



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

## Ficha técnica y metodología de estimación de la ECH 2020<sup>1</sup>

### 1. Objetivo

En este documento se presenta la ficha técnica de la Encuesta Continua de Hogares (ECH) 2020 y la metodología utilizada para estimar distintos indicadores de interés para distintas periodicidades (e.g mensuales y anuales) junto con sus precisiones. En línea con esto último, se presentan los márgenes de error de los principales indicadores incluidos en el *boletín técnico de Estimación de la pobreza por el método del ingreso*.

A su vez, pretende ser una guía para los usuarios que utilizan los microdatos de la ECH ya que la encuesta en el contexto de la emergencia sanitaria adaptó su metodología y operativa para que el Instituto Nacional de Estadística (INE) siguiera brindando, estimaciones de calidad en los indicadores más importantes que se obtienen a partir de la ECH.

Si bien la ECH históricamente ha sido una encuesta multipropósito, para el 2020 solo cuenta con información relativa a las principales características sociodemográficas de las personas (sexo y edad), mercado laboral y los ingresos personales y del hogar.

### 2. Introducción

La ECH es una encuesta con periodicidad mensual, en donde, previo a la emergencia sanitaria producto del COVID-19, los hogares eran seleccionados al azar bajo un diseño complejo (estratificado, por conglomerados y en dos etapas de selección) para participar un mes en el año (i.e. diseño cross-section) y las entrevistas eran llevadas a cabo de forma presencial.

A partir de marzo 2020, la ECH pasó a ser una encuesta de paneles rotativos, en donde los hogares son elegidos al azar utilizando casos respondientes de la ECH 2019 hasta febrero de 2020; es decir, la ECH no presencial utiliza un diseño en dos fases de selección<sup>2</sup>. A su vez, dichos hogares participan de la encuesta por tres meses consecutivos y el relevamiento es llevado a

---

<sup>1</sup> Si bien existe una clara diferencia entre estimador y estimación, en este documento se van a utilizar ambos términos de forma intercambiable para poder facilitar la lectura y evitar frases del estilo: “el estimador de la varianza del estimador”.

<sup>2</sup> Por más detalles sobre la metodología acceda al siguiente link: [Metodología ECH no presencial 2020](#).



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

cabo de forma telefónica para un set reducido de variables (estructura del hogar, mercado de trabajo e ingresos). Lo anterior implica que la muestra de un mes en particular está compuesta por tres paneles o grupos de rotación (GR), en donde cada GR es una muestra representativa de la población. Este nuevo diseño es más eficiente (menores errores estándar) para las estimaciones del cambio neto de los indicadores de interés (e.g. mercado de trabajo e ingresos) en comparación a los diseños cross-section, debido a la correlación existente entre las estimaciones de un mes a otro producto de la dependencia parcial entre muestras (proporción de solapamiento). Es importante tener en cuenta que este tipo de diseño tiene como objetivo mejorar las precisiones de cambios netos y no realizar análisis longitudinales (e.g. transición de los hogares en distintos estatus mientras el hogar y sus integrantes permanecen en la muestra).

### **3. Condición de elegibilidad**

Para ser elegible a fin de participar en la ECH, el hogar debe residir en viviendas particulares ubicadas en cualquier parte del territorio nacional.

### **4. Tamaños de muestra**

El tamaño de muestra efectivo mensual promedio se situó en 4880 hogares aproximadamente. Las tasas de elegibilidad y respuesta fueron estables durante el año; y se situaron en promedio en 84% y 82%, respectivamente.

### **5. Metodología de ponderación**

Teniendo en cuenta que durante el año 2020 sólo se computaron estimaciones mensuales referentes al mercado de trabajo e ingresos, se realizaron únicamente sistemas de ponderadores mensuales. La metodología de ponderación utilizada para la construcción de los ponderadores mensuales depende de si la ECH sigue un diseño cross-section (enero y febrero) o con panel rotativo (marzo a diciembre).

Para el caso de los ponderadores de la ECH cross-section los mismos fueron calculados siguiendo la metodología tradicional de la ECH pre-covid, esto es: los ponderadores basados en el diseño (inverso de la probabilidad de selección de la vivienda, siguiendo el diseño muestral) son ajustados en una primera instancia por no respuesta (asumiendo un modelo común de medias a nivel de estrato de diseño) y luego ajustados a conteos demográficos provenientes de las proyecciones de población por medio de estimadores de regresión/calibración; empleando como



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

variables auxiliares para la modelación/calibración variables dicotómicas pertenecientes a distintos grupos demográficos (sexo, tramo de edad y departamento); y utilizando el método de integración para producir ponderadores constantes para todos los miembros de un mismo hogar.

Por lo tanto, el ponderador del hogar  $i$  perteneciente a la muestra efectiva del mes  $t$  viene dado por:

$$w_{it} = d_i \times \hat{\phi}_h \times g_i ,$$

donde  $d_i$  es el ponderador original del hogar  $i$ ,  $\hat{\phi}_h$  es la propensión estimada de respuesta en el estrato  $h$ , la cual, bajo el modelo de no respuesta utilizado, queda definida como la tasa de respuesta ponderada en el estrato  $h$  y  $g_i$  es el ajuste proveniente de la calibración/regresión para el hogar  $i$ .

Por otra parte, para los ponderadores de la ECH no presencial (marzo a diciembre) la metodología de cálculo fue más compleja, producto de que el diseño implementado fue en dos fases de selección (i.e. sub muestra) y se cuenta con un vector de covariables conocidas  $x$  para los no respondientes (provenientes de la ECH presencial), que permite estimar de forma más sofisticada la propensión que tienen los hogares en participar en la ECH no presencial, en comparación con el modelo de media de grupos que se utiliza de forma tradicional en la ECH presencial, el cual, es llevado a cabo porque en las encuestas presenciales la única información auxiliar de los hogares no respondientes que se cuenta es aquella proveniente del marco de muestreo (e.g. estratos, unidades primarias de muestreo <UPM> y regiones geográficas).

En este caso, el ponderador en el mes  $t$  para el hogar  $i$  viene dado por:

$$w_{it} = w_i^* \times d_{2i} \times \hat{\phi}_i \times g_i$$

donde  $w_i^*$  es el peso final del hogar  $i$  proveniente de la ECH presencial, el cual, fue ajustado por no respuesta y calibrado a conteos poblacionales,  $d_{2i}$  es el ponderador original de la muestra telefónica (fase 2) condicionado a la muestra presencial (fase 1), el cual se encuentra determinado por el inverso de la probabilidad de inclusión del hogar en la fase 2 condicionado a que fue seleccionado (y respondió) en la fase 1,  $\hat{\phi}_i$  es la propensión estimada del hogar  $i$  la cual es estimada utilizando *random forest* como algoritmo para modelar la propensión de respuesta de los hogares; y considerando como variables de entrada  $x$  para la construcción del mismo, información proveniente de la ECH presencial. Finalmente,  $g_i$  es el ajuste por calibración, el cual, es llevado a cabo de la misma forma que en el caso de los ponderadores mensuales de la ECH presencial. Más detalle sobre la construcción de los ponderadores de la ECH no presencial se encuentran en el siguiente link: [Metodología ECH no presencial 2020](#).



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

## 6. Métodos de estimación mensuales y para otros períodos de tiempo

Como en toda encuesta por muestreo, el objetivo es estimar (aproximar) parámetros  $\theta$  por medio de una muestra aleatoria. El parámetro  $\theta$  puede ser definido como función de una o varias variables que son relevadas en la encuesta, es decir,  $\theta = f(\mathbf{y})$  en donde  $\mathbf{y}$  es el vector de variables de interés. La estimación de un parámetro cualquiera para el mes  $t$  el cual se denota como  $\theta_t$  es calculado utilizando los ponderadores mensuales  $w_t$  definidos en el punto anterior

$$\hat{\theta}_t = \hat{f}(w_t \times \mathbf{y})$$

Ejemplos de los parámetros  $\theta_t$  pueden ser aquellos utilizados para describir el mercado de trabajo (e.g. total de ocupados, tasa de desempleo y empleo) o medidas descriptivas de los ingresos, como ser, la media o algún estadístico de posición (e.g. mediana). Por ejemplo, si el interés recae en estimar el total de una variable cualquiera  $y$  en el mes  $t$ , el estimador queda definido como:

$$\hat{Y}_t = \sum_{i \in s_t} w_{it} \times y_{it},$$

donde  $y_{it}$  es el valor que toma la unidad  $i$  en el mes de referencia  $t$  y  $s_t$  denota la muestra de hogares (o personas) en el mes de referencia  $t$ .

Por otra parte, las estimaciones de algunos parámetros de interés pueden ser necesarios computarlos para intervalos de tiempo mayores a la periodicidad de la ECH. Un ejemplo sencillo de un parámetro es la proporción de hogares por debajo de la línea de pobreza (LP) en el año

$$\theta = N^{-1} \times \sum_{i \in U} y_i,$$

donde  $y$  es un variable dicotómica (1= pobre, 0= no pobre),  $y_i$  es el status de pobreza del hogar  $i$ ,  $N$  es la cantidad de hogares en la población y  $U$  denota la población o universo de elegibles para participar en la ECH. Siguiendo el mismo ejemplo y teniendo en cuenta la periodicidad de la encuesta así como el periodo de referencia de relevamiento de los datos de ingresos, la proporción  $p$  de hogares por debajo de la LP es construida como el promedio de las proporciones mensuales, es decir,



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

$$p = T^{-1} \times \sum_{t=1}^T p_t$$

donde  $p_t$  es la proporción de hogares por debajo de la LP en el mes  $t$  y  $T$  es la cantidad de meses que se utilizan para computar el parámetro, en este caso,  $T = 12$ .

Como ya se dijo, hasta el año 2019 la ECH seguía un diseño cross-section y las estimaciones anuales eran computadas calculando un set de ponderadores anuales  $w_i \forall i \in s$  haciendo un pool con las encuestas mensuales ( $s = \cup_{t=1}^{12} s_t$ ); y siguiendo la misma metodología utilizada para el cálculo de los ponderadores mensuales  $w_t$ . Lo anterior era posible ya que las tasas de muestreo en cada uno de los estratos de diseño de forma mensual eran cercanas a cero y por ende las probabilidades de selección de las UPM no se veían influidas y se podía asumir independencia entre muestras (y entre estimadores). Este sistema de ponderadores anuales, lograba simplificar el cómputo de las estimaciones a los usuarios externos debido a que las mismas podían ser realizadas utilizando cualquier software que maneje datos provenientes de encuestas complejas (e.g. R o Stata). De todas formas, el resultado obtenido para las distintas estimaciones de los parámetros, era similar al promedio de las estimaciones mensuales (utilizando el sistema de ponderadores  $w_t$ ) para aquellas áreas o dominios<sup>3</sup> de estimación con tamaños de muestra mensuales similares o con poca variabilidad a lo largo del tiempo. Es decir, para la estimación de un parámetro cualquiera  $\theta$  se cumple aproximadamente que:

$$\hat{\theta} = \hat{f}(w \times \mathbf{y}) = T^{-1} \times \sum_{t=1}^T \hat{f}_t(w_t \times \mathbf{y})$$

Teniendo en cuenta lo anterior, las estimaciones producidas para parámetros que tiene una periodicidad mayor a la mensual (e.g. trimestral, semestral y anual) y que por ende deben utilizar información de varios meses de la encuesta, deben ser computados promediando las estimaciones mensuales para las cuales deben utilizarse el sistema de ponderadores  $w_t$  y por último, estas estimaciones, deben promediarse según la periodicidad que requiera el parámetro en cuestión. A modo de ejemplo, la estimación del total de una variable cualquiera  $y$  para el año puede computarse de las siguientes dos formas:

$$\hat{Y} = \sum_{i \in s} w_i \times y_i ,$$

o

<sup>3</sup> Se denomina dominio o área de estimación a un subconjunto de la población para el cual se quieren obtener estimaciones de distintos indicadores (parámetros).



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

$$\hat{Y} = T^{-1} \times \sum_{t=1}^T \hat{Y}_t = \hat{Y} = T^{-1} \times \sum_{t=1}^T \sum_{i \in s_t} w_{it} \times y_{it}$$

Es importante tener en cuenta que dicha estrategia puede ser utilizada para la estimación de distintos parámetros o para otras áreas de estimación, en donde, en el proceso de cálculo de las estimaciones se deben utilizar únicamente las unidades que pertenecen al dominio o área de estimación. Esta metodología de cálculo mencionada anteriormente fue utilizada por el INE para realizar las estimaciones presentadas en el *boletín técnico de Estimación de la pobreza por el método del ingreso*. Por ejemplo, la estimación de la proporción de hogares por debajo de la LP en el año 2020 a nivel total país ( $U$ ) viene dada por:

$$\hat{p} = T^{-1} \times \sum_{t=1}^T \hat{p}_t$$

donde  $\hat{p}_t$  es la estimación de la proporción de hogares por debajo de la LP en el mes  $t$  y la misma viene dada por

$$\hat{p}_t = (\sum_{i \in s_t} w_{it})^{-1} \times (\sum_{i \in s_t} w_{it} \times y_{it})$$

donde  $y_{it}$  es una variable dicotómica que indica el status de pobreza del hogar  $i$  en el mes de referencia  $t$ , y  $s_t$  indica la muestra efectiva en el mes  $t$  a nivel de hogares.

De igual forma, la estimación del Índice de Gini ( $G$ ) a nivel de toda la población para el año está dada por:

$$\hat{G} = T^{-1} \times \sum_{t=1}^T \hat{G}_t = T^{-1} \times \sum_{t=1}^T \left( \frac{2 \sum_{i=1}^{n_t} (w_i y_{it} \sum_{j=1}^i w_{jt}) - \sum_{i=1}^{n_t} w_{it}^2 y_{it}}{(\sum_{i=1}^{n_t} w_{it})(\sum_{i=1}^{n_t} w_{it} y_{it})} \right)$$

donde  $y_{it}$  es el ingreso per cápita sin valor locativo a precios constantes del hogar  $i$  en el mes  $t$  y  $n_t$  es el tamaño de muestra efectiva en el mes  $t$ .

## 7. Precisiones

El INE siguiendo las recomendaciones internacionales computa márgenes de error (moe) para las estimaciones de los principales indicadores que produce, de forma de brindar a los usuarios, una medida de calidad de las mismas y en base a ellas puedan interpretar los resultados de forma correcta. El insumo principal para el cómputo de los moe es el error estándar del estimador (SE). Como en toda encuesta por muestreo que sigue un diseño muestral complejo, el objetivo es



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

encontrar una estimación del SE que refleje todas las fuentes de variabilidad del estimador, las cuales se encuentran determinadas por el diseño muestral, así como por los distintos ajustes que son llevados a cabo para obtener los ponderadores finales.

De forma general, el moe al 95% de confianza para la estimación de un parámetro cualquiera  $\theta$  viene dado por:

$$\text{moe}(\hat{\theta}) = 1.96 \times \widehat{\text{SE}}(\hat{\theta})$$

donde 1.96 es el valor de una normal estándar que acumula un 0.975 de probabilidad y  $\widehat{\text{SE}}(\hat{\theta})$  es la estimación del SE del estimador  $\hat{\theta}$ .

Para el cómputo de los SEs se utilizan distintas técnicas dependiendo de la complejidad del parámetro en cuestión y la periodicidad del mismo. Para el caso de los indicadores mensuales, el SE de los estimadores de totales, medias y ratios se utiliza el método del último conglomerado junto con la linearización de Taylor. En tanto, para parámetros más complejos que no pueden ser aproximado por linearización de Taylor (e.g. Índice de Gini) se utilizan métodos de remuestreo (e.g. Bootstrap) para el cómputo de los SE.

De todas formas, sin importar el método utilizado para computar los SEs de forma mensual, para el cómputo de los SE de aquellos indicadores que exceden la periodicidad mensual (e.g. proporción de hogares pobres), se realiza el supuesto de que las muestras mensuales  $s_t$  son independientes y por ende los estimadores mensuales son independientes entre sí, es decir, la covarianza entre los estimadores de un mes a otro es cero. Esto es cierto únicamente para aquellas muestras que se encuentran distanciadas por un periodo mayor al sistema de rotación implementado en la ECH no presencial. Para los casos de dos muestras consecutivas, es esperable que la covarianza de los estimadores sea positiva producto del porcentaje de solapamiento entre muestras y en la estabilidad de los indicadores de mercado de trabajo e ingresos. Es importante aclarar que si bien existen formas de poder introducir al menos parcialmente en el cómputo de los SE el solapamiento de los distintos paneles o GR por medio de técnica de *re-muestreo coordinadas* (i.e. se preservan las unidades primarias de muestreo incluidas en cada réplica bootstrap de un mes a otro para incluir el solapamiento en la estimación del SE) el efecto en los estimadores de SEs de parámetros anuales, los cuales, exceden ampliamente el periodo de rotación de la muestra, no amerita en una primera instancia su implementación.



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

Por lo tanto, la estimación del SE de un parámetro cualquiera  $\theta$  que es computado para un período que comprende  $T$  meses consecutivos (e.g. tasa de desempleo trimestral, pobreza anual, etc.) queda definido como:

$$\widehat{SE}(\hat{\theta}) = T^{-1} \sqrt{\sum_{t=1}^T \widehat{SE}^2(\hat{\theta}_t)}$$

en donde  $\widehat{SE}^2(\hat{\theta}_t)$  es calculado por medio de Bootstrap Rao-Wu o el método del último conglomerado dependiendo de la complejidad del parámetro.

Siguiendo la metodología descrita anteriormente a continuación se presentan los márgenes de error para los principales indicadores incluidos en el *boletín técnico de Estimación de la pobreza por el método del ingreso*

En el siguiente cuadro se presentan los márgenes de error al 95% para la incidencia de la pobreza en hogares, personas y el Índice de Gini, por región geográfica.

Área de estimación	márgenes de error al 95%		
	Incidencia de la Pobreza en hogares (en %)	Incidencia de la Pobreza en personas (en %)	Índice de Gini
Total país	0.3%	0.4%	0.0359
Total país (5000 + hab)	0.3%	0.5%	0.0034
Montevideo	0.5%	0.8%	0.0050
Interior. Loc de 5000+ hab	0.4%	0.6%	0.0043
Interior. Loc de menos 5000 hab	0.8%	1.3%	0.0108
Zonas rurales	0.7%	1.0%	0.0114

## 8. Recolección de la información<sup>4</sup>

<sup>4</sup> Por mayor información sobre los cambios en la operativa e instrumento de relevamiento durante la ECH telefónica, dirigirse [Cambios en la operativa y formulario de la Encuesta Continua de Hogares a partir de la emergencia sanitaria](#)





División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

## Tipo de entrevista e informante

Durante el 2020 la entrevista fue realizada de forma presencial solamente en enero y febrero. A partir de marzo, se comienza a relevar de forma telefónica desde los hogares de los entrevistadores. La ECH no presencial solamente relevó información del mercado laboral, ingresos y datos sobre las características sociodemográficas de las personas. El informante, tanto de la ECH presencial como telefónica, es un miembro del hogar de 18 años o más (excepto el servicio doméstico) competente para brindar información acerca del resto de los integrantes del hogar.

### 9. Codificadores

Para las preguntas referidas a la ocupación de la persona se utiliza el CIUO-08 y la CIU Revisión 4 para las preguntas referidas a rama de actividad de la empresa o establecimiento.

### 10. Conceptos de mercado de trabajo

Cabe mencionar que la ECH telefónica mantiene, los mismos conceptos de mercado laboral e ingresos que los utilizados en la ECH habitual que se realizaba de forma presencial.

#### Límite etario

La investigación sobre las características de la mano de obra y los ingresos percibidos se aplica únicamente a las personas de 14 o más años de edad. Este umbral coincide con la edad que define a la Población Económicamente Activa (PEA) en el país.

#### Activo

Persona de 14 o más años de edad, que tiene al menos una ocupación en la que vierte su esfuerzo productivo a la sociedad, o que, sin tenerla, la busca activamente durante el período de referencia elegido para la encuesta. Este grupo incluye la fuerza de trabajo civil y los efectivos de las fuerzas armadas.

#### Ocupado

---



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

Persona que trabajó por lo menos una hora durante el período de referencia de la encuesta, o que no lo hizo por estar de vacaciones, por enfermedad, por accidente, por conflicto de trabajo o interrupción de este a causa del mal tiempo, por averías producidas en las maquinarias o por falta de materias primas, pero tiene empleo al que volverá con seguridad. Se incluyen en esta categoría a los trabajadores familiares no remunerados.

### **Ocupado Ausente**

Son aquellos trabajadores que sin haber trabajado la semana pasada mantienen un vínculo con su puesto de trabajo. La razón más frecuente de ausencia de los ocupados responde a licencias (vacaciones, enfermedad, estudio, etc.). Para otro tipo de razones de ausencia (como poco trabajo, suspensión, conflicto laboral, seguro de desempleo), se valora la fuerza del vínculo con el puesto de trabajo en función de si continúan percibiendo ingresos por su trabajo o si el plazo de retorno a su puesto es menor o igual a los tres meses. Dado el contexto de emergencia sanitaria, se agregó la condición de volver a su puesto de trabajo una vez levantadas las medidas de confinamiento. En caso de cumplir con alguna de estas tres últimas condiciones, se considera a la persona, ocupada ausente temporalmente de su trabajo.

### **Desocupado**

Persona que durante el período de referencia no tenía trabajo, buscaba un trabajo remunerado y estaba disponible en ese momento para comenzar a trabajar. Esta categoría comprende a las personas que trabajaron antes pero perdieron su empleo (desocupados propiamente dichos), desempleados en seguro de paro y aquellas que buscan su primer trabajo.

### **Inactivo**

Persona que no trabaja ni tiene empleo y no desea o no está disponible para trabajar. Se clasifica en las siguientes categorías: personas que se ocupan solamente del cuidado de su hogar, estudiantes, jubilados, pensionistas o rentistas.

## **11. Conceptos de ingresos**

La ECH releva ingresos personales y del hogar, incorporando el criterio de la periodicidad regular de cada partida, para que pueda ser considerada como ingreso corriente (no extraordinario). Los ingresos personales que se relevan se clasifican en:



División Estadísticas Sociodemográficas  
Departamento Encuesta Continua de Hogares

- Ingresos por trabajo: dependiente e independiente, en dinero y en especie (incluye ingreso por Fonasa).
- Ingresos por transferencias: incluye los ingresos percibidos provenientes de instituciones públicas o privadas (en dinero y en especie) así como contribuciones por divorcio o separación que se reciben de otros hogares.
- Otros ingresos personales: se incluye específicamente la devolución de Fonasa y todo otro ingreso corriente.

En cuanto a los ingresos del hogar, las partidas relevadas son:

- Transferencias de otros hogares (en dinero y en especie).
- Rentas de la propiedad de activos físicos y financieros.
- Ingresos por dividendos y utilidades de negocios en los que la persona no trabaja.
- Indemnización por despido.
- Valor locativo (para los propietarios y en áreas urbanas).

## 12. Microdatos de la ECH 2020

Los microdatos de la ECH 2020 corresponde a los datos de la ECH presencial (enero y febrero) y la telefónica (marzo a diciembre). Las variables que contiene esta base de datos son todas aquellas que permiten reproducir los indicadores y aperturas presentados en los informes mensuales de coyuntura, en el boletín de pobreza 2020 y en boletín técnico adjunto a la base de datos 2020.

En el correr del año se publicará la base de datos con los microdatos del resto de las variables que son comunes en los cuestionarios de la encuesta presencial y telefónica, relacionadas a las características sociodemográficas de las personas, al mercado laboral y los ingresos (personales y del hogar). Debido a la incorporación de una encuesta panel y a los cambios operativos y tecnológicos tanto en el relevamiento como en el proceso de revisión de inconsistencias, es necesario una depuración más exhaustiva aún no culminada.