



SOLO PARA PARTICIPANTES
DOCUMENTO DE REFERENCIA

LC/CE.21/DDR/4
29 de junio de 2022

ORIGINAL: ESPAÑOL

22-00513

Vigesimoprimer Reunión del Comité Ejecutivo
de la Conferencia Estadística de las Américas
de la Comisión Económica para América Latina y el Caribe

Santiago, 24 y 25 de agosto de 2022

**RECOMENDACIONES METODOLÓGICAS PARA LA INCLUSIÓN DE INFORMACIÓN
DE LA POBLACIÓN MIGRANTE INTERNACIONAL EN LA CALIBRACIÓN
DE LOS FACTORES DE EXPANSIÓN EN ENCUESTAS DE HOGARES:
EL CASO DE LA ENCUESTA NACIONAL DE EMPLEO DE CHILE**

Grupo de Trabajo sobre Migración Internacional de la Conferencia Estadística de las Américas

Noviembre de 2021

El Grupo de Trabajo sobre Migración Internacional de la Conferencia Estadística de las Américas (2020-2021) fue coordinado por el Instituto Nacional de Estadísticas (INE) de Chile, el Departamento Administrativo Nacional de Estadística (DANE) de Colombia y el Instituto Nacional de Estadística y Geografía (INEGI) de México, y tuvo como Secretaria Técnica a la División de Estadísticas de la CEPAL.

Este documento corresponde a uno de los cuatro productos comprometidos por el Grupo de Trabajo, y fue coordinado por el INE de Chile. Las oficinas nacionales de estadística y países miembros del Grupo son: el Instituto Nacional de Estadística y Censos (INEC) de Argentina, el Instituto Nacional de Estadística (INE) de Bolivia, el Instituto Nacional de Estadística y Censos (INEC) de Costa Rica, la Oficina Nacional de Estadística e Información (ONEI) de Cuba, el Instituto Nacional de Estadística y Censos (INEC) de, el Instituto Nacional de Estadística (INE) de Paraguay, el Instituto Nacional de Estadística e Informática (INEI) de Perú y la Oficina Nacional de Estadística (ONE) de República Dominicana.

Las opiniones expresadas en este documento, que no ha sido sometido a revisión editorial, son de exclusiva responsabilidad de los autores y pueden no coincidir con las de la Organización o las de los países que representa.

CONTENIDO

1.	Introducción	4
2.	Migración: Contexto Regional y Nacional.....	5
3.	Diseño de factores de expansión en encuestas de hogares	6
	3.1. Estimación de Parámetros	7
	3.1.1. Enfoque y Estimadores de Calibración	7
	3.1.2. Métodos de calibración: Calibración de los pesos.....	8
4.	Uso de información auxiliar sobre la población extranjera o inmigrante: El caso de la Encuesta Nacional de Empleo de Chile.....	9
	4.1. Mercado del trabajo y población extranjera en Chile: Antecedentes, hallazgos y conclusiones	9
	4.2. Estimación de población extranjera en Chile	15
	4.3. Incorporación de información sobre la población extranjera: principales resultados.....	18
	4.4. Consideraciones finales.....	21
5.	Recomendaciones	21
6.	Anexo	22
	Referencias.....	23

1. Introducción

Este documento busca describir la experiencia, en lo que respecta a los fundamentos y la metodología empleada para incorporar información sobre la población migrante como parte de los controles poblacionales que considera la *Encuesta Nacional de Empleo* (ENE) del *Instituto Nacional de Estadísticas* (INE) de Chile al momento de calibrar sus factores de expansión. Incorporar información auxiliar en la etapa de ajuste de los factores de expansión a totales poblacionales, tiene al menos dos propósitos: reducir sesgos y mejorar la precisión de las estimaciones. A la luz de sus resultados, se puede señalar que estos se alcanzaron. No obstante, muchas consideraciones metodológicas y análisis temáticos y estadísticos fueron necesarios de realizar y, sin duda, el propósito es que este documento pueda servir como fuente de consulta por parte de países que se encuentren atravesando por una realidad similar a la que se describe en el caso de Chile.

Sin duda, esta experiencia no solo reafirmó la importancia que tienen las características socio demográficas de la población bajo estudio y los fenómenos sociales de interés; sino que, además relevó la importancia que la población migración, justamente por lo diferente de sus características sociodemográficas con la población local, ha tomado en la medición de los fenómenos sociales y económicos que subyacen al mercado laboral. Adicionalmente, el masivo flujo de población migrante hacia Chile, hizo más evidente la necesidad y la importancia que este tipo información auxiliar tiene en las encuestas a gran escala que llevan adelante las Oficinas Nacionales de Estadística de la región.

Sin duda este último hecho fue, parte importante de la fuente que relevó la necesidad de considerar una nueva metodología de calibración que incorporase mayor información auxiliar sobre características auxiliares de las personas. En principio, la mayor información auxiliar busco corregir los efectos indeseables que venía generando el uso de marcos de muestreo desactualizados, los que generaban sub-coberturas en subpoblaciones de particular importancia para la medición del fenómeno, como lo son las personas comprendidas en el grupo que se conoce como fuerza de trabajo primaria. Sin embargo, la creciente importancia relativa que fue ganando la población migrante dentro de la población en edad de trabajar y, a partir de ejercicios de simulación, se permitió establecer que existía la necesidad de incorporar información auxiliar específica sobre la población migrante, información auxiliar con la que, por cierto, no se contaba en tal momento del tiempo, siendo uno de los diversos desafíos que debió enfrentar la implementación de esta exitosa implementación metodológica.

Contar con información auxiliar, confiable y oportuna, sobre la población migrante, sin duda, fue uno de los principales desafíos metodológicos que se debió enfrentar. En tanto, desde la arista temática, fundamentar la necesidad de contar con dicha información, posiblemente, fue un desafío mayor. Específicamente, existía la necesidad de mostrar con suficiente claridad, sin dejar de ser rigurosos en los fundamentos estadísticos y económicos, que la información auxiliar sobre la población migrante o extranjera no podía ser reemplazada (parcial o totalmente) por otra información auxiliar. Más precisamente, se debió fundamentar que, aun cuando una de las características de la población migrante es la de pertenecer a grupos de edad más jóvenes que la población nacional, incorporar información auxiliar sobre totales poblacionales para distintos grupos de edad, incluso bastante desagregados, era insuficiente para subsanar los efectos que venía generando la falta de cobertura de la población migrante, en la medición de los distintos indicadores al interior del mercado del trabajo.

En este sentido, con el propósito de entregar un relato simple, pero riguroso en sus fundamentos estadísticos y temáticos, este documento entrega los antecedentes necesarios para introducir al lector sobre las ventajas de la metodología implementada y las consideraciones que se deben tomar en cuenta al momento de integrar información auxiliar, específicamente, información auxiliar sobre la población migrante en la etapa de ajuste de los pesos finales o factores de expansión de las encuestas de hogares; en particular, en encuestas de fuerza de trabajo. Así, se espera que esta experiencia pueda servir como un antecedente para las encuestas que se encuentran en situaciones similares a las que en este documento se describen.

En relación con esto, a esta breve introducción le sigue un conciso contexto regional sobre la migración dentro de la región y, en particular, sobre el destino que parte de esta migración tuvo dentro de la región: Chile. Inmediatamente, se presentan los conceptos fundamentales que son el marco teórico del muestreo probabilístico y

algunos de los conceptos más relevantes en este campo. Le sigue una breve descripción de las principales etapas del diseño y/o ajustes de los factores de expansión de las encuestas de hogares, con énfasis en la etapa de calibración de los factores de expansión a totales poblacionales. Inmediatamente, se describen algunos antecedentes particulares de la ENE, la implementación metodológica y sus resultados. Finalmente, se entregan algunas recomendaciones finales.

2. Migración: Contexto Regional y Nacional

Las dos últimas décadas se han caracterizado por un aumento de la migración en algunos países de la región, movimiento que puede ser medido por la tasa de migración¹ (ver Tabla 1). En la primera década del presente siglo, sin duda la migración de personas provenientes de los países de Haití y Perú, y en menor medida de Colombia y Venezuela, fueron la principal fuente del incremento de personas extranjeras cuyo destino fue Chile. En tanto, la segunda década, estos mismos países, en particular, la población venezolana y haitiana fueron la principal fuente de origen de población extranjera que migraba hacia Chile.

Este influjo de población extranjera hacia Chile, se ve marcadamente importante durante el último lustro de la década de 2015-2020. Como se muestra en esta sección, esta mayor población extranjera, debido a las motivaciones de la migración –mejores perspectivas económicas- y por las características de las personas que migran hacia Chile –personas en edad de trabajar, sobre todo en los tramos de edades jóvenes- tuvieron y tienen impactos que, si las encuestas no incorporan de manera adecuada y oportuna, pueden verse enfrentados a reportar estimaciones que difieren de la realidad que se exhibe al interior del mercado del trabajo, así como otros fenómenos sociales como la pobreza, entre otros.

Tabla 1: Tasa de migración 2000 - 2020

	2000-2005	2005-2010	2010-2015	2015-2020
Argentina	-0,66	-0,60	0,14	0,11
Bolivia (Estado Plurinacional de)	-1,76	-1,40	-1,10	-0,84
Brasil	0,00	0,00	0,02	0,10
Chile	0,54	1,29	1,87	6,02
Colombia	-1,12	-0,89	-0,83	4,16
Ecuador	-1,81	-0,62	-0,49	2,15
Haití	-3,17	-2,88	-2,91	-3,17
México	-4,31	-1,02	-0,72	-0,48
Nicaragua	-6,62	-5,51	-4,48	-3,31
Paraguay	-1,69	-3,91	-2,69	-2,40
Perú	-6,70	-8,46	-4,20	3,12
Uruguay	-6,26	-2,99	-1,77	-0,87
Venezuela (República Bolivariana de)	-0,18	-1,46	-2,95	-22,33

Fuente: Comisión Económica para América Latina y el Caribe - CELADE. División de Población de la CEPAL. Revisión 2019 y Naciones Unidas, División de Población. Departamento de Asuntos Económicos y Sociales. Panorama Mundial de Población. Revisión 2019. - <https://population.un.org/wpp/DataQuery/>.

¹ Esta tasa corresponde al cociente entre el saldo migratorio (diferencia media anual entre los inmigrantes y los emigrantes de una población) correspondiente a un periodo determinado y la población media del mismo periodo. El resultado se multiplica por 1.000 habitantes.

Es importante notar que, para el periodo 2015-2020, no solo Chile se ha convertido en un país receptor de migración, también lo son los países de Colombia, Ecuador y Perú, por lo que son posiblemente los países, junto con Chile, que mayor atención deben prestarles a los efectos que la migración tiene sobre la realidad de los distintos fenómenos sociales que buscan dar cuenta, a través de mediciones, las encuestas de hogares en los respectivos países.

En lo que respecta específicamente a la realidad chilena², como país receptor de migración, en base a los datos del Censo de 2017, se sabe que, en 1992 se censaron 105.070 personas nacidas en el extranjero residentes habituales de Chile, lo que representó 0,8% del total de población que residía en el país. Esta proporción fue 1,3% en el Censo de 2002 (187.008 personas) y alcanzó a 4,4% en el Censo 2017, con 746.465 inmigrantes internacionales. Según el Censo 2017, 85% de la población inmigrante internacional en Chile se concentra en cuatro regiones, dos del norte y dos centrales, que en orden de prevalencia son: la Metropolitana (65,3%), Antofagasta (8,4%), Tarapacá (5,9%) y Valparaíso (5,4%).

Según el Censo 2017, 85,8% de la población inmigrante internacional tiene entre 15 y 64 años, es decir, entra en la categoría de población potencialmente activa. Esta cifra es 18 puntos porcentuales mayor que en la población residente nativa, donde 67,9% conforma ese grupo etario. La población de inmigrantes internacionales está menos envejecida que la población nativa, con 3,6% y 11,8% de personas con 65 o más años de edad, respectivamente.

De acuerdo con el Censo 2017, 50,4% de los inmigrantes nació en los siguientes tres países: Perú (25,2%), Colombia (14,1%) y Venezuela (11,1%). Los tres países de nacimiento que aportan la mayor proporción de personas al total de inmigrantes, después de Perú, Colombia y Venezuela, son: Bolivia (9,9%), Argentina (8,9%) y Haití (8,4%). En total, estos seis países constituyen 77,6% de la población inmigrante internacional en Chile.

Además, en base al Censo 2017, 66,7% del stock de inmigrantes internacionales llegó al país en el periodo 2010-2017. Los flujos migratorios han cambiado en el tiempo: de los inmigrantes internacionales que declararon llegar antes de 1990, 33,2% nacieron en Argentina. De los que declararon haber llegado entre 2000 y 2009, 47,6% nacieron en Perú. Al observar la distribución de los que declararon haber llegado entre 2010 y 2017, Argentina y Perú pierden peso porcentual y aparecen nuevos países de nacimiento, como por ejemplo Colombia (18,8%), Venezuela (15,7%) y Haití (12,0%).

Finalmente, tanto para hombres como para mujeres, el promedio de años de escolaridad de los inmigrantes internacionales es superior al de la población residente nativa. Los inmigrantes de 25 años o más tienen en total 12,6 años de escolaridad promedio, mientras que los nacidos en Chile tienen 11,0. La brecha de género respecto de los años de escolaridad es ligeramente mayor en la población residente nativa que en la de inmigrantes, ya que los hombres inmigrantes tienen en promedio 0,1 años más que las mujeres inmigrantes, mientras que en los nativos asciende a 0,2 años en promedio.

Lo que sigue a este breve contexto regional sobre el origen y el destino de parte de la migración dentro de la región, se presenta el detalle metodológico sobre el cual se fundamenta la experiencia que se describe sobre la implementación metodológica que se refiere a la incorporación de información auxiliar sobre la población migrante como parte de los controles poblacionales a los que se ajustan los ponderadores de la encuesta.

3. Diseño de factores de expansión en encuestas de hogares

La teoría del muestreo corresponde a un conjunto de métodos que permiten el estudio de una población a partir del estudio de solo una parte de esta (*Tillé (2020), p.13*). A partir de este conjunto de métodos, las encuestas de hogares buscan entregar información acerca de características de la población, a través de la información contenida en una muestra seleccionada de la población. Las muestras son de dos tipos: *probabilísticas* y *no probabilísticas*; sin embargo, solo las primeras poseen propiedades que permiten construir estimaciones y realizar inferencias sobre poblaciones finitas.

² Para mayor detalle ver INE-SDV (2018).

3.1. Estimación de Parámetros

3.1.1. Enfoque y Estimadores de Calibración

El *enfoque de calibración*, como parte de la estimación de parámetros en poblaciones finitas, según Särndal (2007), se define como “el cálculo de pesos que incorporan información auxiliar específica y que satisfacen las ecuaciones de calibración”; en tanto, sobre los pesos calibrados señala que, “... son usados para el cálculo de estimaciones lineales ponderadas para totales y otros parámetros de poblaciones finitas”. Zhang (2000) señala que la calibración es un enfoque bajo el cual “los pesos para una muestra dada, se puede decir que están calibrados con respecto a un conjunto de totales poblacionales conocidos sí, las estimaciones basadas en estos pesos calibrados, reproducen esos totales”. Por su parte, Gutiérrez (2016) señala que “... la calibración provee una forma sistemática para la incorporación de la información auxiliar en la etapa de estimación en una encuesta”. Kott (2006) señala que “La calibración puede ser utilizada para ajustar por no-respuesta y/o errores de cobertura bajo apropiados modelos cuasi-aleatorios”. Entonces, se puede señalar que, la calibración es una técnica que, a través del uso de información auxiliar (totales poblacionales), permite generar pesos o factores de expansión calibrados, los cuales cumplen ciertas condiciones y son utilizados en el proceso de estimación. Por su parte, un *estimador de calibración*, por definición de estimador, es una función de la muestra aleatoria y es utilizado para estimar un parámetro de interés. En este sentido, Särndal (2007) señala que “los estimadores de calibración son una familia de estimadores que tienen como base común la información auxiliar. Un estimador de calibración usa pesos calibrados, los cuales están tan cerca como es posible, de acuerdo a una medida de distancia dada, de los pesos de diseño originales π_k^{-1} , respetando al mismo tiempo un conjunto de restricciones, las ecuaciones de calibración”.

Al respecto, Deville y Särndal (1992) y Deville, et al. (1993) presentan un enfoque general sobre la estimación bajo el enfoque de calibración³. Sea $\mathbf{x}_k = (x_{1k}, \dots, x_{jk})$ el vector de J características auxiliares, para las cuales se asume existe información auxiliar disponible en la forma de totales poblacionales para estas características auxiliares, denotada por \mathbf{t}_x . Los métodos de calibración -que conforman el enfoque de calibración- consisten en la búsqueda de un sistema de pesos o ponderadores para los elementos de la muestra. De esta forma, el *estimador de calibración* para el total es un estimador lineal que se puede escribir como sigue:

$$\hat{t}_{y,Cal} = \sum_{k \in S} w_k(S) y_k \quad (1)$$

donde $w_k(S)$, $k \in S$ son pesos que dependen de la muestra aleatoria S . Por simplicidad, estos pesos se denotan simplemente por w_k , en tanto se tenga en cuenta que dependen de la muestra aleatoria S . El sistema de pesos está determinado por la forma en que el estimador $\hat{t}_{y,Cal}$ está *calibrado* en los totales de las características auxiliares. En otras palabras, los estimadores de totales para las características auxiliares son iguales a sus correspondientes totales poblacionales, los que vienen dados por la información auxiliar; independientemente de la muestra seleccionada. Esto se puede escribir como sigue:

$$\sum_{k \in S} w_k \mathbf{x}_k = \mathbf{t}_x \quad (2)$$

Al sistema de ecuaciones descrito en (2) se lo suele denominar *ecuaciones de calibración* o *restricciones de calibración*. Notemos que los pesos w_k estiman los totales de las características auxiliares \mathbf{x} con absoluta precisión; luego, estos mismos pesos w_k pueden ser utilizados para estimar t_y , tal como se muestra en (1), pero además, como se muestra más adelante, alcanzando mejoras en la precisión de las estimaciones.

³ Ver Tillé (2020), cap.12.

De este modo, el tipo de estimadores que utilizan información auxiliar, como lo hacen los *estimadores de calibración* además de mejorar la *precisión* en las estimaciones y *reducir sesgos*, son también estimadores *lineales* puesto que pueden ser escritos como la combinación lineal de una suma ponderada de términos (*Särndal (2007); p.100*), garantizando estimaciones *consistentes* con la información auxiliar utilizada en el proceso de estimación. Entonces, dado que los pesos calibrados w_k permiten reproducir perfectamente totales poblacionales, de forma tal que, los estimadores de calibración aseguran que cualquier estrategia de muestreo sea una estrategia de muestreo representativa con respecto de las características auxiliares, a partir del conocimiento de información auxiliar sobre esta, en este caso, totales poblacionales. De este modo, *Gutiérrez (2016)* señala que “*un estimador de calibración es aquel estimador lineal que tiene la agradable propiedad de la representatividad bajo cualquier diseño de muestreo*”.

3.1.2. Métodos de calibración: Calibración de los pesos

En la última etapa de ajuste o calibración de los pesos, los pesos ajustados por no respuesta se ajustan de forma que los nuevos pesos reproduzcan totales poblacionales (que corresponden a información auxiliar que es exógena a la encuesta).

Esto es así, pues en los documentos donde se presentan los desarrollos del enfoque de calibración, no se suelen considerar los ajustes aquí presentados; no obstante, se debe tener presente que, al momento de implementar estos métodos, son los pesos ajustados por no respuesta los que se calibran a la información auxiliar, la cual está disponible en la forma de totales poblacionales sobre algunas características auxiliares. Realizar este tipo de ajuste, como ya se ha mencionado antes, diversos autores señalan que, cuando los estudios por muestreo están afectados por problemas como la no-respuesta y la imperfección de los marcos de muestreo, es deseable contar con estimadores que garanticen algunas propiedades como:

- Sesgo pequeño o nulo.
- Estimaciones más precisas: Errores estándares pequeños.

A continuación, se resume muy brevemente los métodos o enfoques más importantes dentro del enfoque de calibración.

- A. Estimador de regresión general, el cual es definido para variables continuas principalmente. Por ejemplo, el ingreso laboral personal, como variable de interés, y un conjunto de variables auxiliares, categóricas y continuas, que explican el comportamiento de la primera. Entre las variables auxiliares podrían corresponder a características individuales como el sexo y edad de las personas, el gasto personal, etc.
- B. Post-estratificación, definido para cruces de variables categóricas. Por ejemplo, estos pueden conformarse a partir de características individuales como la edad (o grupos conformados a partir de esta), el sexo, el estatus marital, los años de educación, la raza, la nacionalidad, entre otras. Es importante enfatizar que, para estas celdas que se conforman, se requiere contar con totales poblacionales para estos mismos. De este modo, entre más características auxiliares se involucran, mayor detalle sobre la información auxiliar es requerida.
- C. Raking, definido para marginales en tablas de contingencia. Por ejemplo, que pueden corresponder a categorías de características individuales conformadas por la edad (o grupos conformados a partir de esta), el sexo, el estatus marital, los años de educación, la raza, la nacionalidad, entre otras. Distinto al anterior estimador, es importante enfatizar que, solo se requiere contar con totales poblacionales para estas marginales; es decir, totales de las marginales de las tablas de contingencia. De este modo, más características auxiliares involucra mayor información auxiliar, no obstante, el detalle no se incrementa tanto como el caso anterior.

4. Uso de información auxiliar sobre la población extranjera o inmigrante: El caso de la Encuesta Nacional de Empleo de Chile⁴

En este apartado se describen brevemente algunos antecedentes sobre cómo el fenómeno la migración, que se produjo principalmente, en la última década, afectó el mercado del trabajo chileno. A esto le sigue la presentación de los detalles sobre la metodología utilizada para la estimación de población extranjera en Chile y, finalmente, los principales resultados de la incorporación de información auxiliar sobre la población extranjera en la Encuesta Nacional de Empleo (*ENE*) de Chile, lo que permitió que los resultados de la encuesta reflejen de manera más adecuada la nueva realidad del mercado del trabajo en Chile.

4.1. Mercado del trabajo y población extranjera en Chile: Antecedentes, hallazgos y conclusiones

La *ENE*, como ocurre en todas las encuestas de hogares, busca entregar información para más de una característica de interés; es decir, es una encuesta multipropósito. En este sentido, el objetivo de la *ENE* es clasificar y caracterizar a la población en edad de trabajar; más precisamente, clasificarlas en tres grupos mutuamente excluyente: ocupados, desocupados e inactivos; además se investigan otras características de interés como las horas habitualmente trabajadas, la subutilización de la fuerza de trabajo, entre muchas otras⁵.

La *ENE* tiene un diseño probabilístico, estratificado y bi-etápico, donde en la primera etapa se seleccionan unidades primarias de muestreo (*UPM*) y en la segunda etapa, la selección corresponde a viviendas al interior de las *UPM* seleccionadas en la etapa anterior. La *ENE* es una encuesta basada en paneles rotativos; por ello las unidades de análisis son entrevistadas más de una vez en varios momentos. Además, la muestra total de la encuesta se distribuye en tres submuestras⁶. Finalmente, se investigan a las personas al interior de las viviendas seleccionadas. De este modo, como parte del proceso de diseño de los factores de expansión, se calculan los pesos de diseño, los que dan cuenta de las probabilidades de selección de las viviendas dentro de las dos etapas de selección antes mencionadas. Previo y posterior al proceso de recolección de datos en terreno, aun cuando se extremen los esfuerzos operativos, suceden una serie de eventos que llevan al necesario ajuste de los pesos de diseño.

En primer término, por conseguir el contacto y la respuesta por parte de las personas que forman parte de las viviendas seleccionadas, es habitual la constatación que no todas las unidades seleccionadas fueron susceptibles de logro y, por tanto, algunas de las unidades seleccionadas no tienen asociada una respuesta. Como medida de mitigación del sesgo que esto podría generar, se realizan una serie de ajustes sobre los pesos de diseño: un ajuste por no respuesta a los factores de expansión re-ponderando el peso de las unidades para las que no se obtuvo respuesta en aquellas para las que sí se obtuvo, con el fin de compensar la pérdida de unidades. Realizado lo anterior, se cuenta con un factor de expansión a nivel de viviendas, con la particularidad de que tanto sus totales como su distribución fue ajustada por la ausencia de respuesta a nivel de *UPM*. Adicionalmente, es posible también pasar de pesos a nivel de viviendas a pesos para individuos; es decir, para esto se calibran los pesos finales o factores de expansión a totales poblacionales.

El factor de expansión se calibra con el objetivo de que las estimaciones de población reproduzcan los totales poblacionales según las proyecciones de población para el mes central del trimestre de recolección de datos. A contar del trimestre móvil noviembre 2019 - enero 2020, la *ENE* cambió el método de calibración, incorporó más información auxiliar, actualizando sus fuentes. Estos cambios no solo buscaban entregar estimaciones más precisas, sino también ser una forma de precautelar sobre los efectos indeseables de la desactualización de los marcos de muestreo.

⁴ Ver *INE (2018a)*, *INE (2018b)*, *INE (2019a)*, *INE (2019b)*, *INE-ENE (2020)*, *INE-DMIE (2020)* e *INE (2020)*.

⁵ Para mayor detalle sobre la temática de la *ENE* se puede consultar *INE-ENE (2020)*.

⁶ Para mayor referencia sobre el diseño muestral de la *ENE* se puede consultar *INE-ENE (2020)*.

Hasta el trimestre móvil octubre – diciembre de 2019, el factor de expansión se calibraba con el objetivo de que las estimaciones de población reprodujeran los totales poblacionales según las proyecciones de población para el mes central del trimestre de recolección de datos, considerando dos tramos etarios. La calibración en la *ENE* se realizaba utilizando el método de post-estratificación, siendo los post-estratos de la encuesta, los grupos originados por el cruce de las variables: estrato geográfico, sexo y tramo de edad. Para cada uno de estos grupos se contaba con información auxiliar proveniente de las proyecciones de población construidas a partir del Censo 2002.

Específicamente, la variable tramo de edad se desagrega en dos categorías: menores de 15 años y 15 años o más. La variable “estrato” corresponde a las delimitaciones geográficas definidas como estratos de muestreo para la *ENE*, considerándose 160 a nivel nacional. La cantidad de post-estratos para la calibración se origina a partir de las interacciones entre las variables estrato (160 grupos), tramo de edad (2 grupos) y sexo (2 grupos), obteniéndose en total 640 post-estratos. De esta manera, en su momento, la forma de calibrar los factores de expansión a nivel de personas en la *ENE* puso énfasis en la desagregación territorial, 160 estratos territoriales, por sobre las características de los individuos en el resguardo de ciertos totales de interés (sólo se distingue por sexo y entre dos grupos de edad). En consecuencia, si la muestra efectiva no tiene características sociodemográficas similares a las de la población real, la metodología de calibración no podía corregir este *problema*, ya que, los factores de expansión ponderaban por igual, independiente de su edad, condición de nacionalidad, u otra variable correlacionada con la situación de las personas en el mercado laboral, a todos los individuos mayores (15 y más años) del mismo sexo y dentro de la misma unidad de muestreo (manzana-sección) en su respectivo estrato geográfico.

Por tanto, no hacer una distinción más precisa en la ponderación, mediante la incorporación en la construcción de los factores de expansión de una mayor cantidad de variables y/o categorías relacionada al fenómeno de estudio, llevaba a que algunas subpoblaciones, terminen siendo sub o sobre ponderadas, con evidentes efectos sobre las estimaciones de la encuesta (Ver *INE (2019a)*), particularmente afectando a dos grupos de extrema sensibilidad para la medición de las estadísticas laborales: *fuerza de trabajo primaria (FTP)* y *población extranjera residente en el país*.

Respecto a los problemas que un marco muestral desactualizado podría causar en los resultados que reportan este tipo de encuestas; durante 2018, previo a la implementación de los vigentes marcos muestrales que utilizaba la *ENE*, se llevó adelante una prueba piloto paralela a la muestra oficial de la *ENE*⁷. La prueba piloto, cuya muestra fue seleccionada desde un marco muestral más actualizado, tuvo el propósito de verificar si las características sociodemográficas de la población que sistemáticamente venía siendo excluida de las selecciones muestrales de la *ENE*, son diferentes de las características de la población que efectivamente venía siendo estudiada en la *ENE*⁸.

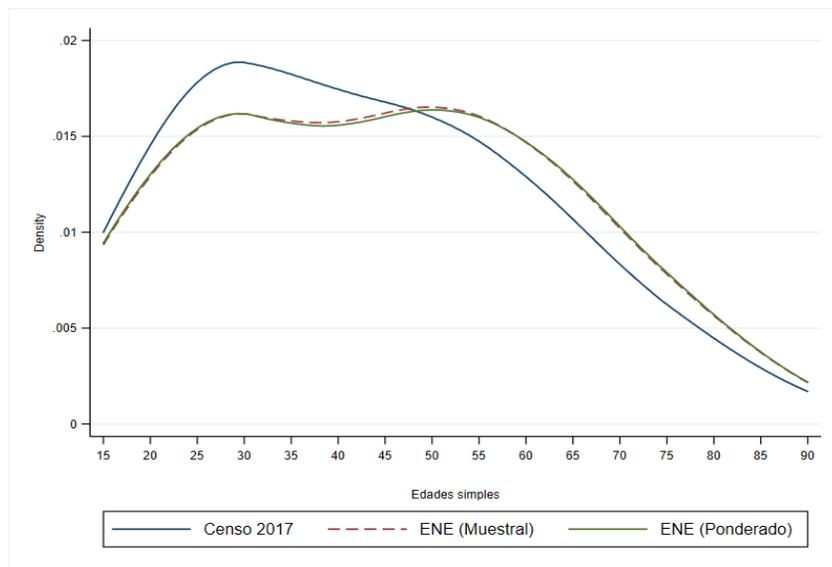
Dentro de los principales resultados, en *INE (2018b)* se encontró evidencia que sugería que la población en edad de trabajar que residía en las manzanas excluidas de las selecciones muestrales de la *ENE*, presentaban una mayor participación en *tramos de edades jóvenes*; en específico, personas entre los 25 y 54 años, la población denominada como *fuerza de trabajo primaria*; al igual que una mayor presencia de *personas extranjeras*, respecto de la participación que reflejan estas dos subpoblaciones dentro de la muestra oficial de la *ENE*; es decir, la desactualización del marco muestral generó que se omitiera a parte de la población que, por un lado, forma parte de la población objeto de estudio de la encuesta y, por otro, presenta características distintas de la población que si ha venido formando parte de las muestras oficiales de la *ENE* y, como se sabe, tienen una particular importancia dentro del fenómeno que busca estudiar la encuesta.

⁷ Para mayor detalle sobre el diseño de la prueba piloto ver *INE (2018a)*; para detalle sobre los principales resultados de la prueba piloto ver *INE (2018b)* e *INE (2019a)*.

⁸ Diversos estudios muestran que las características sociodemográficas y la composición del hogar son determinantes que explican las decisiones de las personas sobre su vínculo con el mercado del trabajo.

La cobertura insuficiente de estas dos subpoblaciones, debido al particular vínculo que tienen estas personas con el mercado del trabajo, ha tenido y tiene efectos no despreciables sobre los principales indicadores que continuamente ha reportado y aun reporta la *ENE*. Evidentemente, aun cuando en los últimos años Chile ha experimentado un significativo influjo de migración internacional (Ver *Aldunate, et. al. (2019)*), la encuesta ha reportado una estimación significativamente menor para una subpoblación que viene principalmente a trabajar, entonces, es natural pensar que los resultados reportados por la *ENE* no correspondían a un adecuado reflejo de la coyuntura del mercado laboral de Chile.

Gráfico 1: Densidad por edades simples para la PET según ENE y Censo



Fuente: Censo 2017 y ENE Marzo 2017 – Mayo 2017. Tomado de *INE (2019b)*.

Cuando se comparan las estadísticas oficiales de la *ENE* con las del Censo 2017, como se constata en el Gráfico 1, la ocurrencia de personas *jóvenes* (< 50 años) en la encuesta es menor a lo que se encuentra en la población y la ocurrencia de personas *mayores* (50 y más años) es mayor a la de la población. Por otro lado, dentro de los hechos estilizados del mercado del trabajo, está ampliamente documentado que la edad determina la situación de las personas dentro y fuera del mercado del trabajo, siendo ésta una variable que está fuertemente correlacionada con las principales variables de interés de la encuesta, por tanto, es importante incorporarla como parte de la información auxiliar para calibrar los factores de expansión en la *ENE*.

Otra subpoblación que ha tenido una subrepresentación importante en la encuesta es la población extranjera residente en Chile. En la encuesta, para el trimestre marzo-mayo 2017 se estimó en 1,4 % del total nacional, mientras que en el Censo de abril de 2017 los extranjeros alcanzaron el 4,3 % de la población residente en Chile. Si se restringe la población a comparar exclusivamente a la FTP, esta diferencia se acrecienta aún más, llegando a consignar 4,3pp. El documento “Estimación de personas extranjeras residentes en Chile al 31 de diciembre de 2018” (Instituto Nacional de Estadísticas y Departamento de Extranjería y Migración, 2019) desarrollado en colaboración con Policía de Investigaciones de Chile, Ministerio de Relaciones Exteriores y Servicio de Registro Civil e Identificación, señala que este importante grupo de la población, se estimó en 1.251.225 personas en la fecha señalada. Esto implicó que los extranjeros a diciembre de 2018 representaban el 6,7 % de la población total, al tomar como referencia las proyecciones de población derivadas del Censo 2017. Como ilustra la Tabla 5, este grupo, según las fuentes antes mencionadas, los extranjeros dentro de la población total llegaban a representar el 10,1 % de la FTP, sin embargo, en la *ENE*, para el mismo periodo comparable, sólo representaban el 3,4 % de la FTP estimada en la encuesta.

Esta brecha, entre lo que reportaba la encuesta y lo que nos señalaban los *benchmark* de comparación como el Censo de 2017 y la Estimación de población extranjera residente en Chile, fueron una constatación de la pérdida de confiabilidad en estimaciones de relevancia para la medición y monitoreo del mercado laboral chileno como consecuencia de la elevada inmigración de los últimos años.

Tabla 2: Participación de la población extranjera: Población, PET y FTP

Extranjeros (%)	Censo Abr. 2017	ENE MAM-2017	Diferencia (p.p.)	Est. Extr. Dic. 2018	ENE NDE-2018	Diferencia (p.p.)
Población	4.3	1.4	-2.8	6.7	2.1	-4.6
PET	4.8	1.6	-3.2	7.5	2.2	-5.2
FTP	6.6	2.3	-4.3	10.1	3.4	-6.7

Fuente: Censo 2017 y ENE. Tomado de *INE (2019b)*.

Sobre esto, en *INE (2019a)* se relevó la necesidad de contar con un nuevo paradigma sobre la forma en la cual se calibran los factores de expansión. En el documento se señala que, “*Un marco muestral actualizado reduce la subcobertura de subpoblaciones que, por sus características, tienen directas implicancias sobre las estimaciones de los principales indicadores del mercado laboral. Sin embargo, una muestra que refleja de manera más adecuada la realidad nacional, para maximizar sus bondades, debe estar acompañada por una adecuada metodología de calibración*”. En este mismo sentido, también se señala que “*... una metodología de calibración de los factores de expansión que considere a la edad como una variable necesaria para recoger de manera eficiente el correcto vínculo y dinámica que se genera al interior del mercado del trabajo*”. De modo general, se enfatiza que: “*... es importante avanzar en una nueva metodología de calibración de los factores de expansión, incorporando variables que estén fuertemente correlacionadas con la situación de las personas dentro y fuera del mercado laboral, como parte de las variables utilizadas para asignar un peso apropiado a las personas de la muestra de la ENE*”.

De este modo, el fundamento temático de la nueva metodología de calibración de la *ENE* reconoce la importancia que tienen las características de las personas sobre su situación dentro y fuera del mercado laboral. Más aun, incorporar información sobre estas características, denotadas por **x**, entregan mejoras sustanciales en cuanto a la reducción de sesgos sobre la medición de indicadores del mercado laboral. En particular, además del sexo de las personas, la importancia que tienen la edad y la condición de las personas según si son chilenos/as o extranjeros/as, como parte de la información que debe ser considerada, dentro del proceso de ponderación y el posterior proceso de estimación.

Distinguir entre la necesidad de contar con información auxiliar para una característica clave dentro del mercado del trabajo, como es la edad de las personas, e información auxiliar para otra característica como es la condición o status de las personas según si son chilenos o extranjeros, resultó ser relevante, en el sentido que permitió reducir sesgos que, de otro modo, no hubieran podido ser resueltos, pero además, permitió una mejor caracterización de los distintos fenómenos y las dinámicas que subyacen al mercado del trabajo chileno, caracterizado por una creciente importancia relativa de la población migrante o extranjera dentro de la población en edad de trabajar y, por tanto, dentro de la población objetivo de la medición del fenómeno de empleo (ver Tabla 2).

Conscientes de los efectos indeseables que el uso de marcos de muestreo desactualizados estaba generando en los principales indicadores reportados por la *ENE*, debido a la cobertura insuficiente de parte de la población en tramos de edades jóvenes (llevando a una sobre representación de la población en tramos de edades avanzadas) y la población extranjera o migrante; y de las exigencias que suponía realizar el ajuste de los pesos finales o factores de expansión mediante el método de post-estratificación, limitando la incorporación de nueva información auxiliar, era necesario contar una aproximación de los efectos que generaba el uso de un marco muestral imperfecto y el uso insuficiente de información poblacional exógena.

De este modo, a partir de datos del Censo de 2017, se simuló⁹ una población que reproducía de manera precisa la estructura demográfica de características auxiliares como: sexo, tramos de edad¹⁰ y dos categorías de procedencia (chileno/a - extranjero/a). Para las personas dentro de la población en edad de trabajar; además, se simularon cuatro estados dentro del mercado del trabajo, denotados por Y_i , con $i = \{1,2,3,4\}$, los cuales corresponden, respectivamente, a estar en la *fuerza de trabajo* ($i = 1$), estar *ocupado* ($i = 2$), estar *desocupado* ($i = 3$) y estar en la *inactividad* ($i = 4$).

A partir de la población simulada, se conocen los totales poblacionales de las características auxiliares, y además se calcularon los siguientes parámetros de interés del mercado laboral: el total personas en la *fuerza de trabajo* (Y_1), *ocupados* (Y_2), *desocupados* (Y_3) e *inactivos* (Y_4); así como, las correspondientes tasas de: *participación* (r_1), *ocupación* (r_2), *desocupación* (r_3) e *inactividad* (r_4).

Con el propósito de visualizar los efectos del marco muestral, una vez calculados los verdaderos parámetros poblacionales de interés, desde la población simulada, se omitieron personas de la población simulada. Para que los efectos puedan ser comparables con los efectos que se producían en la ENE, la omisión de personas se realizó de manera aleatoria, pero restringida a únicamente las dos subpoblaciones sobre las cuales se sabe existe cobertura insuficiente debido a la desactualización del marco muestral: personas extranjeras y personas en la fuerza de trabajo primaria (25 y 54 años).

Con el propósito de mostrar que los resultados corresponden a un efecto estable y robusto, desde la población que presenta omisión en las dos subpoblaciones antes mencionadas y como una aproximación de un marco de muestreo desactualizado, se realizaron 1000 simulaciones de selecciones muestrales, las que son utilizadas para estimar los totales de las cuatro características de interés del mercado del trabajo antes mencionadas y sus tasas correspondientes.

Las estimaciones sobre cada una de las 1000 simulaciones son realizadas a partir del estimador de Horvitz–Thompson y tres estimadores de calibración. El primero incorporó información auxiliar sobre sexo, y edad en dos grupos; el segundo sobre sexo y edad en 16 grupos de edad; el tercero sobre sexo, 16 grupos de edad y nacionalidad.

Tabla 3: Distribución (%) promedio: población en edad de trabajar y población extranjera

Estimadores y Parámetros	Población en edad de trabajar						Nacionalidad	
	15-24	25-34	35-44	45-54	55-64	65+	Chilenos/as	Extranjeros/as
Horvitz-Thompson	19.1	18.7	16.5	16.5	14.2	14.9	98.40	1.60
Sexo & Menores-Mayores	19.1	18.7	16.5	16.5	14.2	14.9	98.39	1.61
Sexo & Edades	18.6	19.7	17.2	16.8	13.5	14.2	98.36	1.64
Sexo & Nacionalidad	18.6	19.7	17.2	16.8	13.5	14.2	95.6	4.4
Parámetros	18.6	19.7	17.2	16.8	13.5	14.2	95.6	4.4

Fuente: Simulaciones en base a Censo 2017. Tomado de *INE (2019b)*.

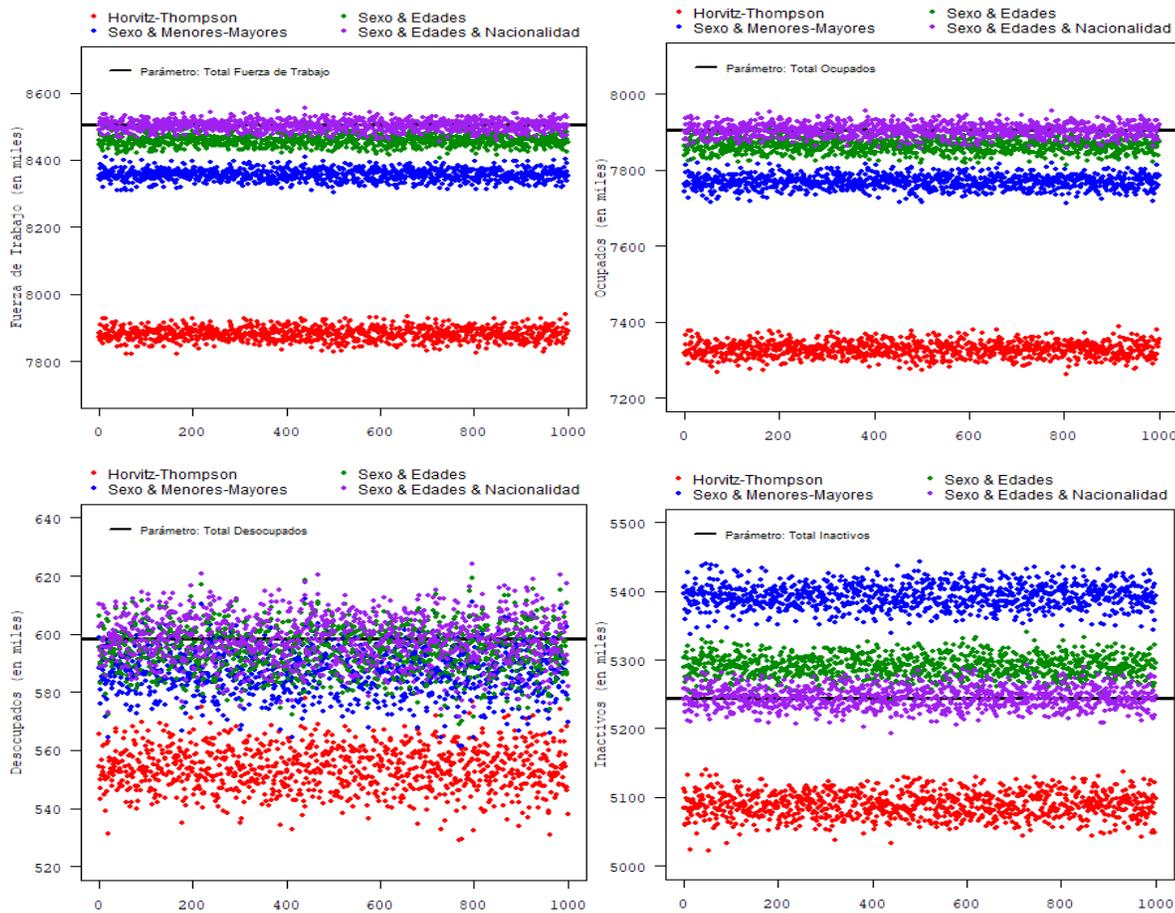
Los resultados mostraron que, como parte de las restricciones, los totales estimados para las características auxiliares coincidían con los totales poblacionales que se utilizaban como parte de los controles poblacionales. Evidentemente, únicamente cuando se incorporaba información auxiliar sobre la nacionalidad, las estimaciones sobre esta característica auxiliar eran consistentes con sus totales poblacionales (Ver Tabla 3).

⁹ Los detalles se pueden encontrar en la sección 4 de *INE (2019b)*.

¹⁰ La población se reproduce considerando 16 grupos de edad: el tramo de 0-14 años, 14 quinquenios al interior del tramo entre 15 y 84 años y el tramo de 85 y más años.

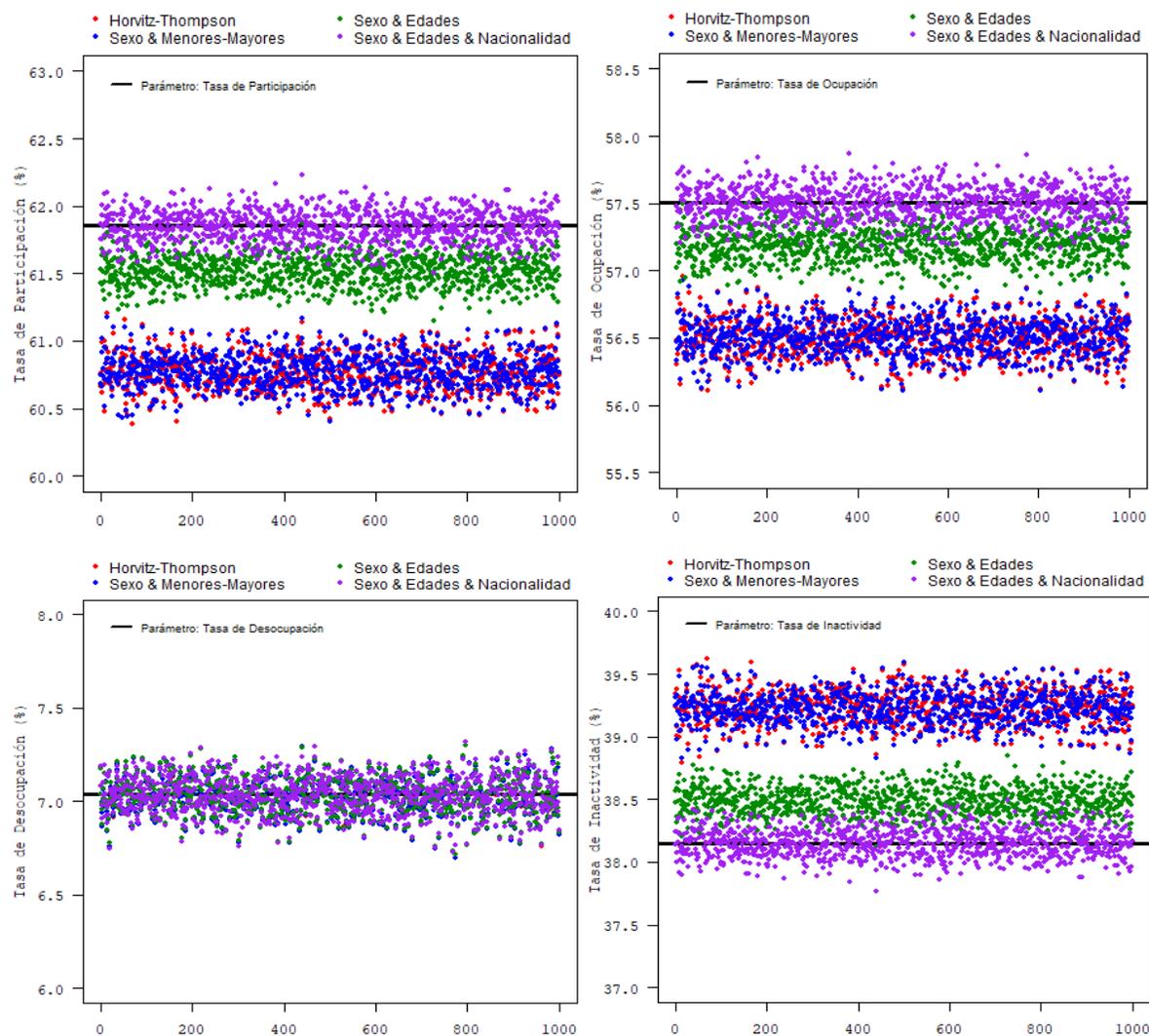
Las características auxiliares, sin embargo, no son el propósito de investigación. No obstante, es importante ver como al cerrar las brechas dentro de la estructura de la población, le siguen el cierre de sesgos dentro de las estimaciones de las características de interés (Ver Gráficos 2 y 3).

Gráfico 2: Parámetros y estimaciones de las muestras simuladas. Totales según estimador



Fuente: Simulaciones en base a Censo 2017. Tomado de INE (2019b).

Como se puede apreciar en ambos gráficos, cuando se incorpora información auxiliar por grupos de edad (puntos verdes), respecto de la que incorpora información para menores y mayores de 15 años (puntos azules), la reducción de los sesgos (la diferencia respecto del parámetro: línea negra) es considerable; sin embargo, aún persisten diferencias en promedio respecto del parámetro poblacional. Únicamente cuando se incorpora información auxiliar, además de totales para los grupos de edad y sexo, sobre la condición de nacionalidad, es que los sesgos se reducen de manera notable y para todos los casos; es decir, tanto para los cuatro totales como sus respectivas tasas.

Gráfico 3: Parámetros y estimaciones de las muestras simuladas. Tasas según estimador

Fuente: Simulaciones en base a Censo 2017. Tomado de *INE (2019c)*.

4.2. Estimación de población extranjera en Chile¹¹

La *ENE*, producto del uso de marcos muestrales desactualizados (hasta antes del trimestre móvil septiembre – noviembre de 2018 y febrero – abril de 2019) y del anterior método para el ajuste de sus factores de expansión, que era insuficiente al momento de resolver problemas asociados a la falta de cobertura y/o respuesta de importantes subpoblaciones para la medición del fenómeno experimentaba sub-cobertura en dos subpoblaciones muy importantes: Población en la fuerza de trabajo primaria (jóvenes) y extranjeros; que por su particular vínculo con el mercado del trabajo, tienen evidentes efectos sobre las principales estimaciones que se derivan acerca del

¹¹ Ver *INE-DMIE (2020)* para mayor detalle sobre este punto.

mercado laboral. En el caso particular del mercado del trabajo, la *ENE* en su metodología de calibración de los factores de expansión, vigente hasta el trimestre móvil octubre – diciembre de 2019, no podía corregir estas deficiencias y, por tanto, tampoco era capaz de corregir los sesgos generados sobre los principales indicadores que reportaba la encuesta.

Es importante destacar el aumento de la participación relativa experimentado por la población extranjera en la población total y, en particular, dentro la población en edad de trabajar y en la fuerza de trabajo primaria. En la Tabla 4 se aprecia cómo la población extranjera, en poco más de 1 año, pasó de corresponder al 4.86 % de la Población en edad de trabajar (PET) al 7.45 % y de corresponder al 6.54 % de la fuerza de trabajo primaria (FTP) al 10.12 %.

Tabla 4: Participación relativa de la población extranjera según fuente de información: Censo 2017, Estimación de personas extranjeras 2019, Estimaciones y proyecciones de la población de Chile (base Censo 2017)

Grupo Población	Participación Extranjeros (%)		
	Censo 2017 ^a (19 abril 2017)	Estimación INE-DEM ^b (respecto de la proyección de población: 30 junio 2018)	Estimación INE- DEM ^b (respecto de la proyección de población: 30 junio 2019)
Población total	4.35	6.67	7.81
Población en edad de trabajar	4.86	7.47	8.72
Fuerza de trabajo primaria	6.54	10.16	12.01

Fuente: Censo 2017 y Estimación de Personas Extranjeras Residentes en Chile, al 31 de diciembre de 2019 y Estimaciones y proyecciones de población de Chile, base Censo 2017. Tomado de *INE-DMIE (2020)*, donde se pueden revisar las notas a y b.

En los últimos años, la población extranjera ha aumentado su peso relativo en la población total. No obstante, lo realmente relevante de esta población para el mercado laboral, es el aumento que ha experimentado dentro de la población en edad de trabajar –que es la población objeto de estudio de la encuesta– y, de manera particular, en la fuerza de trabajo primaria. Si a esto se agrega el hecho de que la población extranjera tiene un vínculo con el mercado del trabajo distinto del que tiene la población nacional se hace necesario que la encuesta no subestime de forma relevante a esta población. En este sentido, potenciar las encuestas muestrales con una calibración de los factores de expansión ad hoc, es reconocer la importancia que tiene este grupo de la población en los distintos fenómenos sociales donde sistemáticamente ha sido subrepresentado, ya sea por razones de marco muestral, diseño estadístico, criterios del trabajo de campo o errores no muestrales propios de este tipo de estudios. Para esto, se hace necesario contar con información auxiliar sobre la población extranjera, de forma que, al incorporar esta información como parte de los totales poblacionales a los cuales se ajusta la encuesta, los sesgos que su falta de cobertura genera, se puedan minimizar.

Con el propósito de contar con un insumo que pueda ser utilizado como una fuente de información auxiliar sobre la población extranjera residente en Chile, se procedió a estimar el total (stock) de personas extranjeras residentes en Chile, con frecuencia mensual, para el período 2012-2025. Esto fue realizado considerando que en su momento no se contaba con ningún registro administrativo que pudiera entregar un reporte periódico sobre el total (stock)

de personas extranjeras que residen en Chile. Para esto, en primer lugar, se definió aproximar este total a partir de parte de la información disponible de las Estimaciones y Proyecciones de Población base 2017; específicamente, el componente del saldo migratorio. Con esto, se definieron stocks anuales de personas extranjeras, tomando como un stock inicial junio de 2012 y, a partir de la suma del flujo de inmigrantes entre los años 1992-2012 (ver Tabla 5), que utilizaron las Estimaciones y proyecciones de población, se construyeron los siguientes stocks para cada mes de junio de los años sucesivos, hasta el 2025. En síntesis, el cálculo para la aproximación del stock de personas extranjeras en Chile, a junio de cada año, fue realizado sumando al stock inicial (junio de 2012) los correspondientes flujos anuales de inmigrantes de las Estimaciones y Proyecciones de Población.

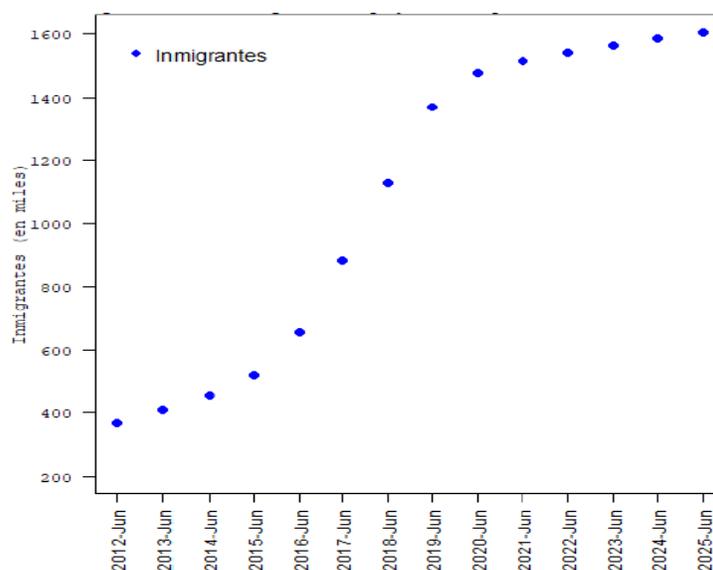
Tabla 5: Estimación y stock anual de población extranjera (al 30 de junio de cada año) a partir del saldo migratorio de las proyecciones de la población base 2017

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

Fuente: Componente Inmigración de las proyecciones de población base 2017. Tomado de *INE-DMIE (2020)*.

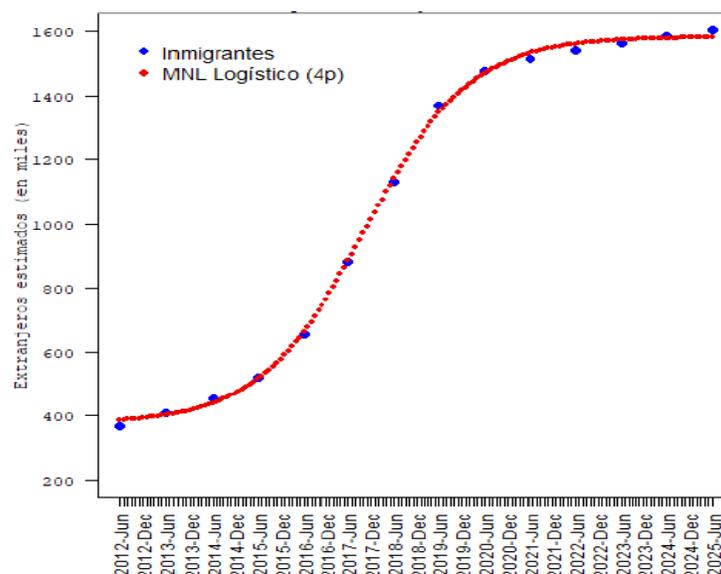
Los totales (stocks) de personas extranjeras a junio de los años comprendidos entre 2012 y 2025, mostraban una trayectoria no lineal, en particular, describe una función logística (ver Gráfico 4). Este tipo de funciones, junto con la exponencial, son típicamente utilizadas en los modelos de crecimiento poblacional.

Gráfico 4: Población inmigrante de las proyecciones de población base 2017



Fuente: Componente Inmigración de las proyecciones de población base 2017. Tomado de *INE-ENE (2020b)*.

Gráfico 5: Población inmigrante de las proyecciones de población base 2017 y estimación de la población extranjera



Fuente: Componente Inmigración de las proyecciones de población base 2017 y estimación del modelo no lineal logístico de 4 parámetros. Tomado de INE-ENE (2020b).

Este tipo de funciones tienen la particularidad de presentar la no linealidad en los parámetros, lo cual justifica realizar el ajuste de los datos, a partir de un modelo no lineal para el crecimiento poblacional. En este sentido, se optó por escoger como modelo final para la estimación de población inmigrante, el correspondiente al del modelo logístico de 4 parámetros. El modelo seleccionado tiene como función de media a una función logística de cuatro parámetros (ver Gráfico 5).

4.3. Incorporación de información sobre la población extranjera: principales resultados¹²

La *ENE*, a contar del trimestre móvil noviembre 2019 - enero 2020, cuenta con una nueva metodología de calibración, la cual, por un lado, se construye sobre sólidos fundamentos en la temática de empleo y, por otro lado, es implementada mediante una robusta técnica estadística. Como se ha mencionado, la calibración es una técnica que permite el uso de información auxiliar (estimaciones y proyecciones de población, registros administrativos, etc.) de un conjunto de variables, para ajustar los pesos o factores de expansión y mejorar la estimación de otras características de interés. Esta técnica, al mismo tiempo, tiene propósitos adicionales: *mejorar la precisión de las estimaciones y disminuir el sesgo* debido a la no-respuesta o a la imperfección de los marcos muestrales.

¹² En un anexo, al final de este documento, se presentan los principales resultados de una encuesta realizada dentro de los países que conforman el Grupo de Trabajo sobre Migración Internacional. Este cuestionario buscaba conocer las principales características en lo que se refiere al ajuste de los factores de expansión sobre totales poblacionales. En dicho anexo se puede apreciar que, distinto al caso de Chile, las encuestas de hogares no incorporan información auxiliar sobre totales poblacionales de las personas migrantes.

La información auxiliar que utiliza la *ENE* tiene fuente principal en dos insumos:

- Las proyecciones y estimaciones de población, según sexo y edad simple, en distintos niveles de desagregación territorial, las cuales se derivan de las proyecciones y estimaciones de población, base Censo 2017.
- La estimación mensual del total de la población extranjera en Chile.

Ambas fuentes de información, en los niveles de desagregación utilizados, corresponden a información confiable y de una construcción metodológicamente robusta. En el primer caso, estas se derivan de las Proyecciones y Estimaciones de Población, base Censo 2017, las cuales son construidas a partir de modelos demográficos que utilizan toda la información disponible, tanto de Censo pasados como de registros administrativos, pero fundamentalmente, de los resultados del reciente Censo de 2017. En el segundo caso, se estudiaron varios modelos antes de seleccionar uno, el cual es estadísticamente válido pues cumple con los supuestos que lo subyacen y, además, sus resultados se encuentran muy próximos a los de la estimación que es reportada al 31 de diciembre de cada año¹³.

La información auxiliar es entregada como un vector de totales poblacionales (x_1, \dots, x_p) , llamadas *marginales de calibración*, bajo el siguiente detalle:

- A nivel nacional: población total, según sexo, para 13 tramos de edad: [0-14], [15-19], [20-24], [25-29], [30-34], [35-39], [40-44], [45-49], [50-54], [55-59], [60-64], [65-69] y [70+].
- A nivel nacional: estimación de la población extranjera.
- A nivel regional: población total, según sexo.
- A nivel regional: población en Edad de Trabajar (PET), según sexo.
- A nivel regional: población total Área Urbana, según sexo.
- Región Metropolitana: población total, según sexo, para 6 grupos de edad: [15-24], [25-34], [35-44], [45-54], [55-64] y [65+] y, para el resto de las regiones, población total para 3 grupos de edad: [15-34], [35-54] y [55+].
- A nivel provincial: población total en edad de trabajar para 29 provincias¹⁴.
- A nivel de grandes conurbaciones: población total en edad de trabajar para 4 conurbaciones¹⁵.

Uno de los propósitos de la nueva metodología, como se señaló dentro de los fundamentos metodológicos, tuvo relación con alcanzar mejoras en cuanto a la precisión de las estimaciones de variables de interés, las cuales, guardan estrecha relación con la información auxiliar (marginales de calibración), pero son distintas entre sí.

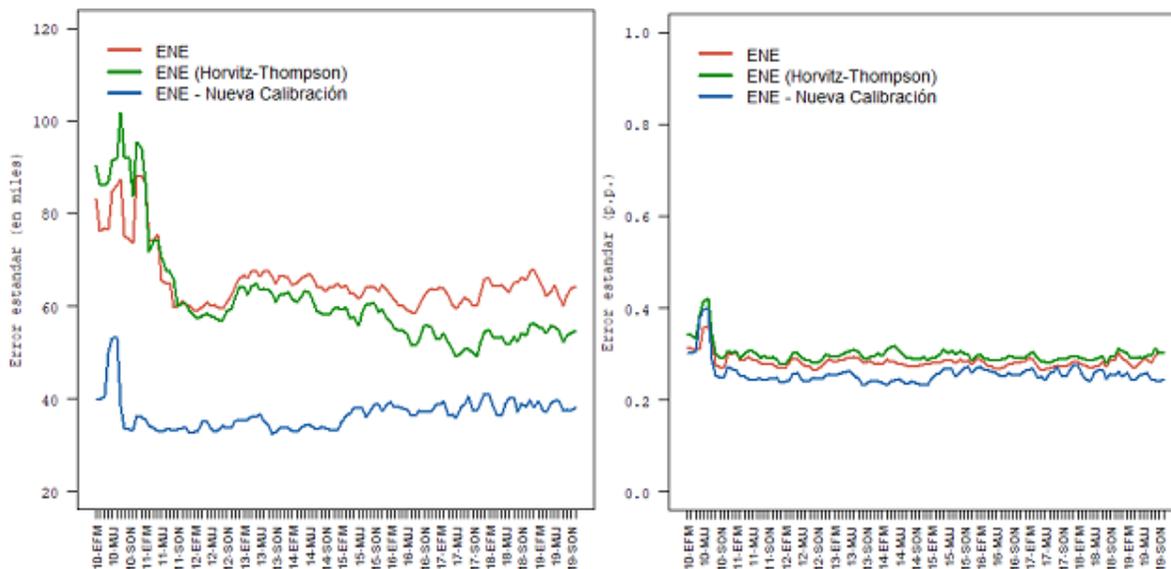
Estas mejoras se tradujeron en obtener estimaciones con un menor error de muestreo, el cual puede ser medido por estadísticos como la varianza, el error estándar, el coeficiente de variación y/o los intervalos de confianza. Como una forma de visualizar estas mejoras en la precisión, en el Gráfico 6, se presentan la nueva serie de los errores estándar para los totales nacionales de las personas ocupadas y la tasa de ocupación.

¹³ La estimación INE-DEM al 31 de diciembre de 2018 correspondió a 1.250.365, en tanto, la estimación empleada en la *ENE*, que corresponde a un modelo no lineal, alcanza la cifra de 1.251.869 personas extranjeras. Por su parte, la estimación INE-DEM al 31 de diciembre de 2019 correspondió a 1.492.522, en tanto, la estimación empleada en la *ENE*, que utiliza el modelo antes mencionado – el cual fue desarrollado con anterioridad a los resultados entregados de población extranjera para 2019 – alcanza la cifra de 1.415.977 personas extranjeras (ver INE - DEM, 2020).

¹⁴ El detalle sobre los criterios y del listado de las provincias se encuentra en la página 76 y en el Anexo 6, respectivamente, en *INE-ENE (2020)*.

¹⁵ El detalle sobre el listado de las conurbaciones se encuentra en la página 76 y 77 en *INE-ENE (2020)*.

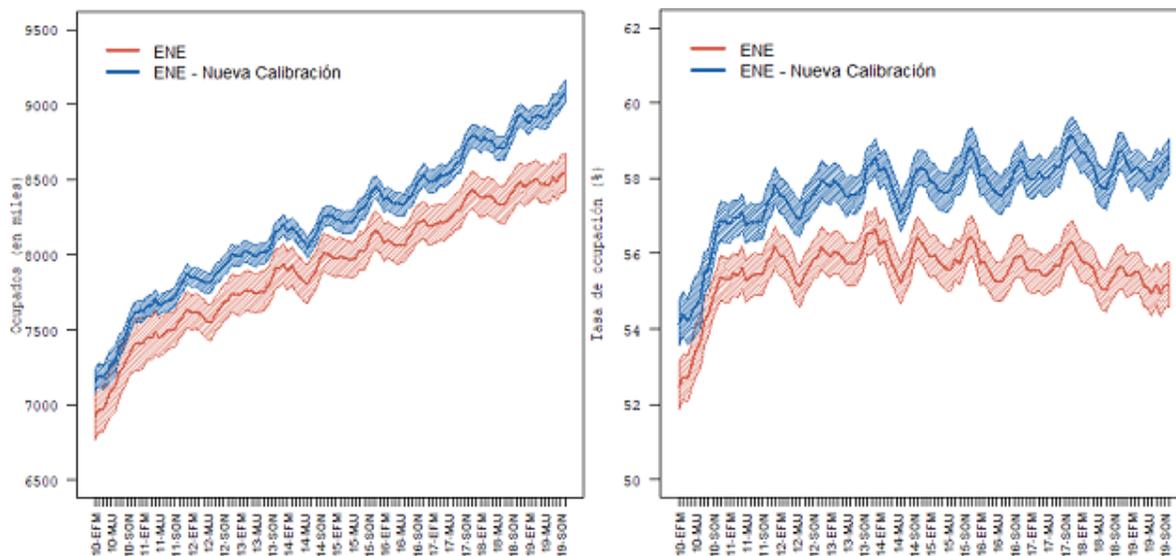
Gráfico 6: Ocupados y Tasa de Ocupación. Errores estándar, según estimador



Fuente: Datos de la ENE. Tomado de *INE (2020)*.

Finalmente, la reducción de sesgos fue otro de los propósitos. En el Gráfico 7 se muestra cómo se corrigen las subestimaciones en los stocks sobre ocupados y la tasa de ocupación.

Gráfico 7: Ocupados y Tasa de Ocupación. Estimación, según estimador



Fuente: Datos de la ENE. Tomado de *INE (2020)*.

relacionadas con la situación en la que se encuentran las personas dentro del mercado del trabajo. Reconocer que las mujeres no se vinculan del mismo modo que lo hacen los hombres dentro del mercado del trabajo, puesto que las primeras tienen menores probabilidades de participar en el mercado del trabajo y probabilidades menores de participar con éxito; es decir, estar ocupadas, forma parte de los hechos estilizados del mercado del trabajo de los países de la región. Del mismo, el vínculo no es el mismo por tramos de edad, incluso controlando por variables como el sexo de las personas (ver *INE (2020)*). Del mismo modo que el vínculo con el mercado del trabajo no es el mismo entre hombres y mujeres o entre mujeres en distinto grupo de edad o entre hombres y mujeres en un mismo grupo de edad y, por tanto, la relevancia de incorporar información auxiliar sobre dichas características auxiliares, el vínculo con el mercado del trabajo no es el mismo entre migrantes y nacionales, no solo porque las características demográficas de los primeros difieren de los segundos (los primeros suelen concentrarse en tramos de edad más jóvenes), sino que también sus habilidades (capital humano), necesidad y motivaciones suelen ser distintas.

Por otro lado, las consideraciones anteriores se hacen más latentes cuanto mayor sea el peso relativo que tenga una subpoblación dentro del fenómeno social de interés. En el caso descrito, la población objetivo corresponde a la población en edad de trabajar y, en particular, la población económicamente activa, es decir, la que forma parte de la fuerza de trabajo. En la Tabla 6 se puede apreciar la importancia relativa que fue ganando la población extranjera dentro de la población en edad de trabajar y, en particular, dentro de la fuerza de trabajo. Del mismo modo y por la misma razón que, en una encuesta económica se seleccionan con certeza empresas importantes dentro de una industria, se debe procurar que, subpoblaciones que son importantes dentro de un fenómeno social, estén adecuadamente representadas dentro de la muestra. Sin embargo, sea por la falta de respuesta o por la imperfección de los marcos de muestreo, eventos que generan subcobertura de algunas subpoblaciones de interés, van en desmedro de la representación de estas subpoblaciones. No obstante, los estimadores de calibración pueden corregir o reducir parte de los efectos que se producen debido a estos errores no muestrales.

En síntesis, los fenómenos sociales de interés: pobreza, empleo, criminalidad, etc., guardan un estrecho vínculo con las características auxiliares de las personas, por lo que estas deben ser incorporadas a través de diseños de muestreo y/o estimadores que así lo permitan, sobre todo cuando subpoblaciones con características auxiliares distintas a las del resto de la población cobran mayor importancia para el fenómeno de interés.

6. Anexo

En este anexo se presenta un breve resumen con los principales resultados de la encuesta realizada dentro de los países que conforman el Grupo de Trabajo. En total participaron 9 países¹⁶ y a continuación se presentan algunos de los resultados de la encuesta, distinguiendo según tipo de encuesta.

Tabla 5: ¿Ajusta factores utilizando información auxiliar?

	Empleo y Fuerza de Trabajo	Ingreso y Gasto	Uso del Tiempo	Pobreza
Si	9	5	3	5
No	0	4	6	4

¹⁶ Argentina, Bolivia, Colombia, Ecuador, México, Perú, República Dominicana y 2 sin identificar.

Tabla 6: ¿Cuál es el método de ajuste utilizando?

	Empleo y Fuerza de Trabajo	Ingreso y Gasto	Uso del Tiempo	Pobreza
Raking	8	2	1	4
Post-Estratificación	1	3	1	1
Otro	0	0	1	0

Tabla 7: ¿Cuál es la fuente de la información auxiliar utilizada?

	Empleo y Fuerza de Trabajo	Ingreso y Gasto	Uso del Tiempo	Pobreza
Proy.Población	8	5	2	5
Proy.Población y est. Escuesta	1	0	1	0

Tabla 8: ¿Considera información específica sobre la población extranjera?

	Empleo y Fuerza de Trabajo	Ingreso y Gasto	Uso del Tiempo	Pobreza
Si	0	0	0	0
No	9	5	3	5

Tabla 9: ¿Considera importante contar con información auxiliar sobre la población extranjera?

	Empleo y Fuerza de Trabajo	Ingreso y Gasto	Uso del Tiempo	Pobreza
Si	5	1	1	2
No	3	3	2	2
No Sabe	1	1	0	1

Referencias

- Aldunate R., Contreras G., de la Huerta C., y Tapia M. (2019) **Characterization of the Recent Immigration to Chile**. *Documentos de Trabajo N° 830*, Banco Central de Chile. Enero, 2019.
- Cassel, C., Särndal, C.E., y Wretman, J. (1976) **Some results on generalized difference estimation and generalized regression estimation for finite populations**. *Biometrika*. Vol. 63, p. 615-620.
- Davies, G. (2018) **Examination of approaches to calibration in survey sampling**. Cardiff University. Thesis for the degree of Doctor of Philosophy.
- Deming, W. y Stephan, F. (1940). **On a least squares adjustment of a sampled frequency table when the expected marginal totals are known**. *Annals of Mathematical Statistics*. Vol. 11, p.427-444.
- Deville, J.C., y Särndal, C.E. (1992) **Calibration Estimators in Survey Sampling**. *Journal of the American Statistical Association*, Vol. 87, No. 418 (June 1992), p. 376-382.
- Deville, J.C., Särndal, C.E., y Sautory, O. (1993) **Generalized Raking Procedures in Survey Sampling**. *Journal of the American Statistical Association*, Vol. 88, No. 423 (Sep. 1992), p. 1013-1020.

- Gutiérrez, A. (2016) **Estrategias de muestreo, diseño de encuestas y estimación de parámetros**. Ediciones de la U, Bogotá.
- Holt, D. y Smith, T. (1979). **Post-stratification**. *Journal of the Royal Statistical Society*. Serie A, Vol. 142, p. 33-46.
- Horvitz, D. y Thompson, D. (1952). **A generalization of sampling without replacement from a finite universe**. *Journal of the American Statistical Association*. Vol. 47, p. 663-685.
- INE-SDV (2018). **Características de la inmigración internacional en Chile, Censo 2017**. Subdepartamento de Demografía y Vitales. Departamento de Demografía y Censos. Instituto Nacional de Estadísticas, noviembre de 2018.
- INE (2018a) **Actualización del Marco Muestral de Manzanas a 2016 en la Encuesta Nacional de Empleo: Diseño muestral prueba piloto**. Documentos de trabajo INE, Serie Metodológico N°4, diciembre de 2018. INE-Chile. Consultar en: https://www.ine.cl/docs/default-source/documentos-de-trabajo/actualizacion_marco_muestral_2016_ene_dise%C3%B1o_muestral_de_la_prueba_piloto.pdf?sfvrsn=f1171920_2.
- _____(2018b) **Actualización del Marco Muestral de Manzanas a 2016 en la Encuesta Nacional de Empleo: Resultados prueba piloto**. Documentos de trabajo INE, Serie Metodológico N°5, diciembre de 2018. INE-Chile. Consultar en: https://www.ine.cl/docs/default-source/documentos-de-trabajo/actualizacion_marco_muestral_2016_ene_resultados_prueba_piloto.pdf?sfvrsn=afe270ee_4.
- _____(2019a) **Actualización del Marco Muestral de Manzanas a 2016 en la Encuesta Nacional de Empleo Aproximación del impacto sobre los principales indicadores de la ENE**. Documentos de trabajo INE, Serie Metodológico N°9, abril de 2019. INE-Chile. Consultar en: https://www.ine.cl/docs/default-source/documentos-de-trabajo/actualizaci%C3%B3n-del-marco-muestral-de-manzanas-a-2016-en-la-encuesta-nacional-de-empleo-aproximaci%C3%B3n-del-impacto-sobre-los-principales-indicadores-de-la-ene.pdf?sfvrsn=a204be57_2.
- _____(2019b) **Fundamentos de la nueva metodología de calibración de los factores de expansión de la Encuesta Nacional de Empleo**. Documentos de trabajo INE, Serie Metodológico N°12, diciembre de 2019. INE-Chile. Consultar en: https://www.ine.cl/docs/default-source/documentos-de-trabajo/documento-de-trabajo-fundamentos-de-la-nueva-calibraci%C3%B3n-de-los-factores-de-expansi%C3%B3n-en-la-ene.pdf?sfvrsn=3de3a0e1_4.
- _____(2019c) **Fundamentos de la nueva metodología de calibración de los factores de expansión de la Encuesta Nacional de Empleo (Datos, programación en R y gráficos complementarios)**. Documentos de trabajo INE, Serie Metodológico N°12, diciembre de 2019. INE-Chile. Consultar en: https://www.ine.cl/docs/default-source/documentos-de-trabajo/documentos-relacionados/documento-fundamento-de-la-nueva-metodolog%C3%ADa-de-calibraci%C3%B3n-de-los-factores-de-expansi%C3%B3n-de-la-encuesta-nacional-de-empleo-datos-programaci%C3%B3n-en-r-y-datos-complementarios.pdf?sfvrsn=fb9f06ab_2.
- _____(2020) **La nueva metodología de calibración de la Encuesta Nacional de Empleo: método de calibración Raking**. Documentos de trabajo INE, Serie Metodológico N°14, marzo de 2020. INE-Chile. Consultar en: https://www.ine.cl/docs/default-source/documentos-de-trabajo/nueva-metodolog%C3%A0a-de-calibraci%C3%B3n-de-la-ene-m%C3%A9todo-de-calibraci%C3%B3n-raking.pdf?sfvrsn=bfed7e52_2.

- INE-DMIE (2020) **Cálculo de la población extranjera en Chile para la calibración de factores de expansión en la Encuesta Nacional de Empleo**. Subdepartamento de Investigación Estadística, Departamento de Metodologías e Innovación Estadística, Julio 2020. INE-Chile. Consultar en: https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/documentos-de-trabajo/documentos-de-trabajo/c\%C3\%A1lculo-de-poblaci\%C3\%B3n-extranjera-en-chile-para-la-calibraci\%C3\%B3n-de-factores-de-expansi\%C3\%B3n-en-la-encuesta-nacional-de-empleo.pdf?sfvrsn=184891be_4.
- INE-ENE (2020) **Documento Metodológico, Encuesta Nacional de Empleo**. Febrero 2020. INE-Chile. Consultar en: https://www.ine.cl/docs/default-source/ocupacion-y-desocupacion/metodologia/espanol/metodolog\%C3\%ADa-encuesta-nacional-de-empleo-ene-2020.pdf?sfvrsn=793380e9_8.
- Kalton, G., y Flores-Cervantes, I. (2003) **Weighting Methods**. *Journal of Official Statistics*, Vol. 19, No. 2, 2003, p. 81-97.
- Kott, P. S. (2006) **Using Calibration Weighting to Adjust for Nonresponse and Coverage Errors**. *Survey Methodology*, Vol. 32, No. 2, (Dec. 2006), p. 133-142.
- Lumley, T. (2010) **Complex Surveys. A guide to analysis using R**. Wiley Series in Survey Methodology.
- _____(2020) **survey: Analysis of Complex Survey Samples**. R package version 3.37. Consultar en: <http://r-survey.r-forge.r-project.org/survey/>.
- Särndal, C.E. (2007) **The calibration approach in survey theory and practice**. *Survey Methodology*, Vol. 33, No. 2, (Dec. 2007), p. 99-119.
- Särndal, C.E., Swensson, B. y Wretman, J. (1989). **The weighted residual technique for estimating the variance of the general regression estimator of the finite population total**. *Biometrika*. Vol. 76, p. 527-537.
- _____(1992). **Model Assisted Survey Sampling**. Springer, New York.
- Silva, P. L. (2004) **Calibration Estimation: When and Why, How Much and How**. *Textos para discussao*, Diretoria de Pesquisas, ISSN 1518-675X, No. 15. Instituto Brasileiro de Geografia e Estatística (IBGE), Rio de Janeiro.
- Stephan, F. (1942). **An iterative method of adjusting sample frequency data tables when expected marginal totals are known**. *Annals of Mathematical Statistics*. Vol. 13, p. 166–178.
- Tillé, Yves (2020) **Sampling and Estimation from Finite Populations**. Wiley Series in Probability and Statistics.
- United Nations (UN). (1982) **Designing Household Survey Samples: Practical Guidelines**. ST/ESA/STAT/SER.F/98, Statistics Division, United Nations, New York.
- Valliant, R. (2004) **The Effect of Multiple Weighting Steps on Variance Estimation**. *Journal of Official Statistics*, Vol. 20, No. 1, 2004, p. 1–18.
- Zhang, Li-Chun. (2000) **Post-Stratification and Calibration - A Synthesis**. *The American Statistician*, Vol. 54, No. 3 (Aug. 2000), p. 178-184.