
50	**1.5**	**2.0**	**2.5**	**3.0**	3.5	3.9	4.4	4.9	5.4	5.9	8.4	10.8	18.2	25.5
75	**1.7**	**2.5**	3.2	4.0	4.7	5.4	6.2	6.9	7.7	8.4	12.1	15.8	26.9	38.0
100	**2.0**	**3.0**	4.0	5.0	6.0	6.9	7.9	8.9	9.9	10.9	15.9	20.8	35.7	50.5
125	**2.2**	3.5	4.7	6.0	7.2	8.4	9.7	10.9	12.2	13.4	19.6	25.8	44.4	63.0
150	**2.5**	4.0	5.5	7.0	8.5	9.9	11.4	12.9	14.4	15.9	23.4	30.8	53.2	75.5
175	**2.7**	4.5	6.2	8.0	9.7	11.4	13.2	14.9	16.7	18.4	27.1	35.8	61.9	88.0
200	**3.0**	5.0	7.0	9.0	11.0	12.9	14.9	16.9	18.9	20.9	30.9	40.8	70.7	100.5
225	3.2	5.5	7.7	10.0	12.2	14.4	16.7	18.9	21.2	23.4	34.6	45.8	79.4	113.0
250	3.5	6.0	8.5	11.0	13.5	15.9	18.4	20.9	23.4	25.9	38.4	50.8	88.2	125.5
275	3.7	6.5	9.2	12.0	14.7	17.4	20.2	22.9	25.7	28.4	42.1	55.8	96.9	138.0

Source: Produced by the National Statistics Institute

According to the United Nations practical guidelines (2009), estimates of the design effect ($Deff$) above 3.0 are not acceptable. Thus, the table above highlights feasible scenarios.

To determine the average number of dwellings per PSU in the redesign of the ENE 2020, intraclass correlation coefficients (ρ) were calculated for each defined domain of the survey, based on data from 2017 Census. Table 11 shows the design effects obtained for each domain of the survey, according to their intraclass correlation coefficients and the different values of average units of the target population in the SSU sample within the clusters[46]. As can be seen, all the design effects are less than three, so a sample of six, eight, ten, or twelve dwellings per PSU guarantees estimates with an acceptable design effect.

[46] This is calculated as the number of dwellings in the cluster (m) multiplied by the average number of persons in the labour force per dwelling (persons, labour force).

Table 11 Design effects by domains of the survey, for varying average numbers of dwellings per PSU

	Urban						Rural					
	ρ	Pers LF x Dwl	Approximate *Deff* by ñ				ρ	Pers LF x Dwl	Approximate *Deff* by ñ			
			6	8	10	12			6	8	10	12
National	**0.010**	**1.5**	**1.08**	**1.11**	**1.15**	**1.18**	**0.012**	**1.3**	**1.08**	**1.10**	**1.13**	**1.16**
Region												
Arica y Parinacota	0.008	1.6	1.07	1.09	1.12	1.14	0.012	1.5	1.09	1.13	1.16	1.19
Tarapacá	0.010	1.7	1.09	1.12	1.15	1.19	0.010	1.3	1.07	1.10	1.13	1.16
Antofagasta	0.014	1.7	1.13	1.17	1.22	1.27	0.025	1.5	1.19	1.27	1.34	1.41
Atacama	0.015	1.5	1.11	1.16	1.20	1.24	0.017	1.3	1.11	1.16	1.20	1.24
Coquimbo	0.007	1.4	1.05	1.08	1.10	1.12	0.011	1.2	1.07	1.09	1.12	1.14
Valparaíso	0.006	1.4	1.05	1.06	1.08	1.10	0.007	1.4	1.05	1.07	1.09	1.11
Metropolitan	0.010	1.7	1.09	1.13	1.16	1.20	0.005	1.6	1.04	1.06	1.07	1.09
O'Higgins	0.008	1.4	1.06	1.08	1.11	1.13	0.007	1.4	1.05	1.07	1.09	1.11
Maule	0.007	1.4	1.05	1.07	1.09	1.11	0.011	1.2	1.07	1.10	1.12	1.15
Ñuble	0.007	1.3	1.05	1.07	1.09	1.11	0.009	1.1	1.05	1.07	1.09	1.11
Biobío	0.008	1.4	1.06	1.08	1.10	1.12	0.016	1.2	1.09	1.13	1.17	1.20
La Araucanía	0.007	1.4	1.05	1.07	1.09	1.11	0.013	1.1	1.07	1.10	1.13	1.15
Los Ríos	0.006	1.4	1.05	1.06	1.08	1.10	0.008	1.2	1.05	1.07	1.09	1.11
Los Lagos	0.007	1.5	1.06	1.08	1.10	1.12	0.008	1.2	1.05	1.07	1.09	1.11
Aysén	0.007	1.5	1.06	1.08	1.10	1.12	0.007	1.3	1.05	1.07	1.09	1.11
Magallanes	0.003	1.5	1.03	1.04	1.05	1.06	0.003	1.5	1.02	1.03	1.04	1.05

Source: Produced by the National Statistics Institute

In order to determine the average number of dwellings per PSU in the new ENE design, simulations were used. The simulations assessed whether sampling errors of the main indicators were within the limits of tolerance[47] and whether the resulting sample sizes were in accordance with operational capacities. From these results was determined the average number of dwellings per PSU that meet the objectives of the new sampling design.

Table 12 shows the absolute errors obtained for the unemployment rate in addition to the relative errors obtained for the level of unemployed persons in the urban domains through the simulations. In some cases, the absolute error is greater than 2%. For example, in the domain of urban Atacama, all estimates of the absolute error are greater than 2%, and in the urban areas of Tarapacá and Ñuble, which have twelve dwellings per PSU, absolute errors are greater than 2%. The relative error is greater than 30% in the urban domain of the Aysén Region in all cases (8, 10, and 12 dwellings) and in the Magallanes Region (10 and 12 dwellings).

[47] In each domain of the survey, a maximum permissible error of two (in absolute terms) has been set for the unemployment rate. A maximum permissible error of 30%, in relative terms, has been set for the level of unemployed persons.

Table 12 Sampling errors of the unemployment rate for urban domains, according to average number of dwellings per PSU

	Unemployment rate			Level of unemployed persons		
	Absolute error[48]			Relative error		
Region	8 Dwellings	10 Dwellings	12 Dwellings	8 Dwellings	10 Dwellings	12 Dwellings
National Urban	**0.30%**	**0.30%**	**0.40%**	**4.60%**	**5.10%**	**5.50%**
Arica y Parinacota	1.50%	1.70%	1.90%	20.50%	22.80%	25.00%
Tarapacá	1.80%	2.00%	2.10%	24.90%	28.00%	30.00%
Antofagasta	1.50%	1.60%	1.80%	18.40%	20.50%	22.40%
Atacama	2.20%	2.40%	2.60%	23.10%	25.90%	28.50%
Coquimbo	1.50%	1.70%	1.80%	18.70%	20.40%	21.90%
Valparaíso	0.80%	0.90%	1.00%	12.50%	13.80%	15.20%
Metropolita	0.50%	0.60%	0.60%	8.40%	9.30%	10.00%
O'Higgins	1.30%	1.40%	1.50%	18.80%	21.10%	23.10%
Maule	1.20%	1.30%	1.40%	17.30%	19.10%	20.90%
Ñuble	1.70%	1.90%	2.10%	21.40%	23.40%	25.60%
Biobío	0.90%	1.00%	1.10%	10.50%	11.70%	12.80%
La Araucanía	1.30%	1.40%	1.60%	18.40%	20.60%	22.40%
Los Ríos	1.50%	1.70%	1.80%	22.00%	24.60%	27.10%
Los Lagos	1.10%	1.20%	1.30%	17.20%	19.40%	21.20%
Aysén	1.70%	1.70%	1.80%	34.30%	34.60%	36.70%
Magallanes	1.30%	1.40%	1.50%	29.50%	32.60%	34.70%

Source: Produced by the National Statistics Institute

In order to make decisions about the rural domains, at least ten dwellings per PSU were simulated instead of eight, as was established for urban areas. The purpose of simulating more dwellings was to increase the efficiency of operational capacities in rural PSUs, which are extensive geographical areas with greater geographical dispersion of their dwellings.

Table 13 shows the absolute errors of the unemployment rate in rural domains as obtained through the simulations. In most domains with ten dwellings per PSU, the absolute errors were greater than 2%. Thus, in the cases of twelve and fourteen dwellings per PSU, the absolute errors were even greater.

[48] For more information, see *Informe de Simulaciones de tamaños muestrales de la ENE 2020.*

Table 13 Sampling errors of the unemployment rate for rural domains, according to average number of dwellings per PSU

	Unemployment rate			Level of unemployed persons		
	Absolute error[49]			Relative error		
Region	10 Dwellings	12 Dwellings	14 Dwellings	10 Dwellings	12 Dwellings	14 Dwellings
National Rural	**0.70%**	**0.80%**	**0.80%**	**11.80%**	**12.90%**	**13.70%**
O'Higgins	2.10%	2.20%	2.30%	37.40%	41.10%	42.40%
Maule	2.10%	2.30%	2.40%	35.10%	38.10%	40.10%
Ñuble	2.40%	2.80%	2.90%	33.90%	38.30%	40.20%
La Araucanía	2.60%	2.80%	3.00%	35.60%	39.70%	41.60%
Los Ríos	2.30%	2.50%	2.70%	39.70%	42.70%	46.10%
Los Lagos	1.60%	1.60%	1.80%	35.40%	37.70%	38.80%
Metropolitan	1.70%	1.90%	2.00%	38.50%	40.70%	45.20%

Source: Produced by the National Statistics Institute

Based on these simulations, the number of dwellings to be surveyed per PSU was set at ten for urban and rural strata, with the exception of the urban strata of the regions of Atacama, Aysén, and Magallanes. In these regions, the number was set at eight dwellings per PSU because sampling errors with ten or more dwellings were above the limits of tolerance.

Finally, table 14 shows the steps in calculation to arrive at the final target sample size, and table 15 shows the sizes at the level of dwelling and PSU and the associated simulation errors for all levels of estimation.

[49] For more information, see *Informe de Simulaciones de tamaños muestrales de la ENE 2020.*

Table 14 Distribution of the sample and results of the simulation

Region	National				Urban				Rural			
	Target dwellings	Over-sample dwellings	PSU	Simulation error	Target dwellings	Over-sample dwellings	PSU	Simulation error	Target dwellings	Over-sample dwellings	PSU	Simulation error
National	**39,480**	**54,126**	**5,511**	**0.30%**	**30,230**	**41,396**	**4,238**	**0.34%**	**9,250**	**12,730**	**1,273**	**0.69%**
Arica y Parinacota	1,130	1,480	148	1.62%	1,010	1,300	130	1.70%	120	180	18	3.75%
Tarapacá	810	1,030	103	1.96%	690	850	85	1.99%	120	180	18	5.65%
Antofagasta	1,270	1,660	166	1.62%	1,150	1,500	150	1.64%	120	160	16	6.25%
Atacama	1,380	1,950	237	1.61%	1,200	1,680	210	1.65%	180	270	27	6.55%
Coquimbo	2,270	3,250	325	1.51%	1,320	1,910	191	1.69%	950	1,340	134	2.61%
Valparaíso	4,540	6,120	612	0.88%	3,700	4,910	491	0.93%	840	1,210	121	1.93%
Metropolitan	7,560	10,660	1,066	0.58%	6,810	9,560	956	0.59%	750	1,100	110	1.76%
O'Higgins	2,400	3,170	317	1.20%	1,500	1,970	197	1.39%	900	1,200	120	2.05%
Maule	2,740	3,810	381	1.12%	1,800	2,600	260	1.29%	940	1,210	121	2.11%
Ñuble	1,890	2,520	252	1.54%	1,090	1,480	148	1.87%	800	1,040	104	2.46%
Biobío	4,490	6,120	612	0.91%	3,800	5,110	511	0.97%	690	1,010	101	2.74%
La Araucanía	2,290	3,060	306	1.24%	1,530	2,110	211	1.43%	760	950	95	2.44%
Los Ríos	1,840	2,630	263	1.37%	1,120	1,640	164	1.63%	720	990	99	2.34%
Los Lagos	2,870	4,090	409	1.02%	1,750	2,520	252	1.22%	1,120	1,570	157	1.52%
Aysén	928	1,316	160	1.35%	808	1,136	142	1.42%	120	180	18	3.47%
Magallanes	1,072	1,260	154	1.19%	952	1,120	140	1.23%	120	140	14	3.41%

Source: Produced by the National Statistics Institute

Table 15. Distribution of the sample and results of the simulation

Region	National				Urban				Rural			
	Target dwellings	Over-sample dwellings	PSU	Simulation error	Target dwellings	Over-sample dwellings	PSU	Simulation error	Target dwellings	Over-sample dwellings	PSU	Simulation error
National	**39,480**	**54,126**	**5,511**	**0.30%**	**30,230**	**41,396**	**4,238**	**0.34%**	**9,250**	**12,730**	**1,273**	**0.69%**
Arica y Parinacota	1,130	1,480	148	1.62%	1,010	1,300	130	1.70%	120	180	18	3.75%
Tarapacá	810	1,030	103	1.96%	690	850	85	1.99%	120	180	18	5.65%
Antofagasta	1,270	1,660	166	1.62%	1,150	1,500	150	1.64%	120	160	16	6.25%
Atacama	1,380	1,950	237	1.61%	1,200	1,680	210	1.65%	180	270	27	6.55%
Coquimbo	2,270	3,250	325	1.51%	1,320	1,910	191	1.69%	950	1,340	134	2.61%
Valparaíso	4,540	6,120	612	0.88%	3,700	4,910	491	0.93%	840	1,210	121	1.93%
Metropolitan	7,560	10,660	1,066	0.58%	6,810	9,560	956	0.59%	750	1,100	110	1.76%
O'Higgins	2,400	3,170	317	1.20%	1,500	1,970	197	1.39%	900	1,200	120	2.05%
Maule	2,740	3,810	381	1.12%	1,800	2,600	260	1.29%	940	1,210	121	2.11%
Ñuble	1,890	2,520	252	1.54%	1,090	1,480	148	1.87%	800	1,040	104	2.46%
Biobío	4,490	6,120	612	0.91%	3,800	5,110	511	0.97%	690	1,010	101	2.74%
La Araucanía	2,290	3,060	306	1.24%	1,530	2,110	211	1.43%	760	950	95	2.44%
Los Ríos	1,840	2,630	263	1.37%	1,120	1,640	164	1.63%	720	990	99	2.34%
Los Lagos	2,870	4,090	409	1.02%	1,750	2,520	252	1.22%	1,120	1,570	157	1.52%
Aysén	928	1,316	160	1.35%	808	1,136	142	1.42%	120	180	18	3.47%
Magallanes	1,072	1,260	154	1.19%	952	1,120	140	1.23%	120	140	14	3.41%

Source: Produced by the National Statistics Institute

5.6 Selection of sampling units

5.6.1 Selection of units of the first stage

For the selection of PSUs or clusters, the sampling design of the ENE uses a systematic random selection with probability proportional to size.

In sample-based household surveys, the method of selection with probability proportional to size (PPS) of PSUs is the most widely used by developing countries (Naciones Unidas, 2007). This method has some practical advantages when PSUs vary considerably in size. First, this method can lead to self-weighted samples. Second, it generates approximately equal sample sizes within PSUs, which results in similar workloads for enumerators, a desirable situation from a fieldwork perspective. In addition, this technique uses auxiliary information to increase the precision of survey estimates and gives greater control over the final sample size, without having to include stratification by size of PSU (Naciones Unidas, 2009).

The ENE defines the auxiliary information by the size of the PSU, which is measured by the number of its dwellings. Thus, the PSUs with a higher number of dwellings will be more likely to be selected.

It should be noted that although the first-stage selection probabilities are unequal, selecting secondary sampling units (SSUs) with equal probability makes the sample self-weighted (i.e., all sample units will have the same probability of selection). However, this assumption is not easily applied in practice because, in most cases, long gaps between censuses and updates to the frame result in outdated information on the size of the PSU.

The problem with such a procedure is that true measures of the size of PSUs are almost never known in practice. However, it is often possible to obtain valid estimates, such as population and household surveys from a recent census or some other reliable source (Naciones Unidas, 2007).

The sampling frame of INE's sample-based household surveys (particularly the ENE) has become outdated over time. Therefore, to mitigate possible distortions, the lists of dwellings of the selected PSUs are updated. The SSU sample will therefore be made from a new list, so the probability of selection of the dwellings is determined in the following manner:

$$\pi_{hij} = \pi_{hi} \cdot \pi_{hi/j} \qquad (11)$$

Where,

π_{hij}: Probability of selecting dwelling j in cluster i in stratum h

π_{hi} : Probability of selecting the cluster i in stratum h

$\pi_{hi/j}$: Probability of selecting dwelling j because it was selected in the cluster i in stratum h

Where the probability of selection of the PSU or cluster with probability proportional to size, according to its number of dwellings within the stratum, is determined in the following manner:

$$\pi_{hi} = \frac{n_h \cdot M_{hi}}{M_h} = \frac{n_h \cdot M_{hi}}{\sum_{\forall i \in \mathcal{M}_h} M_{hi}} \qquad (12)$$

Where,

π_{hi}: Probability of selecting cluster i in stratum h

n_h: Number of clusters in stratum h

M_{hi}: Number of dwellings in cluster i of stratum h

M_h: Number of dwellings in stratum h

$\mathcal{M}_h$: All clusters of the 2017 SFD that belong to stratum h

5.6.2 Selection of units of the second stage

After selecting the PSUs, dwellings within each PSU are selected.

Dwellings within each selected PSU, both urban and rural, are selected independently and systematically, and each dwelling has an equal probability of being selected. As in other surveys, the algorithm for the systematic selection of dwellings uses a random seed. Thus, regardless of who applies the algorithm under the same pre-established sequence[50], the same dwellings will be selected.

The procedure uses the Survey Select function of SAS software, which can be summarized by the following steps:

The period K is calculated as follows:

$$K = \frac{M'_{hi}}{m_{hi}} \qquad (13)$$

[50] The dwellings within each PSU are ordered according to the variable "dwelling sequence number" by which the route of enumerators of the cluster is established.

Where,

m_{hi}: Number of selected dwellings in cluster i of stratum h

M'_{hi}: Number of dwellings in cluster i in stratum h, according to the updating of the lists of dwellings

Then a random whole integer A is generated between one and period K.

Next, the period is successively added at the beginning to obtain different values that when rounded will generate the successive selections:

$$A, A+K, A+2K, A+3K, A, \ldots, A+(m_i-1)K \qquad (14)$$

The first selection is already a whole number, the second is the rounding of $A+K$, the third is the rounding of $A+2K$, and so on until the selection given by the rounding of $A+(m_i-1)K$.

When the number of dwellings of the selected PSU is updated and that the selection is made with equal probability within stratum h, the probability of selecting the dwelling of the selected PSU is defined by the following expression:

$$\pi_{hi/j} = \frac{m_{hi}}{M'_{hi}} \qquad (15)$$

Where,

$\pi_{hi/j}$: Probability of selecting the dwelling because it was selected from the cluster in stratum h

Consequently, the probability of selecting cluster i and dwelling j from stratum h is as follows:

$$\pi_{hij} = \frac{n_h \cdot M_{hi}}{\sum_{\forall i \in \mathcal{M}_h} M_{hi}} \cdot \frac{m_{hi}}{M'_{hi}} \qquad (16)$$

The extent of the difference between M_{hi} and M'_{hi} determines that, within a stratum, the dwellings among clusters have a different probability of selection. Therefore, the sampling design loses the property of being self-weighted.

However, the new PSUs of the ENE sampling design have approximately equal sizes. Therefore, they do not to undermine the use of sampling with probability proportional to size. It should be noted that the difference between M_{hi} and M'_{hi} will show the obsolescence of the frame over time.

On the other hand, the application of systematic sampling to the selection of PSUs, as opposed to simple random sampling, introduces an implicit stratification into the design, which improves

the distribution of the sample within each stratum, provides greater coverage, and increases the reliability of estimates.

In summary, the systematic selection of PSUs with probability proportional to size in terms of dwellings allows for control of the total sample size and therefore of costs. This makes it possible to assign workloads more equitably, to obtain better coverage within each stratum, and to provide information on the updating of the frame.

5.7 Rotation of sampling units

The ENE is a panel survey, thus the units of analysis are interviewed more than once and at various times in order to measure change in the indicators of interest over time. Therefore, in order to improve the precision of these estimates of change, the ENE uses a rotating sampling design[51] that guarantees a determined percentage of overlap or superposition between the samples from one period to another. This, in turn, minimizes the deterioration of the sample from non-response related to respondent fatigue caused by repeated interviews. The designed overlap also refreshes the sample.

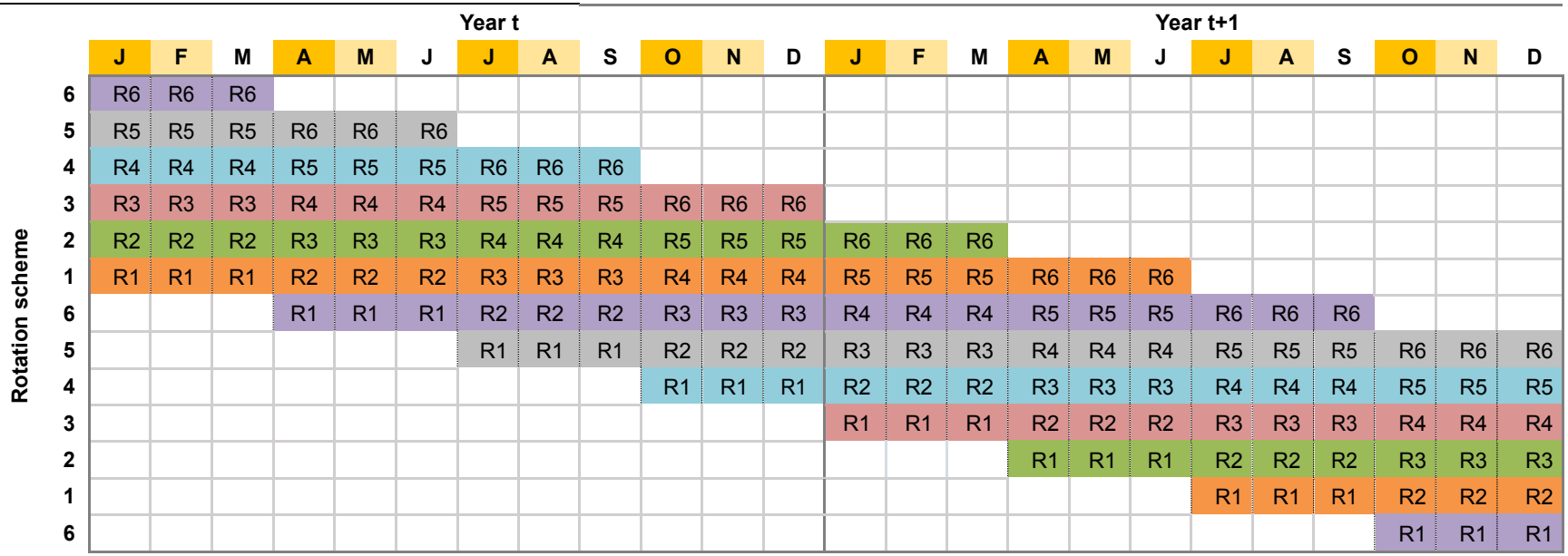
The total sample of the ENE is distributed into three subsamples[52]. During the months of January, April, July, and October of each year, the dwellings of subsample 1 are interviewed. Interviews are conducted in the months of February, May, August, and November for the dwellings of subsample 2, and in the months of March, June, September, and December for the dwellings of subsample 3.

The panel design includes six rounds for each dwelling. Therefore, the PSU sample of each subsample is divided into six rotations. In each moving quarter, a fixed proportion of the dwellings sampled (1/18) is rotated to incorporate a new set of dwellings (refreshment sample), thus rotating the entire sample of dwellings over a period of eighteen months.

[51] After a certain number of interviews, a rotating sampling design refreshes a fraction of the sample with a new set of sample units to update it.

[52] This procedure equally distributes the sample of PSUs from each ENE stratum among the three subsamples.

Figure 10 Rotation scheme of the sample of dwellings



Source: Produced by the National Statistics Institute

Figure 10 shows the rotation scheme of the sample of dwellings. The colors indicate the rotation panel to which each dwelling belongs: orange for panel 1 of each subsample, green for panel 2, pink for panel 3, blue for panel 4, gray for panel 5, and purple for panel 6. The numbers beside the letter "R" in each box indicate the number of periods that the dwellings associated with each of the six rotation panels have been part of the survey. As can be seen, a part of the sample is renewed every month[53]. For example, in April (subsample 1), the dwellings in the purple rotation panel are in their first round, and the dwellings in the gray rotation panel are in their sixth and final round. In July (subsample 1), the dwellings from the gray rotation panel are refreshed. That is, the group of dwellings that has completed six rounds in April leaves the sample, and a new group of dwellings is incorporated, as shown in the diagram. These processes are repeated for each subsample.

The refreshment sample is usually selected from the same primary sampling units; however, an update of PSUs is also included for the refreshment of geographical areas.

5.8 Incorporation of the sample from the 2017 sampling frame of dwellings (SFD)

The incorporation of the new sample from the 2017 SFD begins in January 2020. It will ensure that a fraction of the sample from the new framework can be incorporated into the moving quarter November 2019–January 2020 (NDJ 2019) at a rate of one-sixth of the quarterly sample[54] (approximately 9,000 dwellings).

At the same time, one-sixth of the sample from the previous sampling frame[55] (approximately 6,300 dwellings) is eliminated from the total sample, thus generating an increase in the sample. An analogous procedure will be applied for the subsamples of February, March, April, May, and June 2020, gradually increasing the sample size. Thus, in the moving quarter of April to June 2020 (AMJ 2020), the ENE sample will have a new sampling size[56] of approximately 54,000[57] dwellings, all of which are from the new 2017 SFD.

[53] It should be remembered that the total sample of the ENE consists of three subsamples. Thus, from one month to the next, one eighteenth of the total ENE sample is renewed.

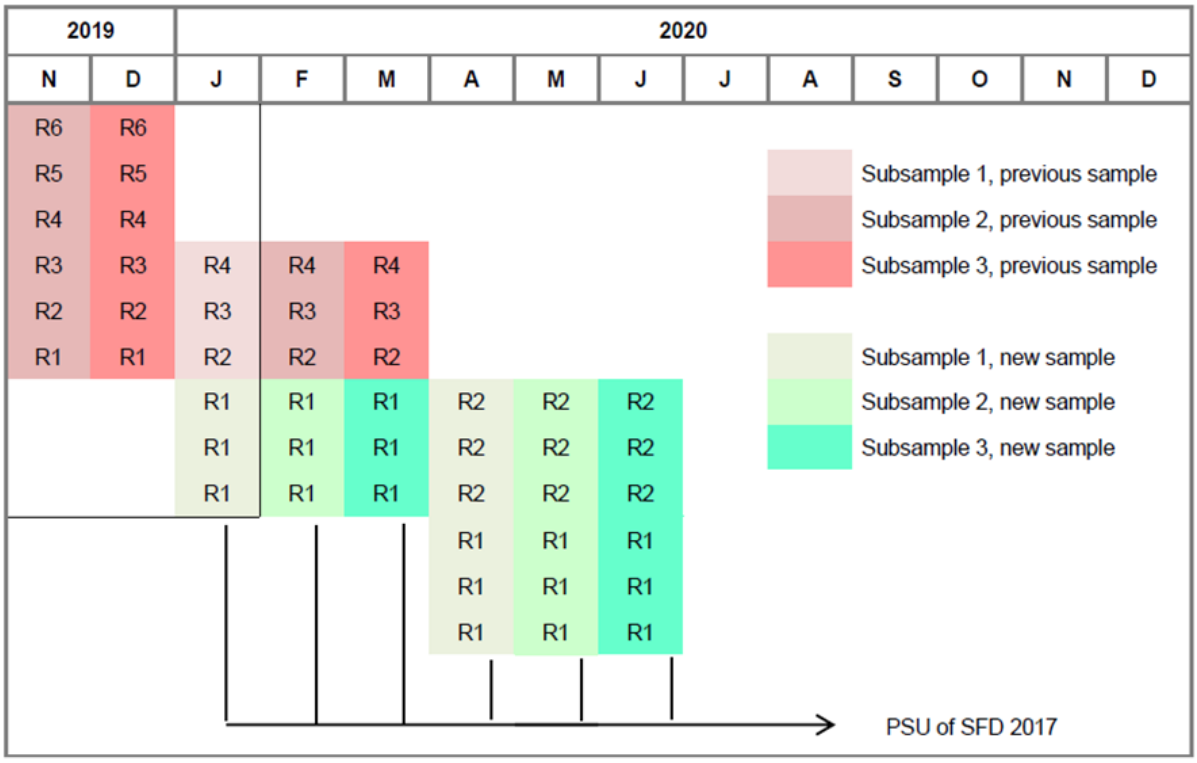
[54] This one-sixth is calculated with the new sample size of the ENE, which includes an oversample.

[55] This one-sixth is calculated with the sample size of the ENE sampling design valid until December 2019.

[56] The sample size of the previous design was 38,000 dwellings, and it did not include an oversample.

[57] This includes the oversample.

Figure 11 Incorporation of the 2017 sampling frame of dwellings (SFD)

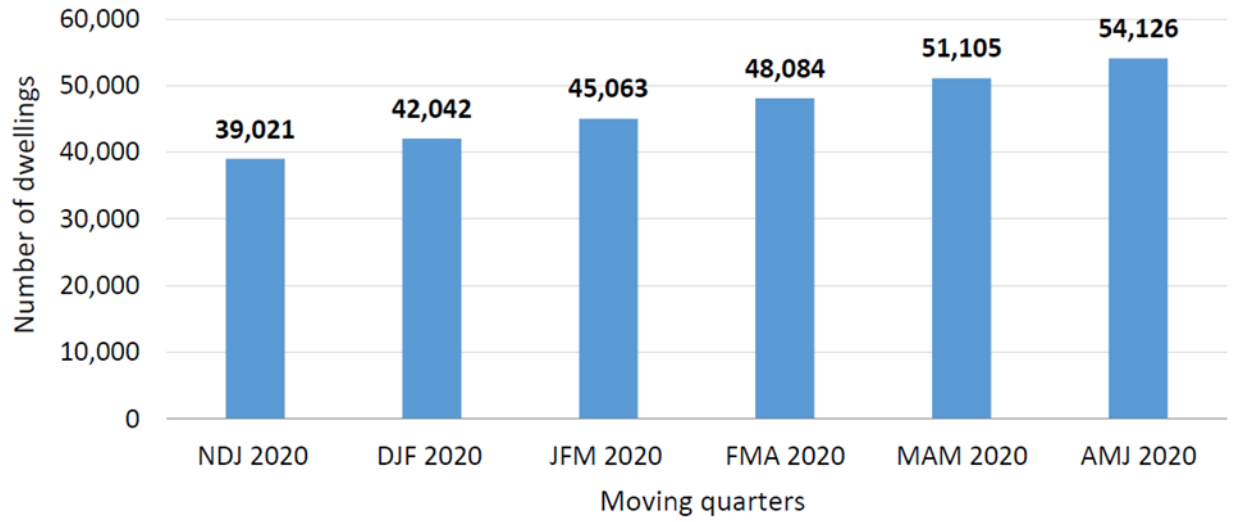


Source: Produced by the National Statistics Institute

This methodology gradually incorporates the new sample, thus controlling for the effects that changes in the sampling design might have on the estimates of the parameters of interest. However, it is important to note that, because the sample sizes will gradually increase during the first five months of 2020, the levels of precision set for the new design cannot be fully guaranteed until the sample size reaches the optimal levels of precision in June 2020[58] (see Graph 1).

[58] During the months of progressive incorporation of the sample (i.e., the transition stage), the level of estimation errors in the study will be evaluated. Some estimates may not reach the optimum level of errors to be official.

Graph 1 Gradual increase of theoretical sample during ENE transition period



Source: Produced by the National Statistics Institute

5.8.1 Incorporation of the new ENE sampling by increments

To ensure that the sample from the 2017 SFD framework is incorporated at a rate of one-sixth per month during the six months of transition, a number between one and six (representing the month in which it is incorporated into the PSU) was systematically assigned within each stratum of the new design after the PSUs were selected. Thus, approximately one-sixth of the 2017 SFD was incorporated in each of the six months. Each increment of incorporation could be distributed at the level of PSU because the new design has established that ten dwellings are to be surveyed per PSU, with the exception of the urban strata of the regions of Atacama, Aysén, and Magallanes, where eight dwellings are to be surveyed.

5.8.2 Selection of dwellings to be removed each month

Because of the great differences in PSU construction among the frames of the 2016 SFS, 2016 SFB, and 2017 SFD, the sample to be eliminated each month from the 2016 sampling frame of sections (SFS) and the 2016 sampling frame of blocks (SFB)[59] was selected from each sampling stratum of the previous sampling design. The sample to be eliminated was selected independently for each subsample and was at the level of dwelling.

[59] Until December 2019, the sample of the ENE was selected from a sampling frame made up of two lists of units. The first list, the sampling frame of blocks (SFB 2016), was updated from the lists of the 2016 Precensus. The 2016 SFB was used for the city (CD) stratum, which consists of cities and large urban centers. The second list, the sampling frame of sections (2016 SFS), was also partially updated from the lists of the 2016 Precensus. The 2016 SFS was used for the RAU (Rest of Urban Area) and rural strata.

During each month of the transition period, priority was given to dwellings from the previous sampling frame. They were selected for elimination from the sample according to the following criteria:

1. Dwellings located within a PSU selected in the sample from the 2017 SFD: each dwelling had to be located within the PSUs of the new framework. In the absence of cartographic automation, only intersections between sections and PSUs could be determined.

The dwellings identified in this first step were prioritized for elimination to reduce the probability of selecting a dwelling from a new sample that already belonged to the ENE sample selected from the previous sampling frame. This reduced the probability of unnecessarily repeating the collection of data.

2. Dwellings with the greatest number of rounds completed: The second order of priority was dwellings with the greatest number of rounds completed. These dwellings were eliminated until 1/6 of the sample was reached. Thus, dwellings that were previously visited three to five times were prioritized for elimination.

In some sampling strata, the number of dwellings from the previous sample located within a PSU selected from the 2017 SFD is greater than one-sixth of the previous sample. In these cases, the proportion was made more flexible so that dwellings from the first group of up to one-fifth of the dwellings within the previous sampling stratum could be eliminated.

6 Development of the expansion factor

Expansion factors can be defined as the number of units in the population that are represented by an element of the sample; the sum of these units equals the size of the population. Therefore, the entire sample represents the entire population (Lohr, 2000).

To generate the estimates for the variables of interest in the ENE, the expansion factor at level of person is used. However, for its calculation, the expansion factor at level of dwelling must also be calculated for all households in the dwelling and for all persons in the household and be calibrated to the population according to the defined calibration methodology. Because the sampling design of the ENE is two-stage, the calculation formula for expansion factors must include the selection probabilities of the first stage units (PSU) and second stage units (dwellings).

The components of the expansion factor are obtained as the product of the inverses of the selection probabilities of the first-stage units (PSUs) and the selection probabilities of the second-stage units (dwellings) selected from the clusters already sampled. Thus, the following figures must be considered:

a) The number of clusters selected within each stratum h and the number of dwellings in stratum h of the selected cluster i, according to the information in the sampling frame
b) The number of selected and existing dwellings in the cluster, according to the updated lists of dwellings of the selected cluster

To control both for sampling frames with outdated information observed in the survey process and for the non-response of units[60], the selection of expansion factors is adjusted in four ways:

1. Adjustment for unknown eligibility
2. Adjustment for ineligibility
3. Adjustment for non-response
4. Adjustment of population groups

With such adjustments, the weights of the units of unknown eligibility can be distributed between the eligible and non-eligible units. In addition, the weights of the non-eligible units can be discarded because, by definition, these units do not contain the target population, making inferences about them incorrect. For various reasons, even though some units are eligible, no response can be

[60] With improvements in collection, detailed information on the outcome (i.e., the disposition status) of each selected sample unit can be gathered. The visitation codes indicate the eligibility of each sampling unit, and they can be used to classify dwellings into three groups: eligible, unknown eligibility, and ineligible. This information is then used to make the corresponding adjustments to the expansion factor (see Appendix 3, Visitation codes of the ENE).

obtained from the informants. Therefore, their weight must be distributed among other units that represent them in the sample.

The following sections describe the methodology for each of the above-mentioned adjustments,

6.1 Expansion factor for the selection of clusters

In the two-stage sampling design of the ENE, the first stage is the selection of the primary sampling units whose probability of selection is proportional to the size of the number of occupied private dwellings in the PSU according to the Population and Housing Census of 2017 or according to the most recent update of the frame. Thus, the selection weight of each PSU is expressed as the inverse of the selection probability, as follows:

$$w_{hi} = (\pi_{hi})^{-1} = \frac{M_h}{n_h \cdot M_{hi}} \qquad (17)$$

Where,

w_{hi}: Expansion factor for cluster i in stratum h

π_{hi}: Probability of selecting cluster i in stratum h

M_h: Number of private dwellings in stratum h according to the 2017 Census or to the year in which the sampling frame was updated

M_{hi}: Number of private dwellings in cluster i of stratum h according to the 2017 Census or to year in which the sampling frame was updated

n_h: Number of selected clusters in the sample of stratum h

This weighting is interpreted as “the total number of units of its type that it represents in the population”.

6.2 Expansion factor for the selection of dwellings

The selection weight of dwellings w_{hij} indicates the number of dwellings in the population that represent the selected dwelling in the sample. This selection weight is calculated as the inverse of the probability of selection of each dwelling.

The total probability of selecting dwelling j in stratum h is determined by the product of the probability of selecting cluster i in stratum h and the probability of selecting dwelling j. This depends upon having selected cluster i (i.e., the selection of dwelling j in the previously selected cluster i). This can be expressed as:

$$\pi_{hij} = \pi_{hi} \cdot \pi_{hi/j}, where\ \pi_{hi/j} = \frac{m_{hi}}{M'_{hi}} \quad (18)$$

Finally, the housing selection weight can be obtained as follows:

$$w_{hij} = \left(\pi_{hij}\right)^{-1} = \frac{1}{\pi_{hi}} \cdot \frac{1}{\pi_{hi/j}} = \frac{M_h}{n_h \cdot M_{hi}} \cdot \frac{M'_{hi}}{m_{hi}} \quad (19)$$

Where,

w_{hij}: Expansion factor of selection of dwelling j in cluster i in stratum h

π_{hij}: Probability of selecting dwelling j in cluster i in stratum h

π_{hi}: Probability of selecting dwelling j because it was selected in cluster i in stratum h

$\pi_{hi/j}$: Number of private dwellings in stratum h according to the 2017 Census or to the year in which the sampling frame was updated

M_h: Number of selected clusters in the sample of stratum h

n_h: Number of private dwellings in the cluster i in the stratum h according to the 2017 Census or to year in which the sampling frame was updated

M_{hi}: Number of selected dwellings in cluster i in stratum h

M'_{hi}: Number of occupied private dwellings in cluster i in stratum h, according to updated lists of dwellings

6.2.1 Eligibility adjustment

Eligibility adjustments are intended to correct for the possible inclusion in the sample of dwellings classified as ineligible and as of unknown eligibility.

6.2.1.1 Eligibility adjustment

In principle, all dwellings have a known eligibility status, whether eligible or ineligible. Moreover, the dwellings may be classified as occupied or unoccupied because their addresses are based on census information or on the updated lists of dwellings. However, enumerators in the field may not be able to confirm the status of some dwellings, because they are inaccessible (for climate, security, or other reasons) or because the sample was simply not collected in a timely manner. In these cases, the status of the dwelling is classified as of “unknown eligibility”. To mitigate this situation, the expansion factors of the dwelling can be adjusted to redistribute the weights of these units proportionally between eligible and non-eligible units.

Thus, the adjustment weight is defined by the following expansion:

$$\hat{R}_{h,know} = \frac{\sum_{i\epsilon\Omega_h}\sum_{j\epsilon\Theta_i} w_{hij}}{\sum_{i\epsilon\Omega_h}\sum_{j\epsilon\Theta_{i,know}} w_{hij}} \quad (20)$$

Where,

Θ_i: Set of dwellings j selected in the cluster i

$\Theta_{i,know}$: Set of dwellings j selected in cluster i and classified as eligible or not eligible (known eligibility)

Ω_h: Set of clusters i selected in stratum h

w_{hij}: Expansion factor of selection of the dwelling j in cluster i in stratum h

Thus, $\hat{R}_{h,know}$ is calculated within each sampling stratum by determining the ratio between the estimate of total number of dwellings (eligible, ineligible, and of unknown eligibility) and the estimate of the total number of dwellings of known eligibility (eligible and ineligible).

The expansion factor for dwelling selection adjusted for unknown eligibility (w'_{hij}) is calculated as follows:

$$w'_{hij} = \hat{R}_{h,know} \cdot w_{hij} \quad (21)$$

6.2.1.2 Adjustment for no response

The target population of the ENE consists of persons aged 15 years old or over who habitually reside in occupied private dwellings. However, despite efforts to keep the sampling frame as up-to-date as possible, errors in updating or changes in the use of the dwelling may result in sample dwellings that do not contain the target population.

This adjustment assigns a null value (missing value) in the dwelling selection weight corrected for unknown eligibility (w'_{hij}) to dwellings classified as ineligible.

This adjustment can be described as follows:

$$R_{hij,elig} = \begin{cases} 1, & if\ dwelling\ j\ of\ cluster\ i\ of\ stratum\ h\ is\ eligible \\ ., & if\ dwelling\ j\ of\ cluster\ i\ of\ stratum\ h\ is\ ineligible \end{cases} \quad (22)$$

The expansion factor of dwellings, adjusted for unknown eligibility and for ineligibility, is as follows:

$$w''_{hij} = w'_{hij} \cdot R_{hij,elig} \quad (23)$$

6.2.2 Adjustment for non-response

This adjustment includes only eligible dwellings (i.e., occupied private dwellings), and it considers the probability of non-response associated with dwellings where the informants refuse to respond to the survey, where it was not possible to establish contact, or where the interview could not be conducted for other reasons[61].

The non-response adjustment seeks to correct the weights so that the survey (based on the units that do respond) is representative of the entire population. As a result, the weights of the units that responded to the survey are increased to compensate for the weight of the non-responding units.

Weight $\hat{R}_{h,R}$, which is defined as the ratio between the estimate of the number of eligible dwellings in the population, is calculated from the expansion factor defined in section 6.2.1.2 (corrected for ineligibility) and from the estimate of the total number of dwellings from which a response from the informants has been obtained. Thus, the adjustment for non-response is determined by the following expression:

$$\hat{R}_{h,R} = \begin{cases} \dfrac{\sum_{i\in\Omega_h}\sum_{j\epsilon\Theta_{i,elig}} W''_{hij}}{\sum_{i\in\Omega_h}\sum_{j\epsilon\Theta_{i,elig,R}} W''_{hij}}, & if\ dwelling\ j\ is\ eligible\ and\ responds \\ . \ , & if\ dwelling\ j\ is\ eligible\ and\ does\ not\ respond \end{cases}$$

Where,

$\Theta_{i,elig}$: Set of dwellings j selected in cluster i classified as eligible

$\Theta_{i,elig,R}$: Set of dwellings j in the cluster i for which results were attained. The set of dwellings is classified as eligible and as having responded to the survey.

Ω_h: Set of clusters i selected in stratum h

W''_{hij}: Expansion factor of selection of the dwelling j within cluster i in stratum h, adjusted for ineligibility

Next, the selection weight of the dwelling adjusted for non-response $\left(w_{hij}^{NR}\right)$ is calculated as follows:

$$w_{hij}^{NR} = \hat{R}_{h,R} \cdot w''_{hij} \qquad (25)$$

This expansion factor is applied to persons k in dwelling j in cluster i of stratum h, which is written as w_{hijk}^{NR}.

[61] This adjustment is in accord with the final disposition codes used in collection since January 2020.

6.2.3 Adjustment of population groups (calibration)

The calculated weights at the level of dwelling are used for the respective population groups. However, the units of analysis of the ENE are the persons aged 15 years or over who reside in occupied private dwellings. Therefore, with population projections based on the population and housing censuses, these groups are used to correct possible imbalances in variables (e.g., age and sex) of the population totals.

If the expansion factor of dwelling selection adjusted for non-response is assigned to each person within the dwelling, an estimate is obtained of the total number of persons. However, this value differs from estimates from other sources, such as population projections generated from demographic models. Therefore, in order to adjust the survey population to the population groups calculated by the models, the survey population is calibrated by incorporating auxiliary information. This results in nearly unbiased estimates with smaller variance than the estimates obtained by the Horvitz and Thompson estimator.

The calibration methods are based on iterative algorithms that successively adjust the expansion factors to reach population totals of the auxiliary variables while minimizing the distance between the original and the calibrated factors.

The idea of keeping the calibrated weights near one another is to be able to associate any good estimation property of the initial weights by approximation. For example, if the initial weights produce unbiased estimators, the new calibrated weights will produce nearly unbiased estimators.

Calibrated estimators have important properties. For example, if the population totals of the auxiliary (marginal) variables have been published, calibration ensures that the survey estimates are consistent with the estimates that have already been published.

Another important property is their simplicity (i.e., the calibration estimators are linear). This means that each element of the sample is assigned a single weight for the estimation of any of the variables of interest in the survey. The estimates for totals, averages, and ratios, among other parameters, can easily be calculated with standard statistical software.

Another property of calibrated estimators is their flexibility in incorporating auxiliary information, including continuous or discrete variables or both at the same time. In addition, calibration estimators somewhat counteract the bias created by non-response and by insufficient coverage of subpopulations that result from the imperfection of the sampling frame (IBGE, 2004).

The multiplicative method, or raking method, ensures positive expansion factors because the distance function is not defined for negative values (Vanderhoeft, 2001). This method has been chosen for the calibration of the ENE.

The distance function $G(w^{cal}, w^{NR})$ is determined as follows:

$$\boldsymbol{G(w^{cal}, w^{NR}) = w^{cal}\log(w^{cal}/w^{NR}) - w^{cal} + w^{NR}} \qquad (26)$$

Where,

w^{cal}: Expansion factor calibrated at the corresponding auxiliary level

w^{NR}: Expansion factor corrected for non-response

This allows the creation of an expansion factor calibrated at the corresponding auxiliary level, which is obtained by the following expression:

$$\boldsymbol{w^{cal} = w^{NR} \cdot F(w^{cal}, w^{NR})} \qquad (27)$$

Where,

$F(w^{cal}, w^{NR})$: Inverse function of $\partial G(w^{cal}, w^{NR})/\partial w^{cal}$ also denoted by $g^{-1}(w^{cal}, w^{NR})$

The defined margins[62] are as follows:

- At the national level: by sex and by thirteen age groups [0-14], [15-19], [20-24], [25-29], [30-34], [35-39], [40-44], [45-49], [50-54], [55-59], [60-64], [65-69] and [70+]
- At the national level: estimation of the foreign population[63]. This auxiliary variable has been included since the FMA 2013 moving quarter. The construction of the tools for calibration using this auxiliary variable is detailed in Annex 8.
- At the regional level: total population by sex
- At the regional level: working-age population (WAP) by sex
- At regional level: total population of the urban area by sex
- Metropolitan Region: total population for six age groups by sex, distributed as follows: [15-24], [25-34], [35-44], [45-54], [55-64] and [65+]

[62] The justification for incorporating population information from various age groups and information on the total foreign population can be found in INE, 2019d.

[63] For more information, see the working document "*Fundamentos de la nueva metodología de calibración de los factores de expansión de la Encuesta Nacional de Empleo*" (INE, 2019d).

- Regions (excluding Metropolitan Region): total population for three age groups: [15-34], [35-54] and [55+].

For the calibration of the weights of the ENE, the population projections based on the 2017 Census are used. The weights are adjusted each month according to the methodology described in Annex 5.

In addition to the previously defined auxiliary variables, there are auxiliary variables for WAP at the level of province and of large conurbations.

The following criteria were used to determine the provinces to be calibrated:

- The coefficient of variation of the working-age population (WAP) of the province, calculated with the base selection weight, must be less than 10% in all 117 samples between the quarters of JFM 2010 and SON 2019.
- The province where the commune of the regional capital is located must be included.
- Provinces that had a WAP of greater than 100,000 inhabitants in October 2019 must be included.

After applying these criteria, some additional points were considered:

- The province of Diguillín has been included in the auxiliary variable since only the ASO 2018 quarter for three reasons:
    - It is the province where the regional capital is located.
    - The Ñuble region was created in September 2018.
    - The province did not meet the minimum coefficient of variation (<10%) prior to 2018.
- Although the province of Arica meets the three criteria, it is not included in the auxiliary variable, because its complement in the region (the province of Parinacota) has a very small population. Thus, since 2016 Arica has not been included in the sample.

To see the list of the auxiliary variables from provinces, see Annex 6.

On the other hand, four large conurbations have been included:

- La Serena-Coquimbo has been included since the JFM 2010 quarter.
- Greater Valparaíso has been included since the JFM 2010 quarter.
- Greater Santiago has been included since the JFM 2010 quarter.

- Greater Concepción has been included since the JFM 2010 quarter.

Initially, the conurbation of Antofagasta was included, but it was later decided to exclude this calibration because the conurbation consists of only one commune, which makes up 95% of the population of the province of Antofagasta. This creates the risk of the remaining communes of the province having a small presence in the sample, thus resulting in the tendency of samples at the level of province and commune to coincide.

7 Estimators

Once the characteristics the sampling design of the ENE have been defined, the formulas for their calculation can be determined.

The nomenclature used for the estimators is the following:

w_{hijk}^{cal}: Weight for person k adjusted for population projections associated with PSU i in stratum h

P_{hijk}: One, for all persons k of dwelling j of PSU i in stratum h

P_{hijk_c}: Person k of dwelling j of PSU i in the stratum h, who has characteristic C

That is, $$P_{hijk_c} = \begin{cases} 1, if\ the\ person\ has\ characteristic\ C \\ 0, for\ any\ other\ case \end{cases}$$

Y_{hijk}: Value of variable Y measured in person k of dwelling j of PSU i in stratum h

Y_{hijk_c}: Value of variable Y measured in person k of dwelling j of PSU i in stratum h, who has characteristic C

That is, $$Y_{hijk_c} = \begin{cases} Y_{hijk}, if the\ person\ has\ characteristic\ C \\ 0, \quad for\ any\ other\ case \end{cases}$$

This can also be expressed in the following manner:

$$Y_{hijk_c} = Y_{hijk} \cdot P_{hijk_c} = \begin{cases} Y_{hijk} \cdot 1 = Y_{hijk}, if\ the\ person\ has\ characteristic\ C \\ Y_{hijk} \cdot 0 = 0, for\ any\ other\ case \end{cases}$$

Note: remember that $\widehat{P}_h = P_h$. By construction, that is, the estimate of the number of persons within each stratum and level coincides with the population projection.

7.1 Estimators of the total for variable Y

The ENE generates estimates for the following totals:

- Population of the labour force
- Employed persons
- Informal employed persons

- Unemployed persons
- Economically inactive persons

The following expressions represent the generalized formula used to obtain estimates of the total of a variable according to the required level of estimation.

Total estimate of variable Y in stratum h

$$\hat{Y}_h = \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk} \qquad (28)$$

Total estimate of variable Y at the required level of estimation

$$\hat{Y} = \sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk} = \sum_h \hat{Y}_h \qquad (29)$$

The estimate of the total $\hat{Y}$ at the required level of estimation is obtained from the sum of the totals of each of the strata that make up the level of estimation.

Total estimate of variable Y in persons with characteristic C in stratum h

$$\hat{Y}_{h_c} = \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk_c} \qquad (30)$$

Total estimate of variable Y of persons with characteristic C at the required level of estimation

$$\hat{Y}_c = \sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk_c} = \sum_h \hat{Y}_{h_c} \qquad (31)$$

7.2 Ratio estimators

A mathematical ratio is a relationship between two magnitudes that are comparable to each other. It is the result obtained when one of the magnitudes or quantities is divided by another. Therefore, it can be expressed as fractions or as decimal numbers. Both the numerator and the denominator can be quantitative or qualitative variables.

The ENE generates ratio estimates to obtain the following indicators:

- Participation rate
- Employment rate
- Unemployment rate
- Unemployment rate with available job starters
- Combined rate of unemployment and involuntary part-time work
- Combined rate of unemployment and potential labour force
- Labour force pressure rate
- Informal employment rate
- Employment rate in the informal sector

The following expression represents the generalized formula used to obtain ratio estimates between two variables according to the required level of estimation.

Ratio estimator between variables X and Y

The ratio between variables X and Y can be expressed in the following manner:

$$\hat{R}_{h_y/x} = \frac{\hat{Y}_h}{\hat{X}_h} = \frac{\sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk}}{\sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot X_{hijk}} \quad (32)$$

The ratio estimator between variables X and Y are equal to the quotient obtained from dividing the total of these variables.

Similarly, the ratio estimator, at the required estimation level, can be obtained as a combined ratio.

$$\hat{R}_{y/x} = \frac{\hat{Y}_h}{\hat{X}_h} = \frac{\sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot Y_{hijk}}{\sum_h \sum_i \sum_j \sum_k w_{hijk}^{cal} \cdot X_{hijk}} \quad (33)$$

8 Precision measurements

Among other applications, the results of surveys are currently used in public policy decision-making. The analytical objectives go far beyond simply presenting results concerning the estimation of a parameter. It is therefore critical that the estimates include measurements of the precision of the estimates.

8.1 Estimator variance

One of the key measurements of precision in sample surveys is sampling variance, an indicator of the variability that occurs when a sample is chosen. Sampling variation is an alternative to listing the entire population. Although there are other precision measurements, they are algebraically related. Thus, it is possible to deduce the expression of any one of them from the others simply by applying mathematical operations.

Exact methods of the estimation of variance are the best way to estimate variance. However, because the new sampling design of the ENE, like its predecessor, uses a complex design, it is impossible to apply the calculation of variance by exact methods.

Consequently, the estimation of the variance of the main estimates of the ENE is obtained from the ultimate cluster method. This method was chosen because the greatest contribution to the variance of an estimator in a multi-stage design originates from the Primary Sampling Units (PSU). The ultimate cluster method simplifies a complex multi-stage design into a special case of cluster sampling without sub-sampling (i.e., into one stage). Ultimate cluster refers to the set of last-stage sample units belonging to a primary sampling unit, regardless of the number of steps within the first stage of selection.

This method approximates the variance of the estimators while also considering the stratification associated with the sampling design.

In the estimation of variance in ENE estimators, expansion factors have been calibrated using the Raking method (see section 6.2.3). Thus, according to Deville & Särndal (1992), the estimation of variance is obtained using approximations based on the generalized regression estimator (GREG)[64], which uses the chi-square distance function.

The GREG is part of a family of calibration estimators in which the Raking method is found. As shown in Deville & Särndal (1992), under conditions of regularity, they are asymptotically equivalent estimates. Therefore, the estimation of the GREG variance can be used. The estimation of variance through GREG has the benefit of being the least complex expression of the

[64] For more information, see Särndal, Swensson, & Wretman (1991), p. 235.

above-mentioned family of measurements. Although efforts have been made to derive expressions for the other members of the family, they have proven to be very difficult. Thus, in practice, the estimate of variance is approximated by the expression associated with GREG (Deville & Särndal, 1992).

The formulas for estimation of variance through GREG have been implemented in the survey package of R software. Thus, the calculation of the ENE estimators has been operationalized in this software, where the characteristics of complex sampling design and the generation of the weights calibrated by the Raking[65] method are required.

8.2 Absolute and relative errors

As mentioned, a sample survey is subject to sampling errors that arise because only a representative part of the population is measured, not the entire population. However, these errors associated with the variables of interest can be measured by different methods (e.g., variance and standard deviation) or by some of their transformations (e.g., absolute error, coefficient of variation, and relative error). To determine the statistical precision of the estimates, the ENE mainly uses the last three measurements of error.

8.2.1 Absolute error and confidence interval

For the calculation of a 95% confidence interval for the estimation of variable Y, the absolute error (i.e., the estimation error expressed in the same units as the variable under analysis[66]) is used.

The maximum expected difference between the estimated value and the true population parameter can be found by using the absolute error. This value can be made at different levels of confidence, the most common being 95%. From this error, the confidence interval of the variable of interest Y is constructed.

The absolute error is obtained from the product of the standard deviation $\left(S(\hat{y})\right)$ of the estimate and the critical value associated with the percentile of the t-distribution with v degrees of freedom ($t_{1-\frac{a}{2},v}$), where the fist concept refers to the desired level of confidence (95%) and the second concept refers to the degrees of freedom with which the critical value is calculated[67].

[65] The development of the variance approach for the Raking method, which is based on the variance of the GREG estimator, is discussed in Särndal, Swensson, & Wretman (1991).

[66] When the variable Y is a percentage or rate, the absolute error is expressed in percentage points.

[67] Degrees of freedom refer to the difference between the number of non-empty clusters (with the presence of the variable of interest) and the number of non-empty strata.

This can be expressed in the following manner:

$$E_{abs}(\hat{y}) = t_{1-\frac{a}{2},v} \cdot S(\hat{y}) \quad (34)$$

Alternatively, in probabilistic terms, it can be expressed as follows:

$$Pr\left(\left|\hat{Y} - Y\right| \leq E_{abs}(\hat{y})\right) = 1 - \alpha \quad (35)$$

Where $\hat{Y}$ and Y are the estimated value and true value in the population, respectively.

The confidence interval, on the other hand, is the estimated interval that includes the average of the estimates of all possible samples with a 95% probability. This interval has a lower and an upper limit within which the actual or population value is expected to be, at a confidence level of 95%.

The confidence interval is constructed through the absolute error, where the minimum value of the population parameter is equal to the estimated parameter minus the absolute error and where the maximum value is equal to the estimated parameter plus the absolute error. The confidence interval can be expressed in the following manner:

$$IC(\hat{y})_{95\%} = \left(\hat{y} - E_{abs}(\hat{y}); \hat{y} + E_{abs}(\hat{y})\right) = \left(\hat{y} - t_{1-\frac{a}{2},v} \cdot S(\hat{y}); \hat{y} + t_{1-\frac{a}{2},v} \cdot S(\hat{y})\right) \quad (36)$$

8.2.2 Relative error and coefficient of variation

In addition, the relative error and the coefficient of variation account for relative sampling errors.

The relative error expresses the magnitude of the sampling error as a percentage, reflecting the maximum percentage variation that can be expected between the estimated value and the population value with a probability of $(1 - \alpha) \cdot 100\%$. It can be obtained as follows:

$$E_{rel}(\hat{y}) = t_{1-\frac{a}{2},v} \cdot \frac{\hat{\sigma}_{\hat{y}}}{\hat{y}} = t_{1-\frac{a}{2},v} \cdot CV_{\hat{y}} \quad (37)$$

Where $\frac{\hat{\sigma}_{\hat{y}}}{\hat{y}}$ represents the coefficient of variation $CV_{\hat{y}}$ associated with the estimate of variable $\hat{y}$.

The great usefulness of both the relative error and the coefficient of variation is that they can be expressed as percentages. This distinguishes them from other measurements of errors, enabling comparisons to be made between different estimates and avoiding differences in units of measurement.

8.3 Representativeness of estimates

Based on *Lineamientos para el Uso de Medidas de Precisión en Encuestas de Hogares* (INE, 2020), the acceptability of an estimate depends upon the approach outlined in the flowchart on whether to publish estimates, including the criteria described below.

8.3.1 Aspects of sample size

Sample size Sample size is the total number of units of analysis included to obtain the estimates. The threshold for acceptance is a sample size greater than or equal to 60 units.

Degrees of freedom Degrees of freedom (df) is defined as the number of independent observations that are free to vary (i.e., the number of individuals in the data) minus the number of estimated parameters (i.e., number of relationships imposed on the data). In a complex sampling design, the df is obtained as the difference between the number of non-empty clusters (i.e., clusters that include the variable of interest) and the number of non-empty strata. According to Heeringa, West, & Berglund (2010), the df also accounts for the precision with which the parameters are being estimated across the sample. Thus, the greater the degrees of freedom, the more accurate the estimates will be, and vice versa. The threshold for acceptance for the evaluated estimate is nine or more degrees of freedom.

If the evaluated level of disaggregation does not meet the criterion of sample size or degrees of freedom, then the estimate for that level should be deleted.

8.3.2 Aspects of estimate precision

If the estimator is a ratio, it will be evaluated by means of the standard error (SE). The thresholds are determined according to the following function**:**

$$Maximum\ tolerable\ SE = \pi \begin{cases} \sqrt[3]{p^2}/9 & if\ p \leq 0.5 \\ \sqrt[3]{(1-p)^2}/9 & if\ p > 0.5 \end{cases}$$

If the evaluated level of disaggregation meets the criteria of sample size and degrees of freedom but does not meet the above criteria, the estimate is published with a warning that it is unreliable.

If the estimator is not proportionate, it will be evaluated by the coefficient of variation as follows:

$$Maximum\ tolerable\ CV = \begin{cases} CV \leq 0.15 \rightarrow acceptable \\ 0.15 < CV \leq 0.30 \rightarrow unreliable\ estimate \end{cases}$$

If the evaluated level of disaggregation meets the criteria for sample size and degrees of freedom but has a coefficient of variation greater than 0.30, the estimate for that level must be eliminated.

Subsequently, the publication of the tabulation depends on the number of estimates that have been deleted or have been determined to be unreliable. It is recommended that if these estimates add up to more than 50% of those forming the tabulation, the tabulation should not be published.

9 Scope of reporting of results

9.1 Dissemination of results

The results and material produced from the survey are available to the public and users at the institutional website (www.ine.cl/), the ENE Data Bank (Banco de Datos ENE), and the INE.STAT platform (https://stat.ine.cl).

Results are published at 9:00 am on the last working day of each month, with the exception of the INE.STAT platform, which is updated forty-eight working hours after the publication of the figures.

The material published by the survey is the following:

- **National Employment Bulletin:** is a document that details the main indicators of the labor market in Chile, describing the variations of the situation inside and outside the labour force in different categories, such as occupational group, branch of economic activity, and occupational category. The bulletin includes underutilization rates, seasonally adjusted rates, and labor informality rates. The bulletin also compares ENE data with data from administrative records.
- **National Employment Bulletin for the Metropolitan Region and greater Santiago:** is a document that details the main indicators of the labor market in the Metropolitan region and greater Santiago. It describes the variations of the situation inside and outside the labour force in different categories, such as occupational group, branch of economic activity, and occupational category.
- **Statistical note on employment:** is a file that contains tables with the main information of the ENE (levels and rates of unemployment), in addition to response rates, absolute error, coefficients of variation, and confidence intervals, as appropriate.
- **Statistical note on labour informality:** is a file that contains tables with the main information on labour informality (levels of informal employed persons and rates of informal employment), in addition to response rates, absolute error, coefficients of variation, and confidence intervals, as appropriate.
- **Bulletin of labor informality:** is a quarterly document that details the main indicators of informal employment in Chile, describing the variations in formal and informal employed persons of various categories, such as occupational group, branch of economic activity, and occupational category. The document includes informal employment rates and the employment rate in the informal sector.
- **Annexes of the ENE:** is a file that contains tables with the main figures of the ENE, as well as their absolute and relative annual variation and impacts.
- **Annual Supplement:** is a document that summarizes the average variations of the main classifications and occupational categories during the year that ended last month, on both a national and regional level. The document highlights the main variations according to

impact in the main variables such as sex, economic branch, occupational category, and main rates).

- **Technical supplement:** is a document that summarizes technical and/or operational aspects of ENE for a particular variation or situation. Their publication is occasional, meeting the objective of disseminating information in an accurate, simple, and timely manner.
- **Database:** is an anonymous database with quarterly information from the ENE from 2010 to the present.
- **Tabulations:** is a data series on the main national and regional information from the survey, from 2010 to the present. Tabulations are updated every quarter.
- **Methodology of the National Employment Survey:** is a document that summarizes the main methodological aspects of the survey. Specifically, it summarizes the conceptual framework for the measurement of employment and the statistical design of the survey, which was updated in 2020.

9.2 Data accessibility and legal framework

According to institutional guidelines, the information produced by the ENE is subject to INE's commitments to maintain the confidentiality and privacy of data protected by statistical secrecy. These guidelines can be summarized in the following legal framework:

- Law 19.628 concerns the protection of private life and personal data.
- Law 20.285 regulates the principle of transparency of public functions and the right of access to information of the bodies of State Administration.
- Article 29 of Law 17.374 states that the National Statistics Institute, fiscal and semi-fiscal bodies, State enterprises, and each of their respective officials, may not disclose facts that refer to specific persons or entities of which they have become aware of in the course of their activities.
- The strict maintenance of these facts constitutes "Statistical Secrecy". Its violation by any person subject to this obligation will result in an offense and punishment in accordance with article 247 of the Criminal Code. Custodial sentencing must be applied in all cases.
- Article 30 of Law 17.374 states that "statistical data may not be published or disseminated with express reference to the persons or entities to which they directly or indirectly refer, if prohibited by the person or persons concerned" (Extracts from Organic Law 17.374 creating the National Statistics Institute, published on 10 December 1970).
- Law 17.336 on Intellectual Property, Article 88, first paragraph states, "The State, municipalities, official corporations, semi-fiscal or autonomous institutions, and other legal persons of the State shall be owners of copyright of works produced by their officials in the course of their duties".

10 Annexes

Annex 1 Number of PSUs, dwellings, and working-age population (WAP) by geographical stratum of the ENE

Region	Area	Code of Geographical stratum	Geographical stratum	PSU	Dwellings	WAP
Arica y Parinacota	Urban	1500001	Región de Arica y Parinacota	331	65,926	161,895
Arica y Parinacota	Rural	1500012	Región de Arica y Parinacota	74	6,676	14,963
Tarapacá	Urban	110011	Conurbación Iquique - Alto Hospicio	477	96,171	224,140
Tarapacá	Urban	120011	Prov. Del Tamarugal	27	5,370	11,128
Tarapacá	Rural	110012	Conurbación Iquique - Alto Hospicio	13	1,027	4,430
Tarapacá	Rural	120012	Prov. Del Tamarugal	53	4,905	13,657
Antofagasta	Urban	210001	Prov. Antofagasta (Conurbación Antofagasta + Mejillones + Tal Tal)	581	116,250	296,879
Antofagasta	Urban	220011	Prov. El Loa (San Pedro de Atacama y Conurbación Calama)	263	54,000	126,604
Antofagasta	Urban	230011	Prov. Tocopilla	51	10,638	22,637
Antofagasta	Rural	210002	Prov. Antofagasta (Conurbación Antofagasta + Mejillones + Sierra Gorda + Tal Tal)	20	1,590	19,916
Antofagasta	Rural	220012	R.R. (Prov. Tocopilla y Prov. El Loa)	50	4,439	13,636
Atacama	Urban	310011	Conurbación Copiapó	279	56,292	123,524
Atacama	Urban	310021	Prov. Chañaral + R.P. Copiapó (Caldera)	79	15,842	30,551
Atacama	Urban	320011	Prov. Huasco (Conurbación Vallenar, Freirina, Huasco)	104	20,800	45,821
Atacama	Rural	310012	Conurbación Copiapó	33	2,876	5,786
Atacama	Rural	310022	R.R Atacama (Prov. Chañaral y R.P. Copiapó (Caldera))	30	2,485	3,030
Atacama	Rural	330012	Prov. Huasco (Vallenar, Alto del Carmen, Freirina, Huasco)	82	7,259	11,834
Coquimbo	Urban	410011	Conurbación La Serena – Coquimbo	734	148,084	324,319
Coquimbo	Urban	410021	R.P. Elqui	53	10,134	21,697
Coquimbo	Urban	420011	Prov. Choapa (Illapel, Canela, Los Vilos, Salamanca)	104	21,254	44,484
Coquimbo	Urban	430011	Conurbación Ovalle	150	30,145	67,570
Coquimbo	Urban	430021	R.P. Limarí	51	10,124	20,676
Coquimbo	Rural	410012	Conurbación La Serena - Coquimbo	131	11,955	26,519
Coquimbo	Rural	410022	R.P. Elqui	87	7,896	15,533
Coquimbo	Rural	420012	Prov. Choapa (Illapel, Canela, Los Vilos, Salamanca)	150	13,567	27,527
Coquimbo	Rural	430012	Conurbación Ovalle	112	10,042	18,816
Coquimbo	Rural	430022	R.P. Limarí	166	14,859	26,114
Valparaíso	Urban	510011	Conurbación Valparaíso (1) (Valparaíso)	549	110,459	243,343
Valparaíso	Urban	510021	Conurbación Valparaíso (2) (Viña del Mar, Concón)	728	148,135	310,830
Valparaíso	Urban	510031	R.P. Valparaíso (Casablanca)	30	6,153	14,243
Valparaíso	Urban	510041	R.P. Valparaíso (Puchuncaví, Quintero)	79	16,383	34,061
Valparaíso	Urban	530011	Prov. Los Andes (Conurbación Los Andes + Rinconada y San Esteban)	158	32,555	71,553
Valparaíso	Urban	540011	Prov. Petorca	99	19,473	41,739
Valparaíso	Urban	550011	Prov. Quillota (Conurbación Quillota - Calera - Hijuelas - La Cruz, Nogales)	304	62,748	140,150
Valparaíso	Urban	560011	Conurbación San Antonio - Cartagena - Las Cruces + Santo Domingo	218	45,301	100,787
Valparaíso	Urban	560021	R.P. San Antonio (Algarrobo, El Quisco)	56	11,409	21,216
Valparaíso	Urban	570011	Prov. San Felipe	199	40,942	91,157
Valparaíso	Urban	580011	Conurbación Valparaíso (3) (Quilpué, Villa Alemana)	483	98,237	222,296
Valparaíso	Urban	580021	R.P. Marga Marga (Limache, Olmué)	89	18,210	41,153
Valparaíso	Rural	510012	Conurbación Valparaíso	33	2,841	5,551
Valparaíso	Rural	510022	R.P. Valparaíso (Casablanca)	38	3,492	7,226
Valparaíso	Rural	510032	R.P. Valparaíso (Puchuncaví, Quintero)	34	3,035	6,096

Region	Area	Code of Geographical stratum	Geographical stratum	PSU	Dwellings	WAP
Valparaíso	Rural	530012	Prov. Los Andes (Los Andes, Rinconada y San Esteban)	77	6,587	15,949
Valparaíso	Rural	540012	Prov. Petorca	107	9,924	20,814
Valparaíso	Rural	550012	Prov. Quillota	104	9,260	21,152
Valparaíso	Rural	560012	Prov. San Antonio (Cartagena - Las Cruces, Algarrobo, El Quisco, Santo Domingo)	72	6,582	12,979
Valparaíso	Rural	570012	Prov. San Felipe	157	13,949	30,862
Valparaíso	Rural	580022	R.P. Marga Marga (Limache, Olmué)	53	4,701	9,771
Metropolitan	Urban	1310011	Conurbación Gran Santiago (1) (Independencia, Quilicura, Recoleta, Huechuraba, Conchalí)	1,044	214,899	553,894
Metropolitan	Urban	1310021	Conurbación Gran Santiago (2) (Providencia, Ñuñoa, Lo Barnechea, Las Condes, Vitacura, La Reina)	1,814	368,243	775,975
Metropolitan	Urban	1310031	Conurbación Gran Santiago (3) (San Miguel, Peñalolén, La Florida, San Joaquín, Macul)	1,497	305,152	757,278
Metropolitan	Urban	1310041	Conurbación Gran Santiago (4) (La Pintana, Pedro Aguirre Cerda, Lo Espejo, La Cisterna, La Granja, El Bosque, San Ramón)	1,197	242,901	662,795
Metropolitan	Urban	1310051	Conurbación Gran Santiago (5) (Estación Central, Pudahuel, Quinta Normal, Renca, Cerrillos, Cerro Navia, Lo Prado, Maipú)	2,222	452,604	1,175,572
Metropolitan	Urban	1310061	Conurbación Gran Santiago (6) (Santiago)	963	191,494	355,195
Metropolitan	Urban	1320011	Conurbación Gran Santiago (7) (Puente Alto)	827	170,156	445,308
Metropolitan	Urban	1320021	R.P. Cordillera	35	7,124	18,234
Metropolitan	Urban	1330011	R.P. Chacabuco (Colina, Tiltil)	190	39,356	97,225
Metropolitan	Urban	1330021	Conurbación Gran Santiago (8) (Lampa)	128	26,200	58,963
Metropolitan	Urban	1340011	Conurbación Gran Santiago (9) (San Bernardo, Calera de Tango)	445	91,635	237,921
Metropolitan	Urban	1340021	Conurbación Buin - Paine	197	40,769	100,519
Metropolitan	Urban	1350011	Prov. Melipilla	192	38,298	90,078
Metropolitan	Urban	1360011	Conurbación Gran Santiago (10) (Talagante, Padre Hurtado, Peñaflor)	297	62,518	154,495
Metropolitan	Urban	1360021	R.P. Talagante	87	18,152	44,097
Metropolitan	Rural	1310012	Conurbación Gran Santiago (1)	37	2,890	9,063
Metropolitan	Rural	1320022	Prov. Cordillera	76	6,724	17,689
Metropolitan	Rural	1330012	R.P. Chacabuco (Colina, Tiltil)	107	9,846	25,842
Metropolitan	Rural	1330022	Conurbación Gran Santiago (8) (Lampa)	74	6,780	16,334
Metropolitan	Rural	1340012	Conurbación Gran Santiago (9) (San Bernardo, Calera de Tango)	60	5,541	15,040
Metropolitan	Rural	1340022	Conurbación Buin - Paine	135	12,682	31,258
Metropolitan	Rural	1350012	Prov. Melipilla	276	24,951	57,199
Metropolitan	Rural	1360012	Conurbación Gran Santiago (10) (Talagante, Padre Hurtado, Peñaflor)	101	9,075	23,114
Metropolitan	Rural	1360022	R.P. Talagante	53	4,968	11,994
O'Higgins	Urban	610011	Conurbación Rancagua - Machalí - Gultro - Los Lirios	597	124,471	266,228
O'Higgins	Urban	610021	R.P. Cachapoal (Codegua, Mostazal)	47	9,394	21,864
O'Higgins	Urban	610031	R.P. Cachapoal (Otras comunas)	255	50,988	115,789
O'Higgins	Urban	620011	Conurbación San Fernando	113	23,774	49,901
O'Higgins	Urban	620021	Prov. Cardenal Caro	47	9,330	18,268
O'Higgins	Urban	630011	R.P. Colchagua	137	28,124	62,308
O'Higgins	Rural	610012	Conurbación Rancagua - Machalí - Gultro - Los Lirios	101	8,874	22,546
O'Higgins	Rural	610022	R.P. Cachapoal (Codegua, Mostazal)	41	3,575	8,195
O'Higgins	Rural	610032	R.P. Cachapoal (Otras comunas)	375	33,393	74,442
O'Higgins	Rural	620012	Conurbación San Fernando	42	3,830	8,290
O'Higgins	Rural	620022	Prov. Cardenal Caro	111	10,347	19,547
O'Higgins	Rural	630012	R.P. Colchagua	284	25,770	57,407
Maule	Urban	710011	Conurbación Talca - Culenar	447	90,726	198,787

Region	Area	Code of Geographical stratum	Geographical stratum	PSU	Dwellings	WAP
Maule	Urban	710021	R.P. Talca (1) (Constitución, Empedrado, Pencahue, Curepto)	88	17,516	36,876
Maule	Urban	710031	R.P. Talca (2) (Pelarco, Rio Claro, San Rafael, San Clemente)	54	11,772	25,756
Maule	Urban	720011	Prov. Cauquenes	83	16,414	32,609
Maule	Urban	730011	R.P. Curicó (1) (Conurbación Curicó, Romeral, Teno)	267	55,067	121,728
Maule	Urban	730021	R.P. Curicó (2) (Licantén, Vichuquén, Hualañe, Molina, Sagrada Familia)	105	20,967	45,425
Maule	Urban	740011	Prov. Linares (1) (Conurbación Linares, Colbún, Yerbas Buenas)	164	33,555	72,830
Maule	Urban	740021	Prov. Linares (2) (Parral, Retiro, Longaví)	86	17,879	36,649
Maule	Urban	740031	Prov. Linares (3) (San Javier, Villa Alegre)	80	15,615	32,896
Maule	Rural	710012	Conurbación Talca - Culenar	79	7,208	16,100
Maule	Rural	710022	R.P. Talca (1) (Constitución, Empedrado, Pencahue, Curepto)	91	8,315	17,100
Maule	Rural	710032	R.P. Talca (3) (Pelarco, Río Claro, San Clemente, San Rafael)	166	15,483	33,695
Maule	Rural	720012	Prov. Cauquenes	76	7,033	13,077
Maule	Rural	730012	R.P. Curicó (Conurbación Curicó, Molina, Sagrada Familia)	167	15,042	33,284
Maule	Rural	730022	R.P. Curicó (Romeral, Teno)	109	10,002	22,864
Maule	Rural	730032	R.P. Curicó (Licantén, Vichuquén, Hualañe)	42	3,801	7,286
Maule	Rural	740012	Prov. Linares (1) (Conurbación Linares, Colbún, Yerbas Buenas)	166	14,811	31,924
Maule	Rural	740022	Prov. Linares (2) (Parral, Retiro, Longaví)	189	17,072	36,692
Maule	Rural	740032	Prov. Linares (3) (San Javier, Villa Alegre)	86	7,803	16,087
Ñuble	Urban	1610011	Conurbación Chillán - Chillán Viejo	336	70,344	156,551
Ñuble	Urban	1610021	R.P. Diguillín	110	22,192	46,739
Ñuble	Urban	1620011	Prov. Itata	53	10,996	23,139
Ñuble	Urban	1630011	Prov. Punilla	92	18,597	39,832
Ñuble	Rural	1610012	Conurbación Chillán - Chillán Viejo	81	7,294	15,760
Ñuble	Rural	1610022	R.P. Diguillín	201	18,251	37,646
Ñuble	Rural	1620012	Prov. Itata	117	10,609	20,853
Ñuble	Rural	1630012	Prov. Punilla	235	21,544	46,203
Biobío	Urban	810011	Conurbación Gran Concepción + Florida + Santa Juana (1) (Santa Juana, Lota, Coronel)	281	57,170	129,702
Biobío	Urban	810021	Conurbación Gran Concepción + Florida + Santa Juana (2) (Florida, Tomé, Hualqui, Penco, Talcahuano)	447	92,101	218,980
Biobío	Urban	810031	Conurbación Gran Concepción + Florida + Santa Juana (3) (Concepción, Hualpén, Chiguayante, San Pedro de la Paz)	931	189,042	429,134
Biobío	Urban	820011	Prov. Arauco	210	44,059	97,483
Biobío	Urban	830011	Conurbación Los Ángeles	264	54,355	117,238
Biobío	Urban	830021	R.P. Biobío (Cabrero, Laja, Nacimiento, Negrete, San Rosendo, Yumbel)	144	29,455	64,354
Biobío	Urban	830031	R.P. Biobío (Antuco, Mulchén, Quilaco, Quilleco, Santa Bárbara, Tucapel)	95	19,479	41,860
Biobío	Rural	810012	Prov. Concepción (Conurbación Gran Concepción, R.P. Concepción (Florida), R.P. Concepción (Santa Juana))	130	11,678	24,546
Biobío	Rural	820012	Prov. Arauco	161	14,628	31,198
Biobío	Rural	830012	Conurbación Los Ángeles	203	18,385	40,670
Biobío	Rural	830022	R.P. Biobío (Cabrero, Laja, Nacimiento, Negrete, San Rosendo, Yumbel)	137	12,273	24,219
Biobío	Rural	830032	R.P. Biobío (Antuco, Mulchén, Quilaco, Quilleco, Santa Bárbara, Tucapel, Alto Biobío)	123	10,913	21,944
La Araucanía	Urban	910011	Conurbación Temuco - Padre Las Casas - Cajón	570	117,806	258,871

Region	Area	Code of Geographical stratum	Geographical stratum	PSU	Dwellings	WAP
La Araucanía	Urban	910021	R.P. Cautín (1) (Carahue, Freire, Galvarino, Lautaro, Nueva Imperial, Perquenco, Saavedra, Teodoro Schmidt, Toltén, Cholchol)	159	32,450	70,617
La Araucanía	Urban	910031	R.P. Cautín (2) (Cunco, Curarrehue, Gorbea, Loncoche, Melipeuco, Pitrufquén, Pucón, Villarrica)	209	42,076	87,898
La Araucanía	Urban	920011	Conurbación Angol	100	20,314	44,819
La Araucanía	Urban	920021	R.P. Malleco (1) (Ercilla, Los Sauces, Lumaco, Purén)	38	7,746	15,386
La Araucanía	Urban	920031	R.P. Malleco (2) (Collipulli, Curacautín, Lonquimay, Traiguén, Victoria)	133	27,646	58,671
La Araucanía	Rural	910012	Conurbación Temuco - Padre Las Casas - Cajón	244	21,407	48,631
La Araucanía	Rural	910022	R.P. Cautín (1) (Carahue, Freire, Galvarino, Lautaro, Nueva Imperial, Perquenco, Saavedra, Teodoro Schmidt, Toltén, Cholchol)	401	36,556	76,307
La Araucanía	Rural	910032	R.P. Cautín (2) (Cunco, Curarrehue, Gorbea, Loncoche, Melipeuco, Pitrufquén, Pucón, Villarrica)	312	28,285	53,358
La Araucanía	Rural	920012	Prov. Malleco (Angol, Ercilla, Los Sauces, Lumaco, Purén, Collipulli, Curacautín, Lonquimay, Traiguén, Victoria)	232	20,806	42,794
Los Ríos	Urban	1410011	Conurbación Valdivia	275	55,012	125,829
Los Ríos	Urban	1410021	R.P. Valdivia (Corral, Los Lagos, Máfil, Mariquina, Paillaco)	69	13,891	31,018
Los Ríos	Urban	1410031	R.P. Valdivia (Lanco, Panguipulli)	50	9,857	20,903
Los Ríos	Urban	1420011	Prov. Ranco (Futrono, Lago Ranco, Río Bueno)	53	10,704	21,759
Los Ríos	Urban	1420021	Prov. Ranco (La Unión)	48	9,801	21,073
Los Ríos	Rural	1410012	Conurbación Valdivia	47	4,301	9,291
Los Ríos	Rural	1410022	R.P. Valdivia	148	13,071	26,601
Los Ríos	Rural	1410032	R.P. Valdivia (Lanco, Panguipulli)	106	9,689	19,366
Los Ríos	Rural	1420012	Prov. Ranco (Futrono, Lago Ranco, Río Bueno)	131	12,006	22,728
Los Ríos	Rural	1420022	Prov. Ranco (La Unión)	50	4,673	9,201
Los Lagos	Urban	1010011	Conurbación Puerto Montt	442	92,056	196,779
Los Lagos	Urban	1010021	R.P. Llanquihue (1) (Calbuco, Fresia, Los Muermos, Maullín)	71	14,234	29,665
Los Lagos	Urban	1010031	R.P. Llanquihue (2) (Frutillar, Llanquihue)	49	9,679	21,458
Los Lagos	Urban	1020011	Conurbación Osorno	269	54,355	117,710
Los Lagos	Urban	1020021	Prov. Chiloé + Palena	207	42,413	85,988
Los Lagos	Urban	1030011	R.P. Osorno	62	12,442	27,181
Los Lagos	Rural	1010012	Conurbación Puerto Montt + Cochamó	184	16,967	33,746
Los Lagos	Rural	1010022	R.P. Llanquihue (1) (Calbuco, Fresia, Los Muermos, Maullín)	179	16,646	32,336
Los Lagos	Rural	1010032	R.P. Llanquihue (2) (Frutillar, Llanquihue)	42	3,699	7,185
Los Lagos	Rural	1020012	Conurbación Osorno	55	5,175	10,987
Los Lagos	Rural	1020022	Prov. Chiloé + Palena	365	32,865	62,013
Los Lagos	Rural	1030012	R.P. Osorno	171	15,763	31,483
Aysén	Urban	1110011	Conurbación Coyhaique	92	18,590	38,040
Aysén	Urban	1110021	R.R. Aysén	61	12,528	24,466
Aysén	Rural	1110012	Conurbación Coyhaique	48	4,282	6,307
Aysén	Rural	1110022	R.R. Aysén	71	6,282	11,142
Magallanes	Urban	1200011	Conurbación Punta Arenas + Cabo de Hornos + Porvenir	243	48,608	108,428
Magallanes	Urban	1200021	Prov. Última Esperanza	39	7,820	15,427
Magallanes	Rural	1200012	Conurbación Punta Arenas + Laguna Blanca + Río Verde + San Gregorio	28	2,609	6,174
Magallanes	Rural	1200022	R.R. Magallanes	23	1,916	5,278

Source: Produced by the National Statistics Institute

Annex 2 Sample stratification of the sampling frame of the ENE

Region	Area	Code of Geographical stratum	Code of sampling stratum	PSU	Dwellings
Arica y Parinacota	Urban	1500001	15000011	108	21,395
Arica y Parinacota	Urban	1500001	15000012	178	35,494
Arica y Parinacota	Urban	1500001	15000013	45	9,037
Arica y Parinacota	Rural	1500012	15000120	66	5,94
Tarapacá	Urban	110011	1100111	110	22,314
Tarapacá	Urban	110011	1100112	244	49,622
Tarapacá	Urban	110011	1100113	123	24,235
Tarapacá	Urban	120011	1200110	27	5,37
Tarapacá	Rural	110012	1100120	13	1,027
Tarapacá	Rural	120012	1200120	40	3,71
Antofagasta	Urban	210001	2100011	126	25,097
Antofagasta	Urban	210001	2100012	275	54,597
Antofagasta	Urban	210001	2100013	180	36,556
Antofagasta	Urban	220011	2200110	263	54
Antofagasta	Urban	230011	2300110	51	10,638
Antofagasta	Rural	210002	2100020	18	1,383
Antofagasta	Rural	220012	2200120	48	4,287
Atacama	Urban	310011	3100111	115	23,084
Atacama	Urban	310011	3100112	105	21,488
Atacama	Urban	310011	3100113	59	11,72
Atacama	Urban	310021	3100210	79	15,842
Atacama	Urban	320011	3200110	104	20,8
Atacama	Rural	310012	3100120	33	2,876
Atacama	Rural	310022	3100220	26	2,147
Atacama	Rural	330012	3300120	80	7,086
Coquimbo	Urban	410011	4100111	186	37,603
Coquimbo	Urban	410011	4100112	392	78,497
Coquimbo	Urban	410011	4100113	156	31,984
Coquimbo	Urban	410021	4100210	53	10,134
Coquimbo	Urban	420011	4200110	104	21,254
Coquimbo	Urban	430011	4300111	48	9,664
Coquimbo	Urban	430011	4300112	102	20,481
Coquimbo	Urban	430021	4300210	51	10,124
Coquimbo	Rural	410012	4100120	131	11,955
Coquimbo	Rural	410022	4100220	87	7,896
Coquimbo	Rural	420012	4200120	150	13,567
Coquimbo	Rural	430012	4300120	112	10,042
Coquimbo	Rural	430022	4300220	166	14,859
Valparaíso	Urban	510011	5100111	148	29,47
Valparaíso	Urban	510011	5100112	290	58,419
Valparaíso	Urban	510011	5100113	111	22,57
Valparaíso	Urban	510021	5100211	90	18,227
Valparaíso	Urban	510021	5100212	381	77,601
Valparaíso	Urban	510021	5100213	257	52,307
Valparaíso	Urban	510031	5100310	30	6,153
Valparaíso	Urban	510041	5100410	79	16,383
Valparaíso	Urban	530011	5300111	50	10,053
Valparaíso	Urban	530011	5300112	108	22,502
Valparaíso	Urban	540011	5400110	99	19,473
Valparaíso	Urban	550011	5500111	87	17,775
Valparaíso	Urban	550011	5500112	179	36,809
Valparaíso	Urban	550011	5500113	38	8,164
Valparaíso	Urban	560011	5600111	78	16,353
Valparaíso	Urban	560011	5600112	140	28,948

Region	Area	Code of Geographical stratum	Code of sampling stratum	PSU	Dwellings
Valparaíso	Urban	560021	5600210	56	11,409
Valparaíso	Urban	570011	5700110	199	40,942
Valparaíso	Urban	580011	5800111	48	9,431
Valparaíso	Urban	580011	5800112	365	74,022
Valparaíso	Urban	580011	5800113	70	14,784
Valparaíso	Urban	580021	5800210	89	18,21
Valparaíso	Rural	510012	5100120	33	2,841
Valparaíso	Rural	510022	5100220	38	3,492
Valparaíso	Rural	510032	5100320	33	2,939
Valparaíso	Rural	530012	5300120	76	6,493
Valparaíso	Rural	540012	5400120	106	9,825
Valparaíso	Rural	550012	5500120	104	9,26
Valparaíso	Rural	560012	5600120	72	6,582
Valparaíso	Rural	570012	5700120	155	13,762
Valparaíso	Rural	580022	5800220	50	4,421
Metropolitan	Urban	1310011	13100111	108	22,402
Metropolitan	Urban	1310011	13100112	642	131,764
Metropolitan	Urban	1310011	13100113	294	60,733
Metropolitan	Urban	1310021	13100210	1,814	368,243
Metropolitan	Urban	1310031	13100311	75	15,091
Metropolitan	Urban	1310031	13100312	835	170,211
Metropolitan	Urban	1310031	13100313	587	119,85
Metropolitan	Urban	1310041	13100411	338	68,167
Metropolitan	Urban	1310041	13100412	788	159,994
Metropolitan	Urban	1310041	13100413	71	14,74
Metropolitan	Urban	1310051	13100511	204	42,132
Metropolitan	Urban	1310051	13100512	1,538	313,245
Metropolitan	Urban	1310051	13100513	480	97,227
Metropolitan	Urban	1310061	13100610	963	191,494
Metropolitan	Urban	1320011	13200111	62	12,733
Metropolitan	Urban	1320011	13200112	588	119,951
Metropolitan	Urban	1320011	13200113	177	37,472
Metropolitan	Urban	1320021	13200210	35	7,124
Metropolitan	Urban	1330011	13300111	37	7,928
Metropolitan	Urban	1330011	13300112	99	20,614
Metropolitan	Urban	1330011	13300113	54	10,814
Metropolitan	Urban	1330021	13300210	128	26,2
Metropolitan	Urban	1340011	13400111	92	18,429
Metropolitan	Urban	1340011	13400112	292	60,013
Metropolitan	Urban	1340011	13400113	61	13,193
Metropolitan	Urban	1340021	13400211	48	9,847
Metropolitan	Urban	1340021	13400212	149	30,922
Metropolitan	Urban	1350011	13500111	54	10,681
Metropolitan	Urban	1350011	13500112	138	27,617
Metropolitan	Urban	1360011	13600111	63	13,157
Metropolitan	Urban	1360011	13600112	188	39,476
Metropolitan	Urban	1360011	13600113	46	9,885
Metropolitan	Urban	1360021	13600210	87	18,152
Metropolitan	Rural	1310012	13100120	37	2,89
Metropolitan	Rural	1320022	13200220	76	6,724
Metropolitan	Rural	1330012	13300120	107	9,846
Metropolitan	Rural	1330022	13300220	74	6,78
Metropolitan	Rural	1340012	13400120	60	5,541
Metropolitan	Rural	1340022	13400220	135	12,682
Metropolitan	Rural	1350012	13500121	171	15,474
Metropolitan	Rural	1350012	13500122	105	9,477

Region	Area	Code of Geographical stratum	Code of sampling stratum	PSU	Dwellings
Metropolitan	Rural	1360012	13600120	101	9,075
Metropolitan	Rural	1360022	13600220	53	4,968
O'Higgins	Urban	610011	6100111	116	23,91
O'Higgins	Urban	610011	6100112	364	76,037
O'Higgins	Urban	610011	6100113	117	24,524
O'Higgins	Urban	610021	6100210	47	9,394
O'Higgins	Urban	610031	6100310	255	50,988
O'Higgins	Urban	620011	6200110	113	23,774
O'Higgins	Urban	620021	6200210	47	9,33
O'Higgins	Urban	630011	6300110	137	28,124
O'Higgins	Rural	610012	6100120	100	8,773
O'Higgins	Rural	610022	6100220	41	3,575
O'Higgins	Rural	610032	6100320	374	33,299
O'Higgins	Rural	620012	6200120	40	3,664
O'Higgins	Rural	620022	6200220	111	10,347
O'Higgins	Rural	630012	6300121	224	20,574
O'Higgins	Rural	630012	6300122	58	5,033
Maule	Urban	710011	7100111	95	19,412
Maule	Urban	710011	7100112	253	51,356
Maule	Urban	710011	7100113	99	19,958
Maule	Urban	710021	7100210	88	17,516
Maule	Urban	710031	7100310	54	11,772
Maule	Urban	720011	7200110	83	16,414
Maule	Urban	730011	7300111	54	11,201
Maule	Urban	730011	7300112	167	34,493
Maule	Urban	730011	7300113	46	9,373
Maule	Urban	730021	7300210	105	20,967
Maule	Urban	740011	7400110	164	33,555
Maule	Urban	740021	7400210	86	17,879
Maule	Urban	740031	7400310	80	15,615
Maule	Rural	710012	7100120	79	7,208
Maule	Rural	710022	7100220	90	8,223
Maule	Rural	710032	7100320	166	15,483
Maule	Rural	720012	7200120	74	6,852
Maule	Rural	730012	7300120	166	14,954
Maule	Rural	730022	7300220	109	10,002
Maule	Rural	730032	7300320	42	3,801
Maule	Rural	740012	7400120	165	14,715
Maule	Rural	740022	7400220	189	17,072
Maule	Rural	740032	7400320	86	7,803
Ñuble	Urban	1610011	16100111	78	16,165
Ñuble	Urban	1610011	16100112	190	39,559
Ñuble	Urban	1610011	16100113	68	14,62
Ñuble	Urban	1610021	16100210	110	22,192
Ñuble	Urban	1620011	16200110	53	10,996
Ñuble	Urban	1630011	16300110	92	18,597
Ñuble	Rural	1610012	16100120	81	7,294
Ñuble	Rural	1610022	16100220	196	17,795
Ñuble	Rural	1620012	16200120	116	10,517
Ñuble	Rural	1630012	16300120	233	21,366
Biobío	Urban	810011	8100111	171	34,372
Biobío	Urban	810011	8100112	110	22,798
Biobío	Urban	810021	8100211	159	32,852
Biobío	Urban	810021	8100212	252	52,006
Biobío	Urban	810021	8100213	36	7,243
Biobío	Urban	810031	8100311	152	30,126

Region	Area	Code of Geographical stratum	Code of sampling stratum	PSU	Dwellings
Biobío	Urban	810031	8100312	458	93,403
Biobío	Urban	810031	8100313	321	65,513
Biobío	Urban	820011	8200111	146	30,444
Biobío	Urban	820011	8200112	64	13,615
Biobío	Urban	830011	8300111	81	16,354
Biobío	Urban	830011	8300112	129	26,454
Biobío	Urban	830011	8300113	54	11,547
Biobío	Urban	830021	8300210	144	29,455
Biobío	Urban	830031	8300310	95	19,479
Biobío	Rural	810012	8100120	128	11,465
Biobío	Rural	820012	8200120	156	14,204
Biobío	Rural	830012	8300120	203	18,385
Biobío	Rural	830022	8300220	131	11,717
Biobío	Rural	830032	8300320	119	10,574
La Araucanía	Urban	910011	9100111	119	24,526
La Araucanía	Urban	910011	9100112	288	59,39
La Araucanía	Urban	910011	9100113	163	33,89
La Araucanía	Urban	910021	9100210	159	32,45
La Araucanía	Urban	910031	9100310	209	42,076
La Araucanía	Urban	920011	9200110	100	20,314
La Araucanía	Urban	920021	9200210	38	7,746
La Araucanía	Urban	920031	9200310	133	27,646
La Araucanía	Rural	910012	9100120	244	21,407
La Araucanía	Rural	910022	9100220	400	36,468
La Araucanía	Rural	910032	9100320	273	24,667
La Araucanía	Rural	920012	9200120	223	20,018
Los Ríos	Urban	1410011	14100111	55	10,799
Los Ríos	Urban	1410011	14100112	154	31,01
Los Ríos	Urban	1410011	14100113	66	13,203
Los Ríos	Urban	1410021	14100210	69	13,891
Los Ríos	Urban	1410031	14100310	50	9,857
Los Ríos	Urban	1420011	14200110	53	10,704
Los Ríos	Urban	1420021	14200210	48	9,801
Los Ríos	Rural	1410012	14100120	45	4,117
Los Ríos	Rural	1410022	14100220	141	12,499
Los Ríos	Rural	1410032	14100320	101	9,227
Los Ríos	Rural	1420012	14200120	124	11,383
Los Ríos	Rural	1420022	14200220	48	4,491
Los Lagos	Urban	1010011	10100111	106	22,37
Los Lagos	Urban	1010011	10100112	221	46,083
Los Lagos	Urban	1010011	10100113	115	23,603
Los Lagos	Urban	1010021	10100210	71	14,234
Los Lagos	Urban	1010031	10100310	49	9,679
Los Lagos	Urban	1020011	10200111	91	18,375
Los Lagos	Urban	1020011	10200112	131	26,16
Los Lagos	Urban	1020011	10200113	47	9,82
Los Lagos	Urban	1020021	10200211	59	12,016
Los Lagos	Urban	1020021	10200212	129	26,726
Los Lagos	Urban	1030011	10300110	62	12,442
Los Lagos	Rural	1010012	10100121	57	5,263
Los Lagos	Rural	1010012	10100122	104	9,547
Los Lagos	Rural	1010022	10100220	152	14,208
Los Lagos	Rural	1010032	10100320	42	3,699
Los Lagos	Rural	1020012	10200120	52	4,921
Los Lagos	Rural	1020022	10200221	188	17,166
Los Lagos	Rural	1020022	10200222	97	8,667

Region	Area	Code of Geographical stratum	Code of sampling stratum	PSU	Dwellings
Los Lagos	Rural	1030012	10300120	156	14,42
Aysén	Urban	1110011	11100110	92	18,59
Aysén	Urban	1110021	11100210	58	11,952
Aysén	Rural	1110012	11100120	48	4,282
Aysén	Rural	1110022	11100220	63	5,681
Magallanes	Urban	1200011	12000112	185	37,054
Magallanes	Urban	1200011	12000113	58	11,554
Magallanes	Urban	1200021	12000210	39	7,82
Magallanes	Rural	1200012	12000120	27	2,52
Magallanes	Rural	1200022	12000220	22	1,846

Source: Produced by the National Statistics Institute

Annex 3 Visitation codes of the ENE

Contact registration codes for the Household Registration Card used in the previous sampling design of the ENE.

Code		Description
Attained		
L	10	Interviewed
Not attained, Type "A"		
A1	11	Absent residents
A2	12	Rejection
A3	13	Inaccessible
A4	14	Special Cases
Not attained, Type "B"		
B1	21	Temporally unoccupied
B2	22	Being repaired, expanded, or remodeled
B3	23	Temporary exclusion
B4	24	Seasonal
Not attained, Type "C"		
C1	31	Demolished
C2	32	Other use
C3	33	On display
C5	35	Incorrectly conducted

Source: Produced by the National Statistics Institute

Contact or disposition codes of the ENE 2020.

Code	Description
Elegible, interviewed	
11	Completed interview
12	Partial interview
Elegible, rejected	
21	Rejected by household
Elegible, not contacted	
31	Occupied dwelling, no residents present
32	Access to dwelling is denied
33	Informant cannot be located, other person contacted
34	Informant is absent during the application of the survey
Elegible, not interviewed for other reasons	
41	Informant physically and/or mentally unable to answer
42	Informant does not speak Spanish
43	Death of the informant or a family member
44	Appointment arranged with informant
49	Other reason for not being interviewed
Unknown eligibility	
51	Was not sent to the field
52	Address could not be located
53	Inaccessible for reasons related to climate or natural disasters
54	Dangerous area
59	Other reason for unknown eligibility
Ineligible	
61	Company, government office, or other organization
62	Institution (hospital, prison, retirement home, etc.)
63	Collective housing (military, work, boarding institution, hostels, lodges, etc.)
64	Dwelling demolished, burned, destroyed, or eradicated
65	Unoccuped dwelling (for example, on sale, for rent)
66	Summer dwelling or for temporary use
67	Unqualified building (for example, under construction or repair, etc.)
69	Other reason for being considered ineligible

Source: Produced by the National Statistics Institute

Annex 4 Non-attainment rate by geographical stratum according to data from the ENE 2018

Region	Area	Code of Geographical stratum	Non-attainment rate	Non-attainment rate (truncated at 30%)
Arica y Parinacota	Urban	1500001	22%	22%
Arica y Parinacota	Rural	1500012	33%	30%
Tarapacá	Urban	120011	22%	22%
Tarapacá	Rural	110011	16%	16%
Tarapacá	Rural	120012	30%	30%
Tarapacá	Rural	110012	33%	30%
Antofagasta	Urban	210001	23%	23%
Antofagasta	Urban	220011	20%	20%
Antofagasta	Urban	230011	31%	30%
Antofagasta	Rural	220012	23%	23%
Antofagasta	Rural	210002	23%	23%
Atacama	Urban	310021	33%	30%
Atacama	Urban	310011	27%	27%
Atacama	Urban	320011	27%	27%
Atacama	Rural	310022	45%	30%
Atacama	Rural	310012	36%	30%
Atacama	Rural	330012	48%	30%
Coquimbo	Urban	420011	31%	30%
Coquimbo	Urban	430011	25%	25%
Coquimbo	Urban	430021	26%	26%
Coquimbo	Urban	410021	37%	30%
Coquimbo	Urban	410011	36%	30%
Coquimbo	Rural	430012	25%	25%
Coquimbo	Rural	420012	29%	29%
Coquimbo	Rural	410012	35%	30%
Coquimbo	Rural	430022	26%	26%
Coquimbo	Rural	410022	36%	30%
Metropolitan	Urban	1310061	35%	30%
Metropolitan	Urban	1320011	24%	24%
Metropolitan	Urban	1310051	25%	25%
Metropolitan	Urban	1310041	28%	28%
Metropolitan	Urban	1310031	25%	25%
Metropolitan	Urban	1310021	39%	30%
Metropolitan	Urban	1310011	31%	30%
Metropolitan	Urban	1360021	25%	25%
Metropolitan	Urban	1360011	28%	28%
Metropolitan	Urban	1350011	23%	23%
Metropolitan	Urban	1340021	23%	23%
Metropolitan	Urban	1340011	31%	30%
Metropolitan	Urban	1330021	25%	25%
Metropolitan	Urban	1330011	28%	28%
Metropolitan	Urban	1320021	36%	30%
Metropolitan	Rural	1310012	36%	30%
Metropolitan	Rural	1360022	36%	30%
Metropolitan	Rural	1360012	36%	30%
Metropolitan	Rural	1350012	50%	30%
Metropolitan	Rural	1340022	36%	30%
Metropolitan	Rural	1340012	36%	30%
Metropolitan	Rural	1330022	23%	23%
Metropolitan	Rural	1330012	37%	30%
Metropolitan	Rural	1320022	36%	30%
Valparaíso	Urban	510031	21%	21%
Valparaíso	Urban	510021	27%	27%
Valparaíso	Urban	510011	25%	25%

Region	Area	Code of Geographical stratum	Non-attainment rate	Non-attainment rate (truncated at 30%)
Valparaíso	Urban	540011	41%	30%
Valparaíso	Urban	580021	25%	25%
Valparaíso	Urban	580011	17%	17%
Valparaíso	Urban	560021	39%	30%
Valparaíso	Urban	570011	18%	18%
Valparaíso	Urban	550011	18%	18%
Valparaíso	Urban	530011	19%	19%
Valparaíso	Urban	560011	18%	18%
Valparaíso	Urban	510041	25%	25%
Valparaíso	Rural	510032	27%	27%
Valparaíso	Rural	530012	29%	29%
Valparaíso	Rural	510022	28%	28%
Valparaíso	Rural	510012	27%	27%
Valparaíso	Rural	580022	25%	25%
Valparaíso	Rural	570012	37%	30%
Valparaíso	Rural	560012	26%	26%
Valparaíso	Rural	550012	26%	26%
Valparaíso	Rural	540012	25%	25%
O'Higgins	Urban	630011	23%	23%
O'Higgins	Urban	620011	20%	20%
O'Higgins	Urban	610031	17%	17%
O'Higgins	Urban	610021	23%	23%
O'Higgins	Urban	620021	22%	22%
O'Higgins	Urban	610011	24%	24%
O'Higgins	Rural	630012	37%	30%
O'Higgins	Rural	620022	20%	20%
O'Higgins	Rural	620012	17%	17%
O'Higgins	Rural	610032	17%	17%
O'Higgins	Rural	610022	20%	20%
O'Higgins	Rural	610012	39%	30%
Maule	Urban	710031	32%	30%
Maule	Urban	710021	43%	30%
Maule	Urban	710011	36%	30%
Maule	Urban	740031	37%	30%
Maule	Urban	740021	35%	30%
Maule	Urban	740011	25%	25%
Maule	Urban	730021	36%	30%
Maule	Urban	730011	40%	30%
Maule	Urban	720011	31%	30%
Maule	Rural	710022	26%	26%
Maule	Rural	710012	23%	23%
Maule	Rural	740032	27%	27%
Maule	Rural	740022	15%	15%
Maule	Rural	740012	14%	14%
Maule	Rural	730022	13%	13%
Maule	Rural	730012	23%	23%
Maule	Rural	730032	17%	17%
Maule	Rural	720012	18%	18%
Maule	Rural	710032	15%	15%
Ñuble	Urban	1630011	22%	22%
Ñuble	Urban	1620011	26%	26%
Ñuble	Urban	1610021	32%	30%
Ñuble	Urban	1610011	23%	23%
Ñuble	Rural	1630012	19%	19%
Ñuble	Rural	1620012	21%	21%
Ñuble	Rural	1610022	21%	21%

Region	Area	Code of Geographical stratum	Non-attainment rate	Non-attainment rate (truncated at 30%)
Ñuble	Rural	1610012	30%	30%
Biobío	Urban	830031	30%	30%
Biobío	Urban	830021	37%	30%
Biobío	Urban	830011	26%	26%
Biobío	Urban	820011	25%	25%
Biobío	Urban	810031	26%	26%
Biobío	Urban	810021	22%	22%
Biobío	Urban	810011	18%	18%
Biobío	Rural	830032	35%	30%
Biobío	Rural	830022	29%	29%
Biobío	Rural	830012	39%	30%
Biobío	Rural	820012	48%	30%
Biobío	Rural	810012	57%	30%
La Araucanía	Urban	920031	27%	27%
La Araucanía	Urban	920021	24%	24%
La Araucanía	Urban	920011	24%	24%
La Araucanía	Urban	910031	23%	23%
La Araucanía	Urban	910021	26%	26%
La Araucanía	Urban	910011	29%	29%
La Araucanía	Rural	920012	15%	15%
La Araucanía	Rural	910032	15%	15%
La Araucanía	Rural	910022	22%	22%
La Araucanía	Rural	910012	17%	17%
Los Ríos	Urban	1420021	30%	30%
Los Ríos	Urban	1420011	30%	30%
Los Ríos	Urban	1410031	30%	30%
Los Ríos	Urban	1410021	35%	30%
Los Ríos	Urban	1410011	30%	30%
Los Ríos	Rural	1420022	23%	23%
Los Ríos	Rural	1420012	30%	30%
Los Ríos	Rural	1410032	32%	30%
Los Ríos	Rural	1410022	19%	19%
Los Ríos	Rural	1410012	31%	30%
Los Lagos	Urban	1030011	39%	30%
Los Lagos	Urban	1020021	35%	30%
Los Lagos	Urban	1020011	29%	29%
Los Lagos	Urban	1010031	32%	30%
Los Lagos	Urban	1010021	26%	26%
Los Lagos	Urban	1010011	29%	29%
Los Lagos	Rural	1030012	23%	23%
Los Lagos	Rural	1020022	29%	29%
Los Lagos	Rural	1020012	32%	30%
Los Lagos	Rural	1010032	42%	30%
Los Lagos	Rural	1010022	23%	23%
Los Lagos	Rural	1010012	27%	27%
Aysén	Urban	1110021	34%	30%
Aysén	Urban	1110011	27%	27%
Aysén	Rural	1110022	45%	30%
Aysén	Rural	1110012	31%	30%
Magallanes	Urban	1200021	21%	21%
Magallanes	Urban	1200011	13%	13%
Magallanes	Rural	1200022	12%	12%
Magallanes	Rural	1200012	12%	12%

Source: Produced by the National Statistics Institute

Annex 5 Population estimates for the National Employment Survey (ENE)

Report on population estimates for the National Employment Survey (ENE)

This document describes the methodology used in the monthly population estimates for the National Employment Survey (ENE). The calculation of the indicators of the ENE requires population estimates at the commune level, by sex and age, and at the urban/rural region level by sex and age. These population estimates by sex and age for communes were interpolated from the annual population projections for communes of 30 June. The population was estimated on the fifteen day of each month. With these communal estimates, the regional estimates by sex and age were obtained. On the other hand, the regional urban population by sex and age was estimated on the fifteenth of each month. This estimation was interpolated from the regional urban population projections to obtain the estimation of the regional rural population by measuring the difference between the urban population and the total regional population, which was previously obtained by adding together the population of all communes in the region. The steps followed in this methodology will be specified below.

Interpolation of the communal on the fifteenth of each month

Population projections are calculated for 30 June of each year. Population estimates for other dates are obtained by an interpolation that uses projections of the years surrounding that date as pivots. The employment survey is continuous. Therefore, in order to obtain the population estimates on the fifteenth of each month, the population projections for two continuous years are successively interpolated, thus obtaining successive population estimates between 15 July and 15 June of the following year. The use of the pivots ensures that the estimates do not deviate from the official population projections published by INE.

In view of the requirements for the calculation of the ENE indicators, the interpolation between the successive pivots was a simple exponential interpolation by sex and age. With the exponential interpolation, eleven estimates were first obtained for the thirtieth day of each month between the two pivot populations. The estimates were rounded to the nearest whole number without decimals, and then simple averages were applied between the thirtieth days of each consecutive month to obtain estimates on the fifteenth day of each month, the date required by the ENE for calculating the indicators of each moving quarter. For each age group, the estimates for fifteenth day of each month in turn were rounded and truncated to prevent the average from approaching either of the two averages of monthly estimates.

$$P_f = P_i \cdot e^{r \cdot \Delta t'}$$

$$r = ln\left(\frac{P_f}{P_i}\right) \cdot \frac{1}{\Delta t}$$

Where,

P_f: Final pivot population

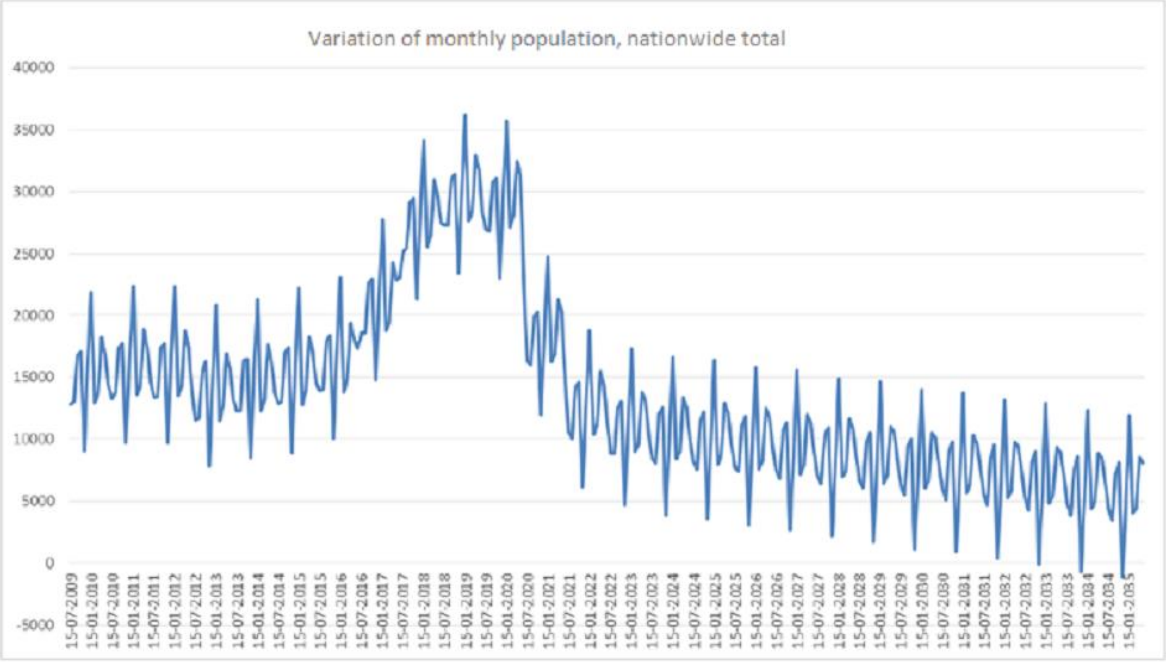
P_i: Initial pivot population

r: Annual growth rate

Δt: Time interval between pivot populations

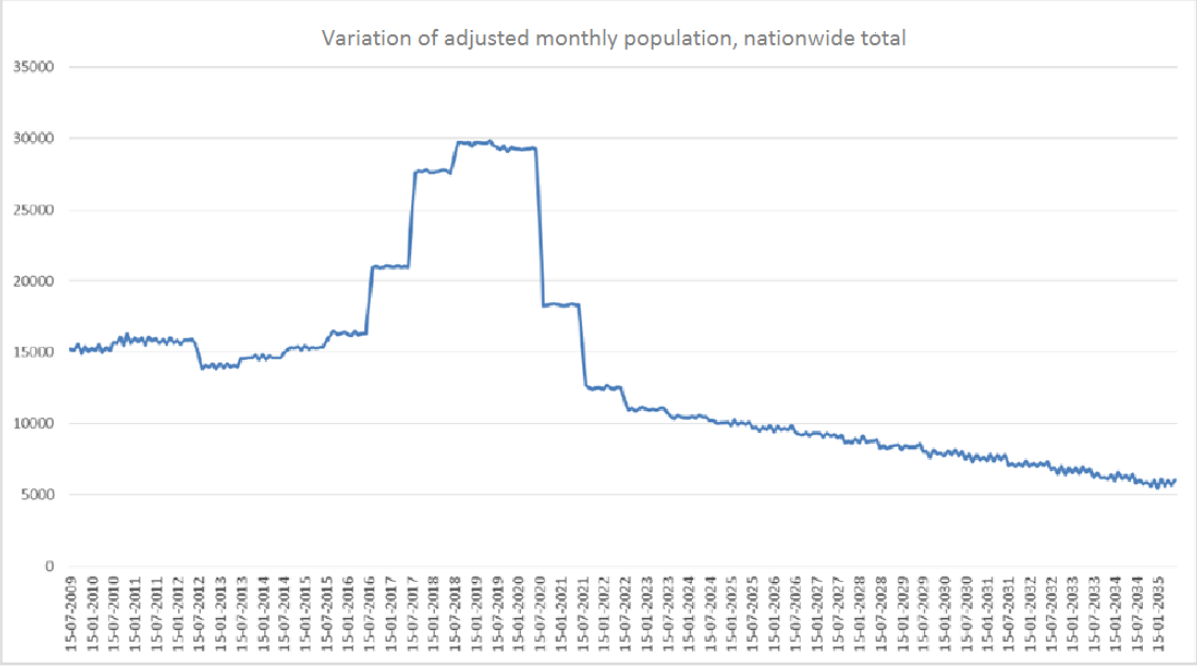
$\Delta t'$: Time interval between the pivot population and the population to be estimated

Because of their exponential interpolation at the level of age, the annual differences between pivot populations were in many cases very small. When the series was rounded to whole numbers, systematical increases (or decreases) of the population were observed in certain months. For example, if the increase in population between two pivots was two persons, the first addition was automatically designated as taking place in October and the second in April. This led to a highly irregular monthly population growth during the months of the year at the national, regional, and communal level, and this irregularity was systematically repeated from year to year.



Source: Produced by the National Statistics Institute

To eliminate these irregularities, a systematic table was used to redistribute the monthly population growth of each specific age for the months indicated in the table. This avoided the sharp fluctuations in population growth observed in the months of the same annual interpolation.



Source: Produced by the National Statistics Institute

Because annual variations come from the national population projection, they cannot be smoothed. However, the fluctuations in the monthly population increases initially observed were avoided. With these populations by sex and age, the communes were established, and with the sum of the communes, the regions were established.

Using the same methodology, population estimates were calculated for urban regions by sex and age. To do this, the regional urban population projections published by INE were used as pivots, and the same procedure described above was followed for the estimation of the communal population. The rural population at the regional level was obtained by the difference between the total regional population (by sex and age) and the regional urban population.

Annex 6 List of auxiliary variables from provinces

Region	Province Code	Province	Population projection WAP, October 2019	Quarter	Year of entry
Tarapacá	11	Iquique	265,356	JFM	2010
Antofagasta	21	Antofagasta	352,067	JFM	2010
Antofagasta	22	El Loa	152,521	JFM	2010
Atacama	31	Copiapó	157,71	JFM	2010
Coquimbo	41	Elqui	435,302	JFM	2010
Coquimbo	43	Limarí	145,175	JFM	2010
Valparaíso	51	Valparaíso	665,584	JFM	2010
Valparaíso	56	San Antonio	145,215	JFM	2010
Valparaíso	57	San Felipe	131,504	JFM	2010
Valparaíso	58	Marga Marga	302,039	JFM	2010
Metropolitan	131	Santiago	4,845,166	JFM	2010
Metropolitan	132	Cordillera	539,844	JFM	2010
Metropolitan	134	Maipo	428,787	JFM	2010
Metropolitan	135	Melipilla	167,073	JFM	2010
Metropolitan	136	Talagante	261,044	JFM	2010
O'Higgins	61	Cachapoal	556,719	JFM	2010
Maule	71	Talca	357,162	JFM	2010
Maule	73	Curicó	249,679	JFM	2010
Maule	74	Linares	245,335	JFM	2010
Ñuble	161	Diguillín	275,163	ASO	2010
Biobío	81	Concepción	864,48	JFM	2010
Biobío	82	Arauco	138,379	JFM	2010
Biobío	83	Biobío	335,921	JFM	2010
La Araucanía	91	Cautín	635,515	JFM	2010
Los Ríos	141	Valdivia	248,062	JFM	2010
Los Lagos	101	Llanquihue	350,096	JFM	2010
Los Lagos	103	Osorno	200,285	JFM	2010
Aysén	111	Coyhaique	48,022	JFM	2010
Magallanes	121	Magallanes	115,869	JFM	2010

Source: Produced by the National Statistics Institute

Annex 7 Calculation of expansion factors in the transition period

As noted in section 5.8, between January and May 2020, the ENE sample will transition towards two independent samples: the previous sample (where the sections come from the 2016 SFS and the blocks from the 2016 SFB) and the new sample, which uses the 2017 SFD.

For nomenclature purposes we will call frame “A” the combination of the 2016 SFS and 2016 SFB and frame “B” the 2017 SFD.

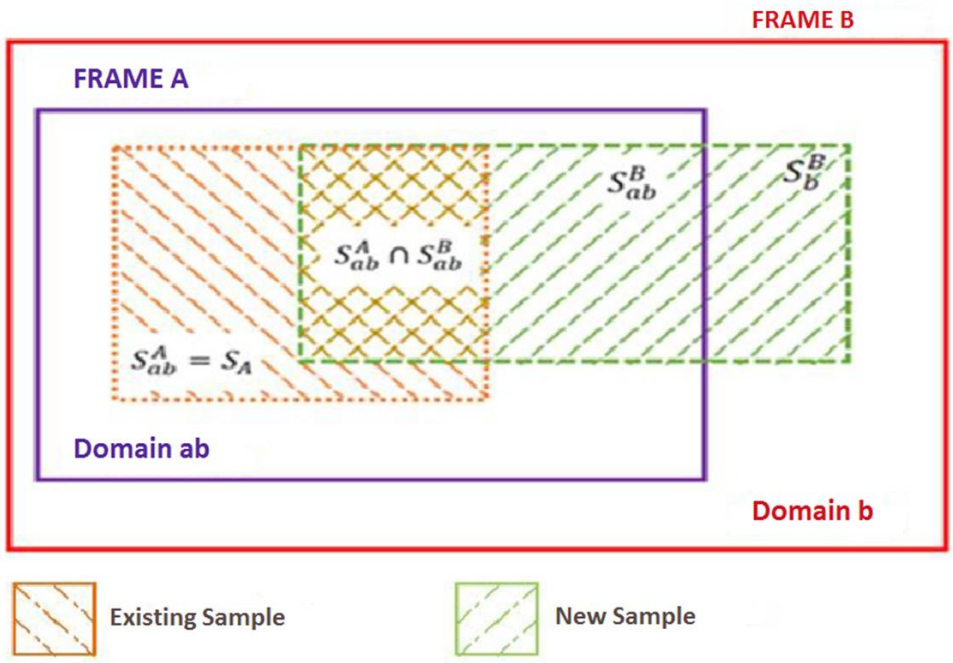
In the first quarter of the transition, approximately five-sixths of the dwellings from the previous sample will be preserved, and one-sixth of the new sample will be incorporated. As a result, the calculated design weights of the new sample from frame B will be larger than the full sample

weights. Therefore, if we were to combine the old sample and the new sample, we would obtain a figure of approximately twice the population[68].

The simultaneous use of several samples drawn from different frames to estimate the target population in a domain or stratum is called multi-frame estimation. One or several frames used in multi-frame estimation may have incomplete coverage (i.e., some domains may only be covered by one frame) (Hartley, 1962). However, the challenge of multiple-frame estimates arises in domains where the frames overlap.

As shown in Figure 12, Frame A is completely contained in Frame B (i.e., it has no units that are not covered by Frame B), while Frame B does have units that are not covered by Frame A (e.g., units in the ADA exclusion areas that have now been incorporated into Frame B).

Figure 12 Diagram of frame and sample configuration



Source: Produced by the National Statistics Institute

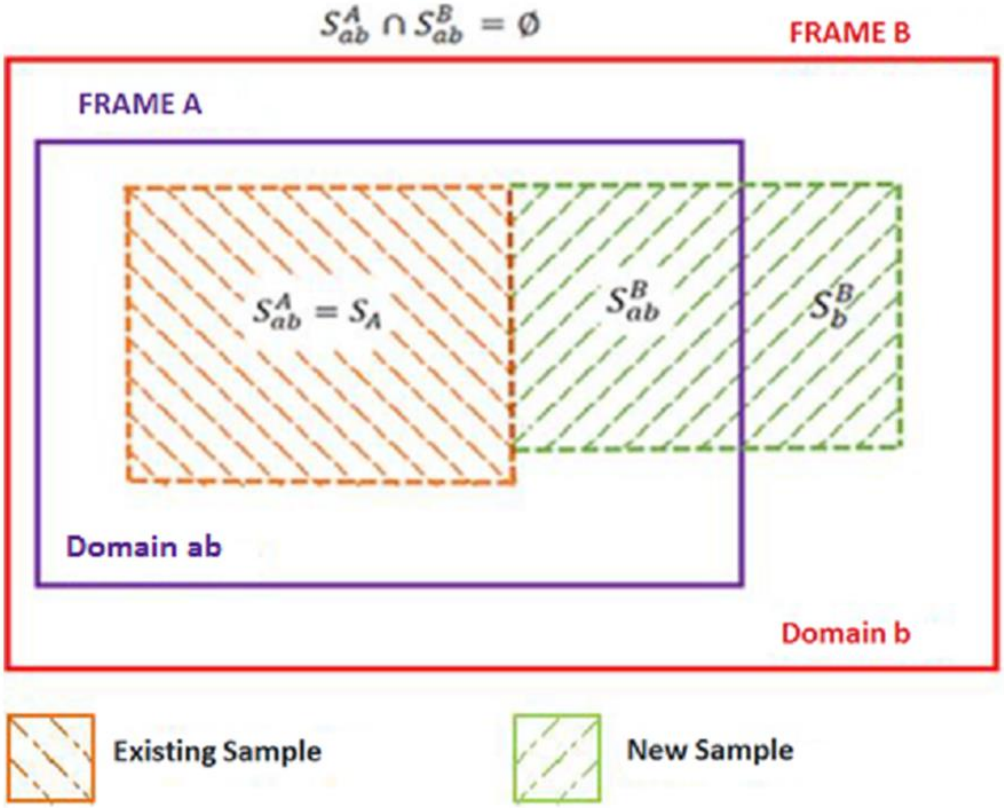
If independent samples are taken in each stratum, then S_A is the sample taken from frame A (previous sample), which can be broken down into $S_A = S_a^A \cup S_{ab}^A = S_{ab}^A, where\ S_a^A = \emptyset$ because it does not have frame A domains that are not covered by frame B. The sample taken from frame B is called $S_B = S_b^B \cup S_{ab}^B$, where S_b^B is the sample taken from frame B that corresponds to domain b or domains that are not covered in frame A because they are excluded (ADA).

[68] This assumes that there is no lack of response and that coverage is sufficient.

In the common domain (ab), samples $S_{ab}^{A} \cap S_{ab}^{B}$ can be intersected because the selection of dwellings is independent for each frame and because these units partially or completely overlap. To avoid the repeating measurement of a dwelling, the sample from frame A (current) was already selected, and the dwelling is not measured for the sample from frame B, whose sampling loss is expected to be offset by oversampling[69].

In this way, the set of duplicate dwellings resulting from independent selection would be empty $S_{ab}^{A} \cap S_{ab}^{B} = \emptyset$, and the diagram in Figure 12 is transformed into the diagram in Figure 13.

Figure 13 Updated diagram of frame and sample configuration



Source: Produced by the National Statistics Institute

Because sample units from different frames S_{ab}^{A} and S_{ab}^{B} can generate estimates, it is in the domain where the expansion factors need to be adjusted.

As the weights of dwelling selection indicate the population of dwellings of each frame in the common domain (i.e., M_{ab}^{A} in the case of frame A and M_{ab}^{B} in the case of frame B), and if $M_{ab}^{A} \approx$

[69] According to the forecasts to determine the fraction of the previous sample to be eliminated in each increment, the cases where a double measurement occurs are expected to be minimal.

M_{ab}^{B}, the sum of the dwelling factors of both samples[70] approximately doubles the population of dwellings of one frame or the other.

To remedy this situation, the weights for dwelling selection (adjusted for non-response) are adjusted. The adjustment depends on the proportion of each of the samples (old and new) in the total sample in each quarter of the transition. This can be expressed as follows[71]:

$$\boldsymbol{w_{Lij}^{MM}} = \begin{cases} \boldsymbol{W_{h_A ij}^{NR} \cdot \alpha, if\ the\ sample\ is\ from\ frame\ A\ (previous\ sample)} \\ \boldsymbol{W_{hBij}^{NR} \cdot (1-\alpha), if\ the\ sample\ is\ from\ frame\ B\ (new\ sample)} \end{cases} \quad \boldsymbol{(38)}$$

$$in\ which\ \alpha = \frac{m_A}{m_A + m_B}\ and\ L = \{h_A\ and\ h_B \in of\ common\ domain\ ab\}$$

Where,

w_{Lij}^{MM}: Selection weight of dwelling j in cluster i in common domain L, adjusted for multiple frames (MF)

$W_{h_A ij}^{NR}$: Selection weight of dwelling j in cluster i in stratum h_A, adjusted for non-response of the previous sample

W_{hBij}^{NR}: Selection weight of dwelling j in cluster i in stratum h_B, adjusted for non-response of the previous sample

m_A: Number of dwellings obtained from the previous sample

m_B: Number of dwellings obtained from the new sample

h_A , h_B: h_B: strata defined in frame B (new); h_A: strata defined in frame A (current)

It should be noted that the selection weights of the previous sample are calculated according to what was defined in the previous sampling design because, during the first three mobile quarters of the transition period, the inputs for making adjustments for eligibility of the entire sample from frame A were not available. This situation can also be attributed to the lack of oversampling[72] and to the new classification of final disposition codes, which came into effect in January 2020.

[70] If $M_{ab}^{A} = M_{ab}^{B}$ in the first incorporation of the new sample (or the first elimination of the previous sample), the sum of the factors of dwellings of both samples (five-sixths of the previous sample and one-sixth of the new sample) would exactly double the population of dwellings of the Frame (A or B).

[71] This methodology was implemented for the New Zealand Labour Force Survey (Statistics New Zealand, 2015) with the weight at the level of person. It was shown that the expression minimizes the variance for the estimation of the totals.

[72] The previous sampling design used replacement of dwellings, not oversampling. In addition, the classification of visitation codes used in the Household Registration Card (HRT) did not follow international guidelines. Instead, it was based on an *ad-hoc* classification of groups, using categories of attainment and non-attainment.

Therefore, the theoretical expansion factor associated with the cluster in stratum h_A of the previous ENE sample is calculated by the following formula:

$$w_{h_A i} = \frac{M_{h_A}}{n_{h_A}^{sel} \cdot M_{h_A i}} \cdot \frac{M'_{h_A i}}{m_{h_A i}^{sel}} \quad (39)$$

Where,

M_{h_A}: Number of private dwellings in stratum h_A according to the appropriate census[73] or to the year in which the sampling frame was updated

$M'_{h_A i}$: Number of occupied private dwellings according to updated lists of dwellings, in the unit of the first stage i in stratum h_A

$M_{h_A i}$: Number of private dwellings in the unit of the first stage i in stratum h_A, according to the appropriate census or to the year in which the frame was updated

$n_{h_A}^{sel}$: Number of selected units of the first stage i in stratum h_A

$m_{h_A i}^{sel}$: Number of selected dwellings in the unit of the first stage i in stratum h_A

To account for the quality of the field survey and to adjust the expansion factor to the response actually obtained on the ground, the survey is adjusted for non-response of the clusters and the number of dwellings for which a response was attained. Thus, the following two adjustment weights are calculated:

1. Weights for clusters (or first-stage units): is the ratio between the number of clusters selected and the number of clusters actually attained.

$$Cluster\ weight = \frac{n_{h_A}^{sel}}{n_{h_A}} \quad (40)$$

$n_{h_A}^{sel}$: Number of selected units of the first stage i in stratum h_A

n_{h_A}: Number of responses attained from selected units of the first stage i in stratum h_A

2. Weight for dwellings: is the ratio between the number of selected dwellings and the number of dwelling from which a response was attained.

[73] It should be remembered that during the period of incorporation of the new sample and the elimination of the previous sample, the total sample comes from distinct frames of different censuses, which have distinct first-stage units and a different stratification.

$$Dwelling\ weight = \frac{m_{h_A i}^{sel}}{m_{h_A i}} \quad (41)$$

Where,

$m_{h_A i}^{sel}$: Number of selected dwellings in the unit of the first stage i in stratum h_A

$m_{h_A i}$: Number of responses obtained from selected dwellings in the unit of the first stage i in stratum h_A

Thus, the expansion factor is adjusted for non-response using the following formula:

$$w_{h_A ij} = w_{h_A i} \cdot \frac{n_{h_A}^{sel}}{n_{h_A}} \cdot \frac{m_{h_A i}^{sel}}{m_{h_A i}} = \frac{M_{h_A}}{n_{h_A}^{sel} \cdot M_{h_A i}} \cdot \frac{M'_{h_A i}}{m_{h_A i}^{sel}} \cdot \frac{n_{h_A}^{sel}}{n_{h_A}} \cdot \frac{m_{h_A i}^{sel}}{m_{h_A i}} = \frac{M_{h_A}}{n_{h_A} \cdot M_{h_A i}} \cdot \frac{M'_{h_A i}}{m_{h_A i}} \quad (42)$$

Where,

M_{h_A}: Number of private dwellings in stratum h_A according to the appropriate census or to the year in which the frame was updated

$M'_{h_A i}$: Number of occupied private dwellings, according to updated lists of dwellings, in the unit of the first stage i in the stratum h_A

$M_{h_A i}$: Number of private dwellings in the unit of the first stage i in the stratum h, according to the appropriate census or to the year in which the frame was updated

$n_{h_A}^{sel}$: Number of units of the first stage selected from stratum h_A

$m_{h_A i}^{sel}$: Number of dwellings selected in the unit of the first stage i in stratum h_A

n_{h_A}: Number of responses obtained from selected units of the first stage i in stratum h_A

$m_{h_A i}$: Number of responses obtained from selected dwellings in the unit of the first stage i in stratum h_A

As can been seen, the base expansion factor is a function of the number of clusters and dwellings attained $(n_{h_A}; m_{h_A i})$ and not those that were initially selected $(n_{h_A}^{sel};\ m_{h_A i}^{sel})$. This factor implicitly includes the two adjustments mentioned above, and it is analogous to the non-response adjustment applied to the factors in Framework B (see chapter 6). Then,

$$w_{h_A ij} = w_{h_A ij}^{NR} \quad (43)$$

Last, the calibrated weight for population projections $(w_{h_A ij}^{cal})$ during the transition period can be obtained from weight w_{Lij}^{MM}(selection weight adjusted for multiple frames).

Annex 8 Development of tools for the calibration of the foreign population in expansion factors of the ENE

The arrival of foreign-born persons to Chile has increased steadily, from 0.8% of the population in 1992, to 1.3% in the 2002 Census, and 4.4% in the 2017 census. This last census shows that 66.7% of the population of international immigrants arrived in the country in the period 2010–2017, mainly between 2016 and 2017 (INE, 2018c). According to an estimate of foreign residents in Chile (INE-DEM, 2019), the percentage of the foreign population in relation to the country's population was 6.6[74] as of December 31, 2018, and this percentage is expected to increase in 2019.

As the foreign-born population is younger than the population born in Chile, the census shows that it made up 4.9% of the working-age population and 6.5% of the primary labour force (INE, 2018c). In 2018, these figures were 7.5% working-age population and 10.1% of the primary labour force[75].

In addition to being a younger population, persons born abroad have a different relationship to work than persons born in Chile. Evidence of this is that, in the 2017 Census, 74.1% of the foreign-born population declared that they worked during the week before the census, but for persons born in Chile, the percentage was 56.9%[76] (INE, 2018c).

Immigration has had and will continue to have an important impact on the various indicators measured household surveys, particularly those of the labour market and of the number of foreign people residing in Chile. A monthly estimate of the foreign population residing in Chile was deemed a necessary benchmark of exogenous information for the adjustment or calibration of expansion factors in INE's household surveys. Particularly, the estimate can be used as input data for the new calibration methodology of expansion factors of the ENE.

To achieve this objective, the Subdepartment of Statistical Research, with the collaboration of the Subdepartment of Demography, has developed a model for calculating the monthly variation of the foreign population that can be used in the calibration of the expansion factors of the ENE.

[74] Strictly speaking, because the relationship between the two populations is not a proportion but a ratio, the percentage indicated is not a percentage as such, because population estimates and projections only distinguish the sex and age of the population, not nationality or any other variable.

[75] Strictly speaking, the percentages for 2018 are a ratio whose numerator is the estimate as of 31 December 2018 and whose denominator is the estimate and projection as of 30 June 2018.

[76] There is a difference by sex. For those born in Chile, the percentage is 69.0% for men and 45.6% for women, while those born in another country the percentage is 82.5% for men and 66% for women.

Methodology for obtaining monthly foreign population figures

In Chile, the most widely used source of information to measure international migration has been the population and housing censuses, which consider habitual residents to be persons who have resided or intend to reside six months or more in the country. However, administrative records (in which theoretical residence is related to the duration of the visa (generally for periods of one year)) and specialized demographic surveys are both recognized as sources for demographic analysis. Each source uses different units of analysis and methodologies, so readers must be aware of these differences when they are compared. In addition to the differences in habitual residence, the censuses enable the identification of persons born abroad (nationals or foreigners), while the administrative registers can identify the foreign population (born in the country or abroad).

Currently, Chile does not have any administrative record that can provide a monthly report on the total population of foreign persons residing in Chile. This information, as mentioned, could be used to adjust or "calibrate" the expansion factors of the ENE and, thus, incorporate the effect of the increase in the foreign population within the total resident population. As a complementary measure, we sought to define a model for finding the total number of the foreign population in order to obtain a monthly estimate of this subpopulation. There are, however, not enough observations of the population, because direct information on the population of foreign persons is only available through the 2017 Census. Furthermore, a corresponding adjustment for census omission must be included, and the INE-DEM estimate[77] of 31 December 2018 (INE-DEM, 2019) must be considered. Because the quantity of data is evidently insufficient, the information provided by population estimates and projections had to be adapted as a complementary data measure.

Population estimates and projections use an estimate of international migration flows as one of its inputs. It must be remembered that this flow of immigrants (both estimates and projections) does not differentiate the Chilean population from the foreign population. That is, among both immigrants and emigrants, there are Chileans and foreigners, so it is necessary to emphasize that "foreigners" and "immigrants" are two different populations. However, in the methodology presented here, the total number of immigrants is considered an approximation of the total number of foreigners. Thus, in the design of the methodology proposed here, the flow of foreigners entering Chile will be approximated by the estimated flow of immigration, which is used by the population estimates and projections based on the year 2017.

To make a monthly approximation of the total population of foreign persons residing in Chile in the period 2012–2025[78], the annual population of foreign persons was first determined. For this purpose, an initial population was determined for June 2012 based on the sum of the flow of

[77] Department of Immigration (Departamento de Extranjería y Migración (DEM)).

[78] Having a monthly estimate after 2025 is not relevant for two reasons. First, in 2022, a new census will be conducted, and therefore in 2025 there will be new population projections with accompanying data on the migration balance. Second, by incorporating information far from the time of the Census (2017), precision is lost. Thus, by incorporating figures after 2025, the problem of under-adjustment of the model arises.

immigrants between the years 1992–2012 (see Table 15), using population estimates and projections. Then, the populations for June of each successive year until 2025 were calculated by adding the corresponding annual flows of immigrants from the population estimates and projections to the initial population.

Table 16 Estimate of the total foreign population (as of 30 June each year) from migration balance of the population projections of base year 2017

Year	1992	1993	1994	1995	1996	1997	1998
Immigrant flow	1,646	2,081	2,523	4,895	5,588	6,199	6,629
Foreign pop.	-	3,727	6,250	11,145	16,733	22,932	29,561
Year	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Immigrant flow	6,929	10,903	11,393	12,380	17,915	19,354	22,996
Foreign pop.	36,490	47,393	58,786	71,166	89,081	108,435	131,431
Year	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Immigrant flow	25,603	29,786	28,506	28,385	43,303	48,449	32,633
Foreign pop.	157,034	186,820	215,326	243,711	287,014	335,463	368,096
Year	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**
Immigrant flow	40,546	46,277	65,701	136,529	224,195	246,861	239,656
Foreign pop.	408,642	454,919	520,620	657,149	881,344	1,128,205	1,367,861
Year	**2020**	**2021**	**2022**	**2023**	**2024**	**2025**	
Immigrant flow	107,936	39,868	25,022	21,796	20,936	20,658	
Foreign pop.	1,475,797	1,515,665	1,540,687	1,562,483	1,583,419	1,604,077	

Source: Own elaboration based on the immigration component of the Population Projections base 2017

Once the total population of foreign persons as of June was estimated for each of the years between 2012 and 2025, a nonlinear model was adjusted. This model—a logistic function of four parameters (hereinafter, the logistic model)—estimated monthly totals, which are an approximation for the total number of foreign persons resident in Chile[79].

Results of the model

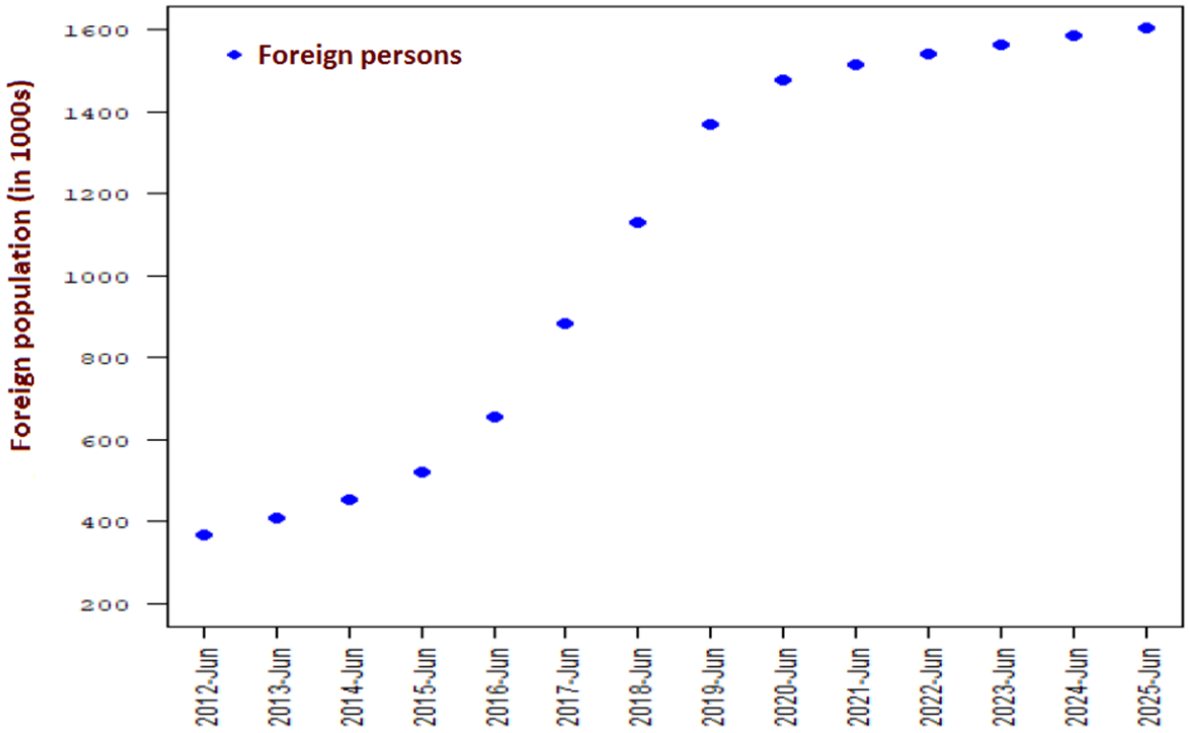
This section presents the results of the logistic model. The technical detail can be found in the final section of this annex.

[79] Because they refer to people, the estimates generated by the model are rounded to the first whole number, without decimals, by means of the function $round(\,,\ digits = 0)$.

In the data described above and in Graph 2 below, two particularities can be seen:

1. The functional form of x is nonlinear.
2. The functional form of x takes the form of a logistic function.

Graph 2 Annual estimate of the foreign population based on the migration balance of the population projections of base year 2017

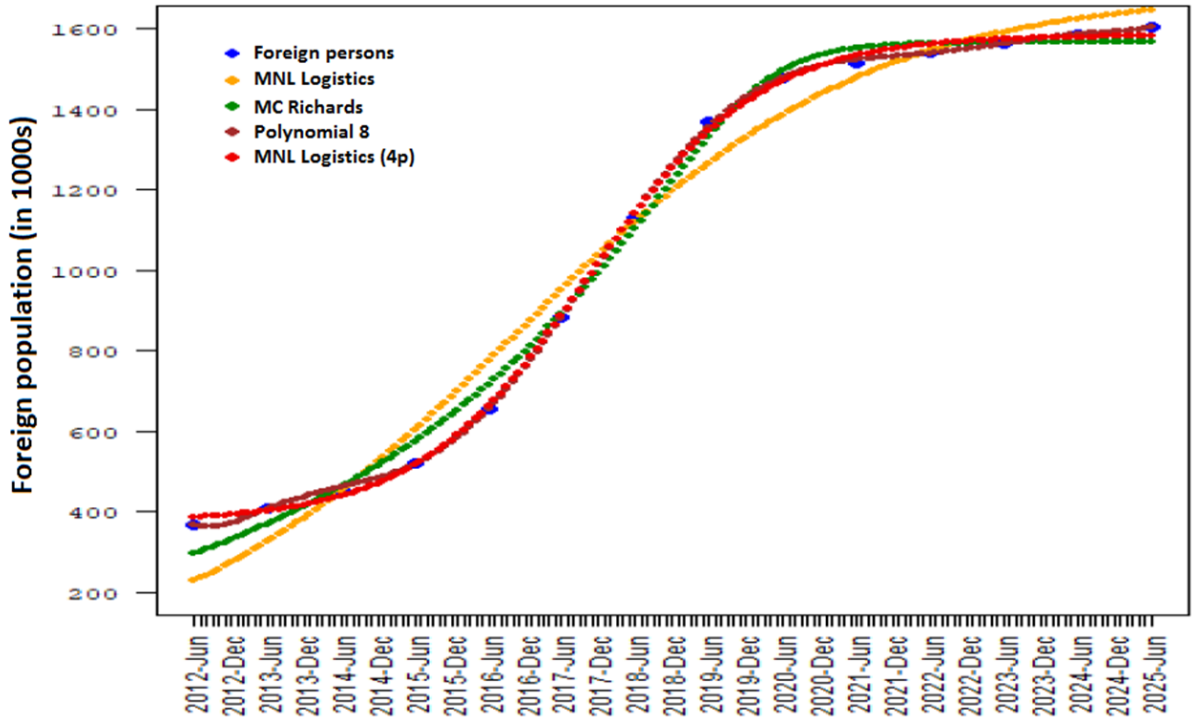


Source: Own elaboration based on the immigration component of the population projections based on the year 2017

In the estimation process, different measurements were tested for the mean function to be estimated. Specifically, three mean functions were adjusted, all corresponding to a logistic form: the Richards growth curve, the three-parameter and four-parameter logistic function, and an eight-degree polynomial model[80]. Graph 3 summarizes these results

[80] Because of asymptotes in the path followed by the data, the exponential growth curve was discarded. Other models such as the Gompertz growth model did not achieve convergence for their numerical solution.

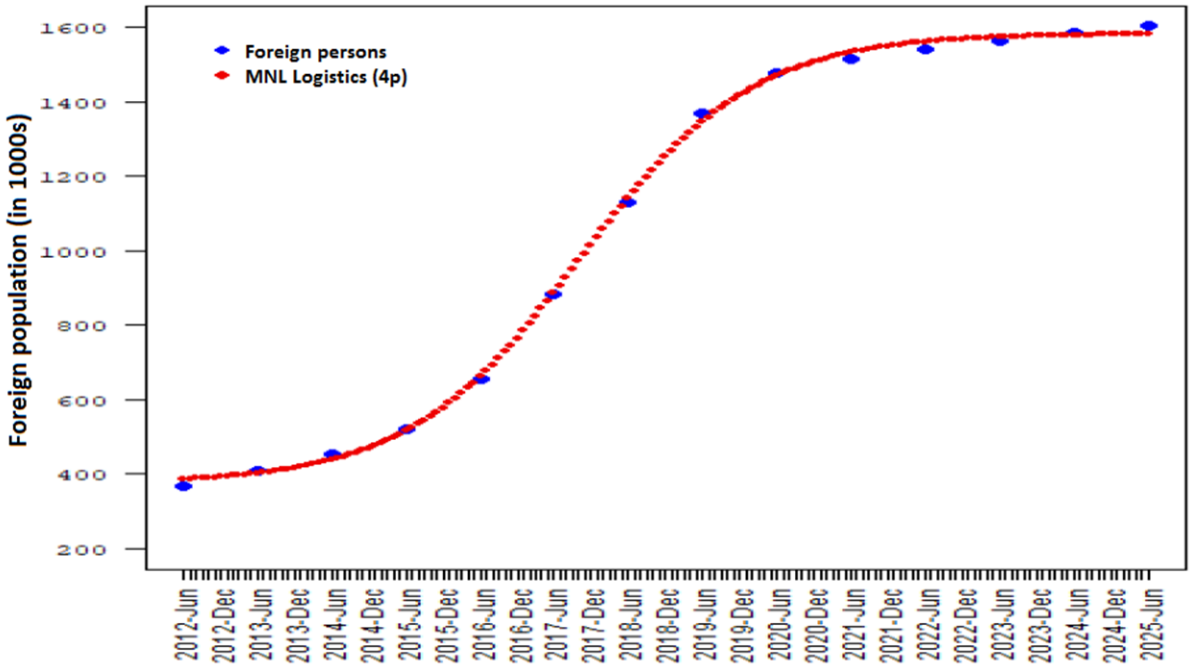
Graph 3 Annual estimate of foreign population based on the migration balance of population projections of base year 2017 and on the estimates of selected models



Source: Own elaboration based on the immigration component of the Population Projections base 2017 and on the models under analysis

From the tables, the sum of the differences of the estimate and of the annual estimates were evaluated by contrasting them with figures for the population born abroad at the time of the 2017 Census (746,465 persons) and with the official estimate of foreign residents for 31 December 2018 of INE and the Department of Immigration (1,251,225 persons). Based upon this evaluation and considering the parsimony that a logistic model provides in this type of estimate, the four-parameter logistic model was chosen as the final model. Graph 4 illustrates this series.

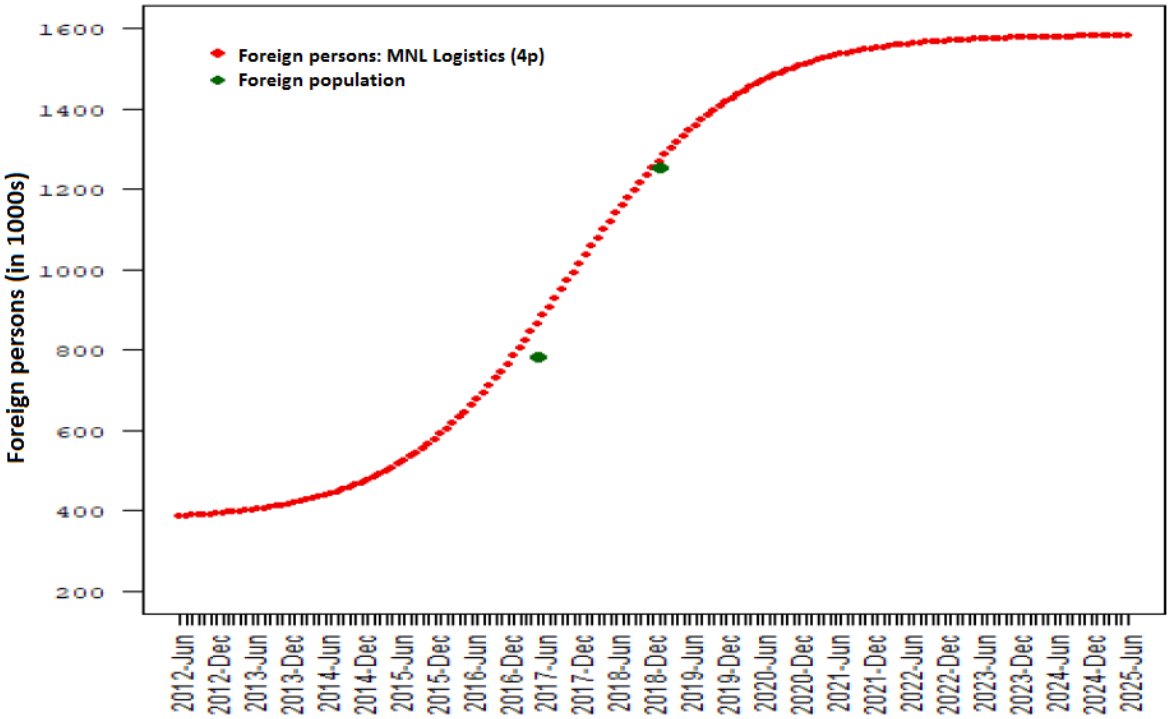
Graph 4 Monthly estimate of the foreign population according to the logistic model and annual estimate based on the migration balance



Source: Own elaboration based on the Immigration component of the population projections based on the year 2017 and on the estimate of the four-parameter nonlinear logistic model

Graph 5 contrasts the results of the selected model and the population of the foreign population resident in Chile reported in the 2017 Census by incorporating the census omission and the foreign population estimate as of 31 December 2018.

Graph 5 Estimate of foreign population according to the 2017 Census logistic model and the estimate of resident foreign population as of 31 December 2018



Source: Own elaboration based on the 2017 Census (with census omission), the foreign population estimate on 31 December, and the estimation of the nonlinear logistic model of four parameters

Technical scopes, current and prospective uses of the estimate

The methodology proposed in this document seeks to improve the calibration of household surveys by accounting for the impact that international migration has on social phenomena. Its use is oriented towards the employment survey and does not seek to replace the official annual estimates provided by INE, the Department of Immigration (DEM), and other State services, whose conceptual and methodological construction are different. Because of changes in migration policies, the annual estimates include updated information on the phenomenon. It should be noted that, from a demographic point of view, persons who immigrate might also leave the country or die, something that the model does not consider. Thus, the estimates derived from the model will be monitored continuously and, its results can be updated with new information included in the model.

Secondly, incorporating exogenous information on the foreign population does not mean that specific estimates regarding this subpopulation, which are based on data of the ENE, have per se the appropriate statistical quality required for the rest of the estimates that are derived from the survey. Statistical quality must be verified by observing precision measures (standard errors, coefficients of variation, and/or confidence intervals), the degrees of freedom within the sub-

population, and the number of sample observations with which the indicators are estimated for this sub-population.

Technical specifications of the model

The model selected is a nonlinear model whose mean function is a logistic function of four parameters, as follows[81]:

$$f(x, A, B, xmid, scal) = A + \frac{B - A}{1 + e^{(xmid - x)/scal}} \qquad (44)$$

Where,

x: Numerical vector of values where the model is evaluated

A: The horizontal asymptote (left) is represented by this parameter

B: The horizontal asymptote (right) is represented by this parameter

$xmid$: Parameter that represents the value of x at the inflection point of the curve

$scal$: Numerical scale parameter on the axis of variable x

The estimation of the four parameters of the mean function, described above, uses R software from the package *nls()* and the following routine:

```
library(foreign)
library(NISTnls)
library(proto)
library(nls2)
library(nlrwr)
library(NRAIA)
library(nlstools)
library(nlsMicrobio)

rm(list = objects())
setwd("C://Estimacion_Extranjeros//Estimacion_OFICIAL")
inm2012 <- read.csv(file = "Inmigrantes2012.csv", header = TRUE, dec = ",")
attach(inm2012)
set.seed(5782)
NLlog4_2012 <- nls(immigration ~ SSfpl(t, A, B, xmid, scal), data = inm2012)
summary(NLlog4_2012)
```

The results described in Figure 14 show that the four estimated parameters present coefficients with a level of significance of 1%. Therefore, the null hypothesis of zero coefficients is rejected (i.e., the estimated parameters are statistically significant).

[81] For more detail on nonlinear model estimation, see Bates & Watts, *Nonlinear regression analysis and its applications* (1988), Bates & Chambers, *Nonlinear models* (1992), and Crawley (2013).

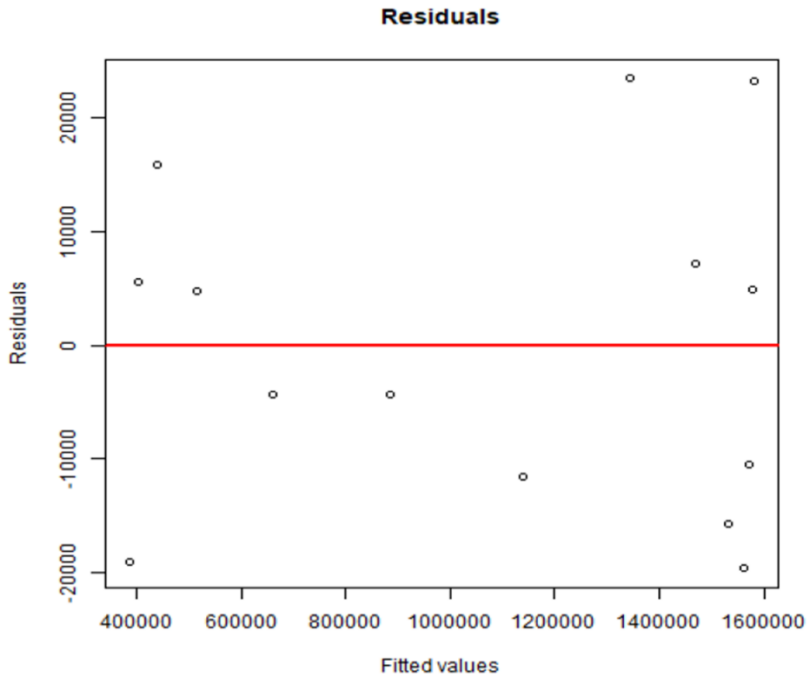
Figure 14 Results of the nonlinear logistic model of four parameters

```
> summary(NLlog4_2012)

Formula: inmigracion ~ SSfpl(t, A, B, xmid, scal)

Parameters:
      Estimate Std. Error t value Pr(>|t|)
A    3.751e+05  1.261e+04   29.75 4.31e-11 ***
B    1.583e+06  8.661e+03  182.72  < 2e-16 ***
xmid 6.435e+01  6.348e-01  101.37  < 2e-16 ***
scal 1.400e+01  5.801e-01   24.13 3.41e-10 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16580 on 10 degrees of freedom

Number of iterations to convergence: 0
Achieved convergence tolerance: 2.741e-06
```

Source: Own elaboration based on the immigration component of the population projections, base 2017

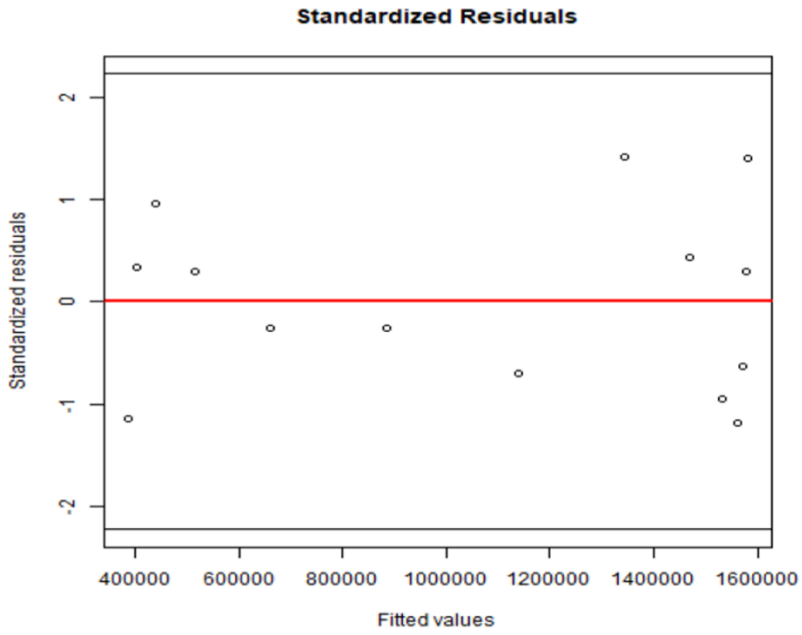
Finally, as Ritz & Streibig (2008) observe, the assumptions behind the nonlinear regression model must be validated. Thus, it must be verified that the mean function is correct and that homoscedasticity (with the standardized residues), normal distribution of the residues, and independence of the residues are satisfied in the adjusted model.

Graph 6 Assumptions of the nonlinear regression model: correct mean function



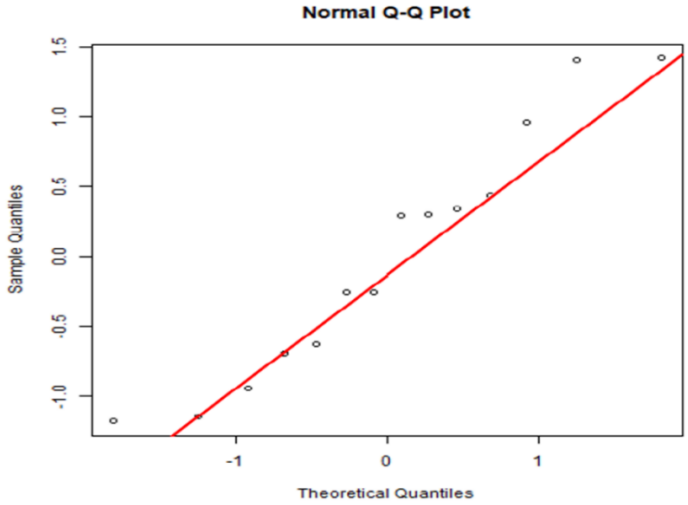
Source: Own preparation based on the results of the adjusted nonlinear model

Graph 7 Assumptions of the nonlinear regression model: homoscedasticity



Source: Own preparation based on the results of the adjusted nonlinear model

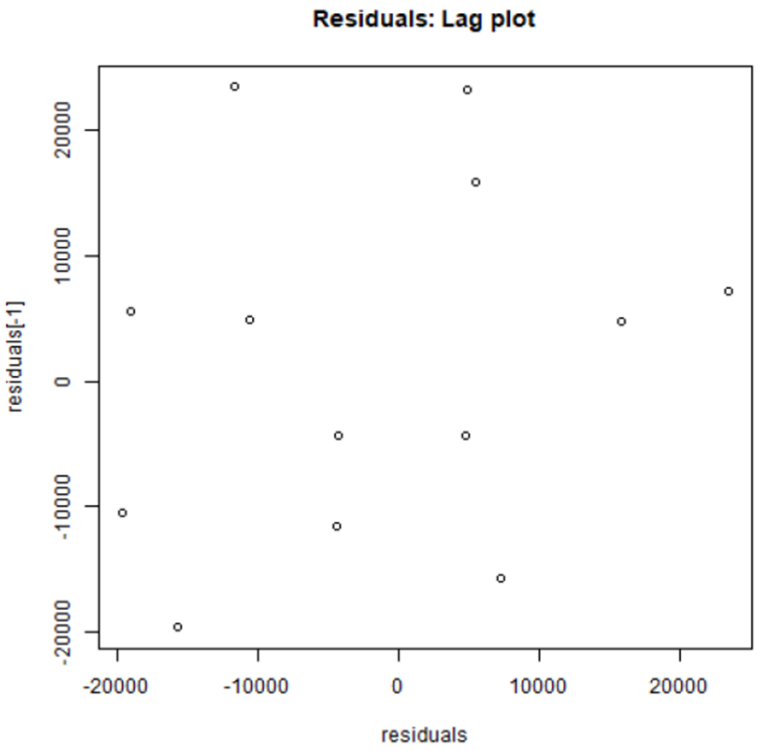
Graph 8 Assumptions of the nonlinear regression model: normality




```
> shapiro.test(standardRES)

        Shapiro-Wilk normality test

data:  standardRES
W = 0.93438, p-value = 0.3511
```

Source: Own preparation based on the results of the adjusted nonlinear model

Graph 9 Assumptions of the nonlinear regression model: independence



Source: Own preparation based on the results of the adjusted nonlinear model

11 Bibliography

13° CIET. (1982). Retrieved from https://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/resolutions-adopted-by-international-conferences-of-labour-statisticians/WCMS_087483/lang--es/index.htm

19° CIET. (2013). Retrieved from https://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/resolutions-adopted-by-international-conferences-of-labour-statisticians/WCMS_234036/lang--es/index.htm

20°CIET. (febrero de 2018). *Resolución sobre las estadísticas de las relaciones de trabajo*. Retrieved from https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms_619087.pdf

AAPOR. (2016). *Standard Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys*. Retrieved from https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

Antía, M., & Coimbra, A. (2009). *Tratamiento de la no respuesta en encuestas de panel en el caso de poblaciones finitas: "Las damas perdidas"*. Universidad de la República, Uruguay.

Bankier, M. D., (1986). *Estimators Based on Several Stratified Samples, with applications to Multiple Frame Surveys*. Journal of American Statistical Association, 81, 1074-1079.

Bates, D., & Chambers, J. (1992). *Nonlinear models*. Wadswoth & Brooks/Cole.

Bates, D., & Watts, D. (1988). *Nonlinear regression analysis and its applications*. Wiley & Sons.

BCCh. (2001). *Indicadores Económicos y Sociales en Chile 1960 – 2000*. Santiago, Chile: Departamento Publicaciones de la Gerencia de Investigación Económica, División Estudios, Banco Central de Chile.

Binder, D.A., & Theberge, A. (1988). *Estimating the Variance of Raking Ratio Estimators*. Canadian Journal of Statistics, 16 (supplement), 47-55.

Brackstone, G.J., & Rao, J.N. (1979). *An Investigation or Raking Ratio Estimators*. Sankhya C, 41, 97-114.

Bravo, D. (2010). *Encuestas Longitudinales en Chile*. Santiago, Chile.

Cecchini, S. (2005). *Indicadores Sociales en América Latina*. Santiago, Chile: División de Estadística y Proyecciones Económicas, CEPAL.

Choudhry, G.H., & Lee, H. (1987). *Variance Estimation for the Canadian Labour Force Survey*. Survey Methodology, 13, 147-161.

Cochran, W. (1998). *Técnicas de Muestreo*. México DF: Compañía Editorial Continental, S.A. de C.V.

Crawley, M, (2013). *The R book*. Chichester: Wiley & Sons.

Department of Commerce, Bureau of the Census. (1971). *Curso Suplementario Para Un Estudio de Caso Sobre Encuestas Y Censos: Conferencias Sobre Muestreo*.

Deville J., S.C. (1993). *Generalized Raking Procedures in Survey Sampling*.

Deville, J.C., & Särndal, C.E. (1992). *Calibration Estimators in Survey Sampling*. Journal of the American Statistical Association, Vol. 87 (418): 376-382.

Dobson, A. (2002). *An Introduction to Generalized Linear Models*. CRC Press.

Glejberman, D. (2015). *Taller Regional. Las Encuestas de Hogares en América Latina: Estado de situación y prospectiva. Usos de las Encuestas Longitudinales*. Santiago de Chile.

Groves, R. (2004). *Survey Errors and Survey Cost*. New York: John Wiley & Sons.

Hartley, H.O. (1962). *Multiple Frame Surveys*. Proceedings of the American Statistical Association, Social Statistics Sections.

Heeringa, S.G., West, B.T., & Berglund, P.A. (2010). *Applied Survey Data Analysis*. Boca Raton, FL: Taylor and Francis Group, LLC.

Heeringa, S.G., West, B.T., & Berglund, P.A. (2010). *Applied Survey Data Analysis*. USA: Chapman & Hall / CRC Press.

Hidiroglou, M. (1986). The Construction of a Self-Representing Stratum of Large Units in Survey Design. The American Statistician, 40(1), 27-31.

IBGE. (2004). *Semantic Scholar*. Retrieved on 15 January 2020 from https://pdfs.semanticscholar.org/21bb/a469c8b217f276755a827006d31b4255bdc5,pdf?_ga=2.132283335.871766543.1578866611-143326825.1578866611

IESTA. (s.f.). *IESTA*. Retrieved on 12 December 2019 from http://www.iesta.edu,uy/wp-content/uploads/2014/05/TJA2010_ant%C3%ADa-coimbra.pdf

INE. (Abril de 2016a). *CAENES, Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas*. Retrieved from

https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/documento-clasificador-de-rama-de-actividad-econ%C3%B3mica-caenes-(abril-2016),pdf?sfvrsn=1b1cc016_3

INE. (Abril de 2016b). *Separata Técnica, Clasificador de Actividades Económicas Nacional para Encuestas Sociodemográficas (CAENES) en la Encuesta Nacional de Empleo*. Retrieved from

https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/publicaciones-y-anuarios/separatas/tem%C3%A1ticas/separata-tecnica-caenes-empleo-abril-2016,pdf?sfvrsn=25455d76_3

INE. (2018a). *CIUO 08.CL, Clasificador Chileno de Ocupaciones*. Retrieved from http://webanterior.ine.cl/docs/default-source/publicaciones/2018/ciuo-08,cl-clasificador-chileno-de-ocupaciones,pdf?sfvrsn=50bd58d2_4

INE. (Febrero de 2018b). *Estadísticas de informalidad laboral: Marco conceptual y manual metodológico*. Retrieved from https://www.ine.cl/docs/default-source/informalidad-y-condiciones-laborales/metodologia/antecedentes-metodologicos/manual-conceptual-y-metodol%C3%B3gico-informalidad-laboral.pdf?sfvrsn=afad6bfc_5

INE. (Noviembre de 2018c). *Características de la imigración internacional en Chile, Censo 2017*. Retrieved from http://www.censo2017.cl/descargas/inmigracion/181123-documento-migracion.pdf

INE. (Marzo de 2019a). *Ciudades, pueblos, aldeas y caseríos 2019*. Retrieved from https://geoarchivos.ine.cl/File/pub/Cd_Pb_Al_Cs_2019.pdf

INE. (Junio de 2019b). *Unidades primarias de muestreo, Nuevo marco muestral de viviendas, Documento interno*. Instituto Nacional de Estadísticas (INE).

INE. (Octubre de 2019c). *Estratificación socioeconómica del marco muestral de viviendas 2017 (MMV 2017). Documento de Trabajo, N°11*. Retrieved from

https://www.ine.cl/docs/default-source/documentos-de-trabajo/estratificaci%C3%B3n-socioecon%C3%B3mica-del-marco-muestral-de-viviendas-2017.pdf?sfvrsn=e4482f3a_2

INE. (Diciembre de 2019d). *Fundamentos de la nueva metodología de calibración de los factores de expansión de la Encuesta Nacional de Empleo, Documento de Trabajo, N° 12*. Retrieved from

https://www.ine.cl/inicio/documentos-de-trabajo/documento/documento-de-trabajo-fundamentos-de-la-nueva-metodolog%C3%ADa-de-calibraci%C3%B3n-de-los-factores-de-expansi%C3%B3n-de-la-encuesta-nacional-de-empleo

INE. (marzo de 2020). *Lineamientos para el Uso de Medidas de Precisión en Encuestas de Hogares, Documento de Trabajo (próximo a publicar)*. Retrieved from https://www.ine.cl/inicio/documentos-de-trabajo

INE-DEM. (febrero de 2019). *Estimaciones de Personas Extranjeras residentes en Chile, al 31 de diciembre 2018*. Retrieved from https://www.extranjeria.gob.cl/media/2019/07/Estimaci%C3%B3n-Poblaci%C3%B3n-Extranjera-en-Chile.pdf

Kish, L. (1963). *Changing strata and selection probabilities, Proceedings of the Social Statistics Section*, *American Statistical Association*, 124-131.

Konijin, H.S. (1981). *Biases, Variances, and Covariances of Raking Ratio Estimators for Marginal and Cell Totals and Averages of Observed Characteristics*. Metrika, 28, 109-121.

Lohr, S.L. (2000). Muestreo: Diseño y Análisis. México: International Thomson Editores.

Lynn, P. (2005). Metodología de Encuestas Longitudinales. *Seminario Internacional de Estadística Eustat*.

Medina, F. (2001). Las Encuestas de Panel. *Diseño e Implementación de un Sistema de Encuestas de Hogares*. (págs, 139-145), Aguascalientes, México. Retrieved on 15 November 2017 from https://www.cepal.org/deype/mecovi/docs/TALLER7/12.pdf

Medina, F. (2012). Esquemas de Rotación de la Muestra en Encuestas Longitudinales. Diseño de encuestas de población activa y módulos de empleo en las encuestas de hogares para medir el trabajo decente. Turín, Italia.

Montgomery, D., Peck, E., & Vining, G. (2006). *Introducción al Análisis de Regresión Lineal*. México: Compañía Editorial Continental.

Naciones Unidas. (2007). *Encuestas de hogares en los países en desarrollo y en transición* (Vol. Serie F N°96). New York. Retrieved from https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/surveys/seriesf_96-S.pdf

Naciones Unidas. (2009). *Diseño de muestras para encuestas de hogares, Directrices prácticas* (Vol. Serie F N°98). New York. Retrieved from https://unstats.un.org/unsd/publication/seriesf/Seriesf_98s.pdf

Observatorio Social Universidad Alberto Hurtado. (2008). *Entrega Nº5: Documento Metodológico Encuesta Panel Casen Medición 2007*. Santiago, Chile.

OECD. (11 de 2015). *Recomendación del Consejo de la OCDE sobre Buenas Prácticas Estadísticas*. Retrieved from http://www.oecd.org/statistics/good-practice-toolkit/OECD-LEGAL-0417-spa.pdf

OIT. (enero de 1993). *Clasificación Internacional de la Situación en el Empleo (CISE)*.Retrieved from https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087564.pdf

OIT. (2013). *La Medición de la Informalidad: Manual Estadístico sobre el Sector Informal y el Empleo Informal*. Ginebra: Organización Internacional del Trabajo.

Ritz, C., & Streibig, J.C. (2008). *Nonlinear regression with R*. Springer.

Särndal, C.E., Swensson, B., & Wretman, J. (1991). *Model Assisted Survey Sampling*. Springer.

Statistics New Zealand. (2015). *Household Labour Force Survey*. Retrieved from http://archive.stats.govt,nz/browse_for_stats/income-and-work/employment_and_unemployment/hlfs-sources-methods-2015.aspx

UNESCO. (2006). *Clasificación Internacional Normalizada de la Educación CINE 1997*. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000147002

Valliant, R., Dever, J.A., & Kreuter, F. (2013). *Practical Tools for Designing and Weighting Survey Samples*. New York: Springer.

Vanderhoeft, C. (2001). *Generalised calibration at Statistics Belgium: SPSS® module g-CALIB-S and current practices*. Statistics Belgium.

Wolter, K. (2007). *Introduction to Variance Estimation (Statistics for Social and Behavioral Sciences)*. New York: Springer.