1. Read the EPS documentation. This is the first time we work with this data, and we would like to have a brief document describing the data. Key elements to highlight:
2. Coverage: how many individuals are followed across the years. Is this a rolling panel?
3. Representativeness.
4. Definition of informality.
5. Any limitation this data source can have.

https://www.previsionsocial.gob.cl/sps/biblioteca/encuesta-de-proteccion-social/documentos-eps/

¿Cómo se tomó las muestras de la Encuesta de Protección Social (EPS)?:

* Se destaca que la EPS es la mayor encuesta panel de Chile, con una muestra aproximada de **20.000 personas distribuidas en todas las regiones del país.** Las respuestas se dividieron en 11 bases de datos distintas. Estas bases de datos poseen una **variable llamada Folio** y cada uno de estos Folios representa las respuestas de un individuo en particular, es de tipo float de largo 9. A través de este Folio se puede realizar la unión longitudinal con las versiones anteriores de la EPS, o también se puede efectuar la unión con los factores de expansión correspondientes.

**Las incorporaciones de muestra se dieron en las olas de 2002, 2004 y 2015:**

Las cuatro muestras de la EPS corresponden a:

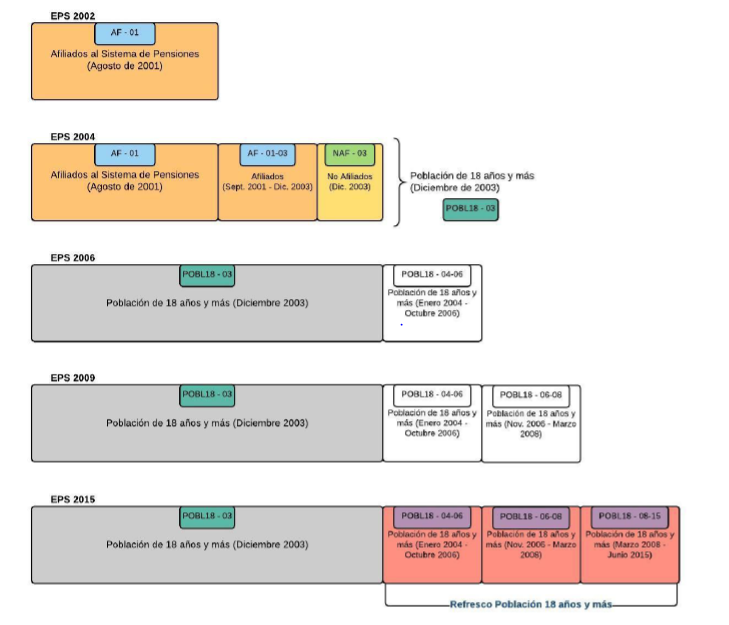
1. Afiliados 2002 – EPS I Ronda

2. Nuevos Afiliados 2004 – EPS II Ronda

3. No afiliados 2004 – EPS II Ronda

4. Nuevos mayores de 18 años y más (entre 2004 y 2015) 2015 – EPS VI Ronda

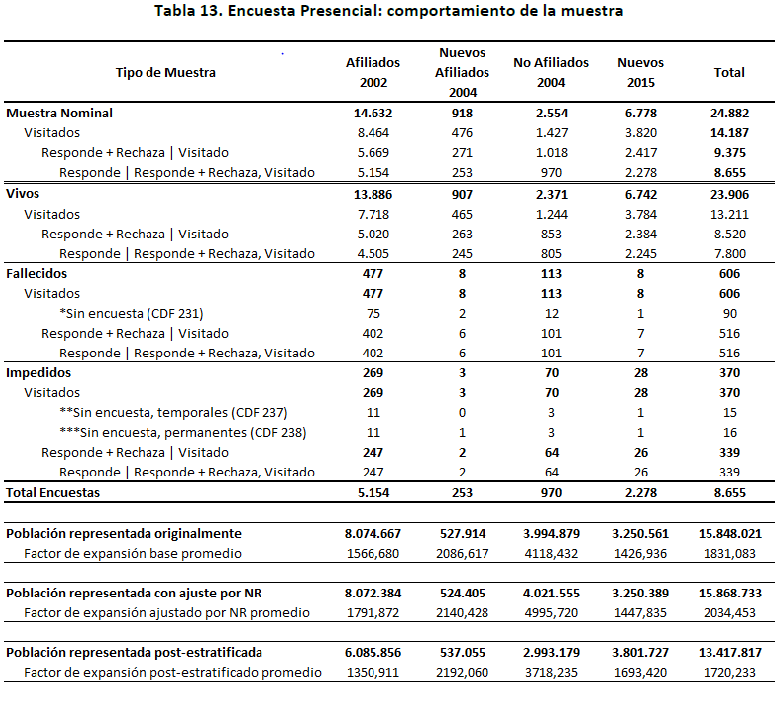
* Encuesta longitudinal del tipo panel fijo (seguimiento) más nacimientos (nuevos individuos que cumplen 18), donde la unidad de análisis son los residentes en Chile que tengan 18 años o más. Después de la selección inicial de la muestra, no se realizan adiciones a la muestra original exceptuando muestras con “nacimientos” de la población. (permite representar a la población de corte transversal en cada ronda de recolección de datos además de la del panel). En las rondas 2006 y 2009, no se hizo refresco de la muestra (nacimientos), no sirven para obtener estadísticas de corte transversal representativas de la población de 18 años y más del país. Si bien, el año 2012, se hizo un refresco éste no cumplió con las características de diseño de la EPS: fue necesario descartarlo y no se utilizó.
* 2002 (Representatividad de solo afiliados (17,246)): La encuesta fue aplicada por primera vez el año 2002 con representatividad a nivel de afiliados al sistema de pensiones (AFP o INP)
* 2004 (representatividad de afiliados y no afiliados (tiene a los 17,246 iniciales más 3,830 nuevos de los cuales 940 son nuevos afiliados y 2,860 son no afiliados)): Su segunda ronda a fines de 2004 y comienzos de 2005, en la cual una nueva muestra de no afiliados al sistema de pensiones fue incorporada junto con una actualización de los afiliados al sistema de AFP, para de esta forma en el año 2004 representar toda la población de 18 años y más. En la primera ronda los no afiliados no estaban incorporados, por lo cual en el año 2002 este estudio no tiene representatividad nacional. Dada la importancia de conocer quiénes son las personas que quedan marginadas del sistema de pensiones, en el año 2004 se incorporó una muestra de no afiliados al sistema de pensiones, y a su vez de actualizó la muestra de afiliados con los nuevos afiliados al sistema de pensiones. Así, la Encuesta de Protección Social 2004 tiene representatividad nacional (excluyendo las fuerzas armadas y carabineros quienes tienen su propio sistema previsional).
* 2006 (representatividad de afiliados y no afiliados): Una tercera ronda fue levantada a fines del año 2006 y comienzos del 2007.
* 2009 (representatividad de afiliados y no afiliados): : La cuarta ronda del estudio fue levantada entre Abril y Diciembre de 2009
* 2012 (base con problemas de muestreo, no recomiendan uso de datos): La quinta ronda fue levantada entre Octubre de 2012 y Agosto de 2013, sin embargo esta ronda fue publicada con reparos debido a los malos resultados del trabajo de campo (baja tasa de respuesta) y a la falta de consistencia con rondas anteriores en la selección de la muestra de refresco.
* 2016 (representatividad de afiliados y no afiliados): Entre los meses de abril y julio del año 2016 fue levantada la sexta ronda de la encuesta, donde se intentó contactar y entrevistar a todas las personas pertenecientes al panel, independientemente de la última ronda en que hayan sido entrevistados, y además se seleccionó una muestra de refresco para incorporar todas las personas que cumplieron 18 años desde la última actualización de la muestra en el año 2004.
  + Muestra de la VI Ronda EPS:
    - 18.649 personas del Panel seleccionadas en el año 2002 o en el año 2004, y que fueron entrevistadas al menos en una ronda de la EPS
      * 15.060 “Afiliados 2002”
      * 2.673 “No Afiliado 2004”
      * 916 “Nuevo Afiliado 2004”

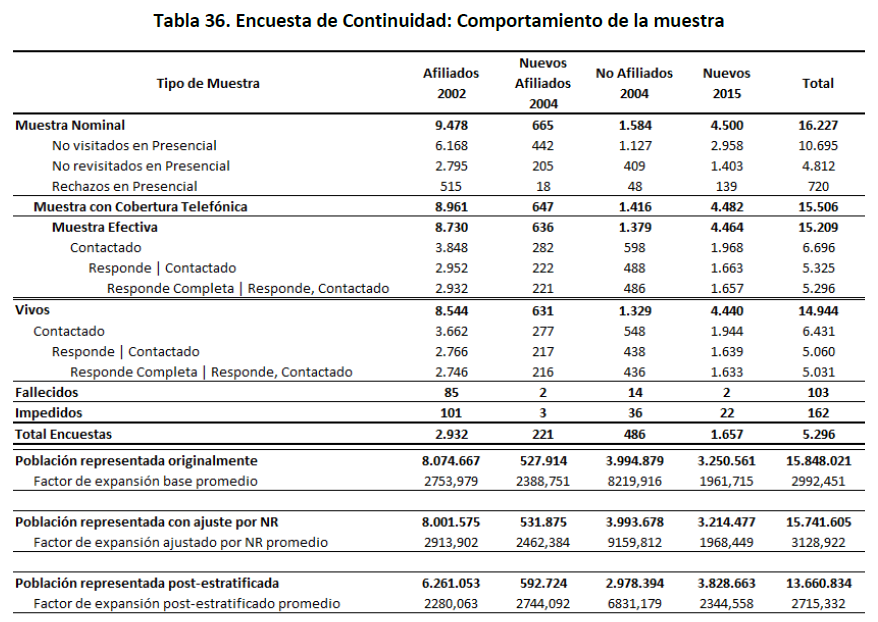


* 2020: No se incluyó refresco ni actualización del marco muestral y sólo se buscó fortalecer el vínculo con el panel. Total **de 24.882 observaciones disponibles para la VII Ronda de la EPS.**

Por lo tanto, la EPS VII Ronda queda finalmente conformada por tres encuestas:

1. **Encuesta Presencial:** 8.655 casos obtenidos de forma presencial entre los meses de diciembre de 2019 y marzo de 2020.
2. **Encuesta de Continuidad:** encuesta telefónica realizada a quienes no respondieron la encuesta Presencial, con un total de 5.296 entrevistas completas.
3. **Encuesta de Re-Entrevista:** encuesta telefónica realizada a quienes respondieron la encuesta Presencial, con un total de 2.225 entrevistas completas.





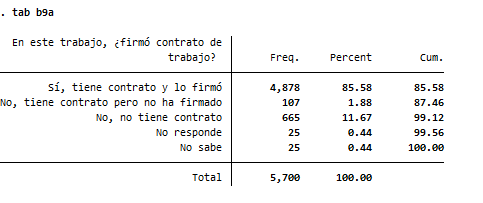
**Variables que se necesitan para replicar el ejercicio con la ELCA:** (El módulo de historia laboral es de nivel individuo-trabajo, es decir, el individuo aparece en la base varias veces según el número de trabajos en su vida)

ELCA: (2010,2013,2016) Asalariados trabajando urbanos de 15 a 64 años. (19,696 personas)

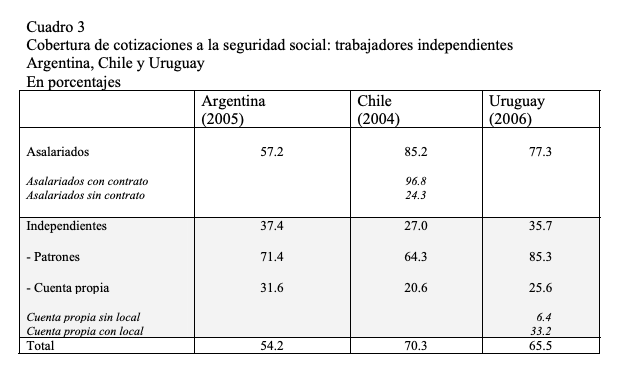
EPS: (2002 (solo afiliados), 2004, 2006, 2009, 2015, 2018, 2020) Trabajadores mayores de 18 años en todas las regiones de Chile.

**Folio: identificador en el tiempo del individuo**

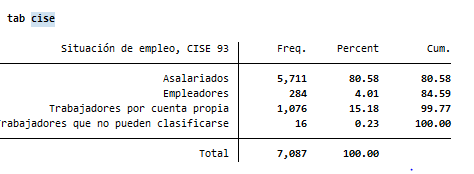
* **Informalidad:** Modulo de historia laboral: b9a (En este trabajo, ¿firmó contrato de trabajo?) y b18 (En este periodo, ¿Se encontraba cotizando en algún sistema previsional?)



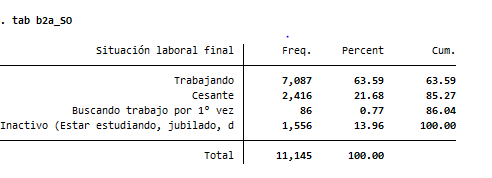
Paper que usa con y sin contrato en Chile como medida de informalidad.



* **Edad:** Modulo entrevistado:a9 (Edad)
* **Nivel educativo** Modulo entrevistado:a12n (categórica último nivel educativo aprobado)
* **Salario** Modulo de historia laboral: b12 (VAR\_ed: ¿Cuál fue el ingreso líquido mensual promedio en su trabajo?) y b13 (¿Cuántas horas semanales trabajaba en este empleo?)
* **Edad del primer trabajo** Modulo de historia laboral: b2a\_inicio\_a (Inicio del periodo laboral: Año) y b2a\_termino\_ax (Fin del periodo laboral: Año). Buscar el mínimo por individuo.
* **Ocupación (dejar solo a los asalariados, sacando a los que trabajan por cuenta propia)** Modulo de historia laboral: cise (situación de empleo)



* **Urbano** Modulo entrevistado: a6a (¿En qué lugar nació usted?) y Modulo de historia laboral: b4(¿En qué región trabajaba?) 15 regiones.
* **Trabajando actualmente** Modulo de historia laboral: b2a\_SO (situación laboral final)



**5. Factores de expansión longitudinales**

**5.1. Factores de expansión longitudinales de la encuesta Presencial**

Dado que la EPS ha acumulado 7 rondas, las combinaciones de muestras longitudinales que se pueden generar son abundantes. Por lo tanto, es necesario decidir qué muestras longitudinales pueden ser interesantes de estudiar. Para ellas, se deben crear los factores de expansión longitudinales correspondientes.

En el presente contexto, sabiendo que las incorporaciones de muestra se dieron en las olas de 2002, 2004 y 2015, se propuso generar los siguientes factores longitudinales:

1. Factor longitudinal para la muestra presente en 2002-2004-2006-2009-2015-2019

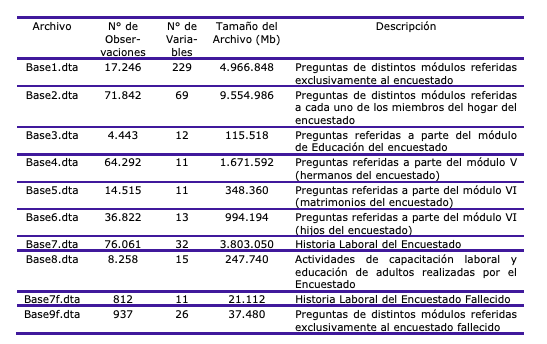
2. Factor longitudinal para la muestra presente en 2004-2006-2009-2015-2019

3. Factor longitudinal para la muestra presente en 2015-2019

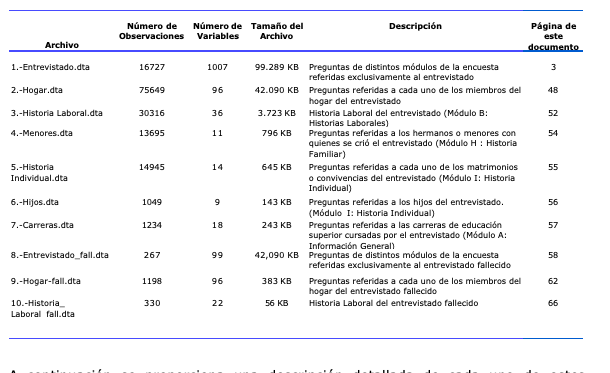
Para generar los factores longitudinales, cada una de las sub-muestras originales es expandida para representar la población total que le corresponda (Participantes en 2002, Participantes en 2004, Participantes en 2015). Los factores, además, son corregidos por la propensión marginal a participar en cada ola, es decir, Pr(𝑝𝑎𝑟𝑡𝑖𝑐𝑖𝑝𝑎𝑟 𝑒𝑛 𝑜𝑙𝑎 𝑊), donde 𝑊 es la ola particular o conjunto de olas, según corresponda.

**Variables/bases por año:**

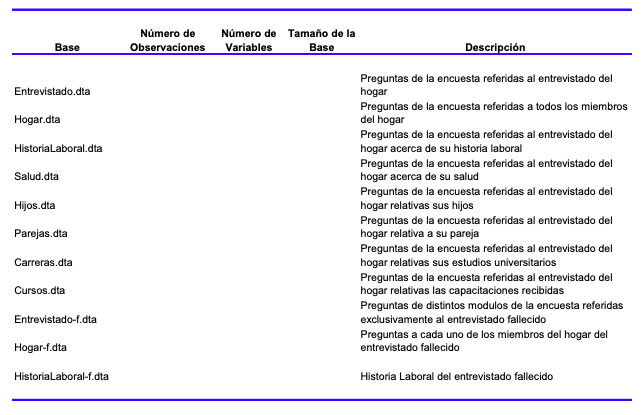
* **2002:https://www.previsionsocial.gob.cl/sps/download/estudios-previsionales/encuesta-de-proteccion-social/documentos-eps/documentos-eps-2002/instrucciones-eps-2002.pdf**

****

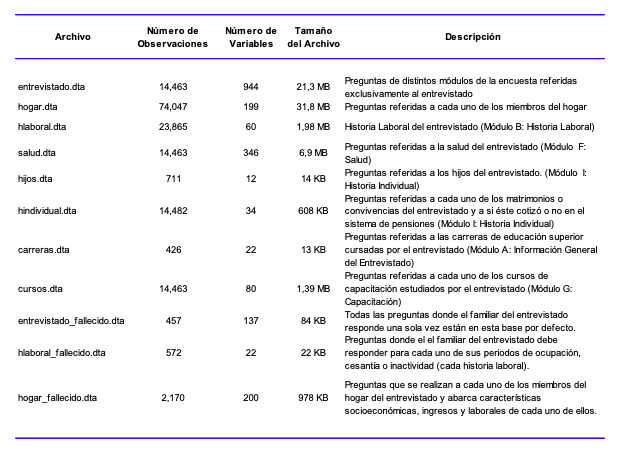
**2004:**

****

**2006**

****

**2009:**

****