

Documentos de trabajo

**Reclasificación de la no respuesta:
distinción entre la no respuesta al
ítem y la no respuesta a la unidad**

Autoras:

Verónica Canales

Camila Mena

Javiera Quinteros



INSTITUTO NACIONAL DE ESTADÍSTICAS

Morandé 801, Santiago de Chile

Teléfono: 562 3246 1000

ine@ine.cl

Facebook: @ChileINE

Twitter: @INE_Chile

Los Documentos de Trabajo del INE están dirigidos a investigadores, académicos, estudiantes y público especializado en materias económicas, y tienen como objetivo proporcionar un análisis exhaustivo sobre aspectos conceptuales, analíticos y metodológicos claves de los productos estadísticos que elabora la institución y, de esta forma, contribuir al intercambio de ideas entre los distintos componentes del Sistema Estadístico Nacional.

Las interpretaciones y opiniones que se expresan en los Documentos de Trabajo pertenecen en forma exclusiva a los autores y colaboradores y no reflejan necesariamente el punto de vista oficial del INE ni de la institución a la que pertenecen los colaboradores de los documentos. El uso de un lenguaje que no discrimine ni marque diferencias entre hombres y mujeres ha sido una preocupación en la elaboración de este documento. Sin embargo, y con el fin de evitar la sobrecarga gráfica que supondría utilizar en castellano “o/a” para marcar la existencia de ambos sexos, se ha optado por utilizar -en la mayor parte de los casos- el masculino genérico, en el entendido de que todas las menciones en tal género representan siempre a hombres y mujeres, abarcando claramente ambos sexos.

Reclasificación de la no respuesta: distinción entre la no respuesta al ítem y la no respuesta a la unidad

Resumen

Una de las estrategias más comunes para mitigar el efecto de la no respuesta a un ítem es utilizar métodos de imputación. Sin embargo, en ocasiones, dado el alto nivel de omisión parcial, resulta preferible reclasificar estos hogares, inicialmente codificados como entrevistados, en una de las categorías de rechazo. Altos niveles de no respuesta parcial pueden amenazar la calidad de las estimaciones y, en el caso de las encuestas de gasto, ser un probable indicio de una importante subdeclaración por parte de los informantes. El presente documento describe los criterios utilizados en la VIII Encuesta de Presupuestos Familiares (EPF) para distinguir entre un hogar entrevistado y aquel que debe ser reclasificado como *break-off*. La VII versión de la encuesta fue la primera en implementar un método de evaluación de la no respuesta parcial. Debido a los ajustes metodológicos entre esa y la VIII versión, este método debió ser actualizado. Los resultados de esta actualización indican que, aunque al establecer criterios más exigentes la tasa de respuesta disminuye, esto no genera diferencias significativas en las estimaciones de gasto de la encuesta.

Palabras claves: *No respuesta al ítem, break-off, calidad, Encuesta de Presupuestos Familiares (EPF).*

Abstract

Imputation is a common strategy to reduce the effect of item nonresponse. However, sometimes given the high levels of item nonresponse, it is preferable to reclassify initially coded interviewed households as part of the *break-off* category. High levels of partial omission can affect the quality of the estimations, and in the case of consumer expenditure surveys, they can be an indicator of an important degree of expense underreport. This document describe the criteria used at the Household Budget Survey 2016-17 to distinguish between a partial interview from a *break-off*. The previous version (2011-2012) was the first one where a method to evaluate the degree of item nonresponse was implemented. Due to methodological adjustments between these versions, these criteria needed to be revisited. Although more exigent criteria diminish to response rate, the change in this evaluation method does not significantly affect the consumer expenditure estimation.

Key words: *Item nonresponse, break-off, quality, Household Budget Survey (HBS).*

Índice general

1	Introducción	7
2	La Encuesta de Presupuestos Familiares	8
3	Falta de respuesta en las encuestas a hogares	9
3.1	Definición de la falta de respuesta	9
3.2	Falta de respuesta en la VIII EPF	9
4	Unidades entrevistadas y <i>break-off</i>	13
4.1	¿Cuándo la no respuesta al ítem pasa a comprometer la información de una unidad completa?	13
4.2	Criterios para la evaluación de la no respuesta al ítem en la VII EPF (2011-2012) . .	14
4.2.1	Método de aplicación	14
4.2.2	Ponderación por libretas	15
4.2.3	Definición de un valor umbral para distinguir entre entrevistados y <i>break-off</i>	20
5	Actualización de los criterios para la VIII EPF	22
5.1	Ponderación de los cuestionarios	22
5.2	Ponderación de los ítems de cada cuestionario	24
5.2.1	Registro de Personas en el Hogar (RPH)	24
5.2.2	Libreta de Gastos del Hogar (LGH)	25
5.2.3	Libreta de Gastos del Recuerdo (LGR)	25
5.2.4	Libreta de Ingresos (LI)	26
5.2.5	Libreta de Gastos en Servicios Financieros y Seguros (LSF)	26
5.3	Criterios de evaluación del registro diario de gastos	27
6	Plan de análisis	31
7	Evaluación de escenarios	35
7.1	Indicadores de respuesta	35
7.1.1	Análisis de dominancia	35
7.1.2	Tasa de respuesta	37
7.2	Efectos en los indicadores de gasto de la encuesta	39
7.3	Caracterización sociodemográfica de los hogares por nivel de colaboración	42
7.3.1	Características de los hogares reclasificados como <i>break-off</i>	45
7.4	Selección de un nivel de exigencia	47
8	Efecto de la actualización en los resultados de la VII EPF	50
8.1	Tasa de respuesta	50
8.2	Indicadores de gasto promedio	51
9	Conclusiones	54
10	Referencias	56

Índice de cuadros

1	Tasa de respuesta a nivel de hogares según área de estimación	11
2	Unidad de análisis por instrumento, VII EPF	15
3	Puntajes para la evaluación de la no respuesta al ítem, VII EPF	16
4	Puntajes asignados a ítems prioritarios para la evaluación de la no respuesta en cuestionario de Registro de Personas en el Hogar, VII EPF	16
5	Ejemplo de puntajes en libreta Registro de Personas en el Hogar	17
6	Puntajes para la evaluación de la no respuesta al ítem en Libreta de Gastos Individuales, VII EPF	18
7	Puntajes para la evaluación de la no respuesta al ítem en Libreta de Gastos del Hogar, VII EPF	18
8	Puntajes para la evaluación de la no respuesta al ítem en Libreta de Gastos del Recuerdo, VII EPF	19
9	Puntajes para la evaluación de la no respuesta al ítem en Libreta de Ingresos, VII EPF	19
10	Puntajes para la evaluación de la no respuesta al ítem en Libreta de Gastos en Servicios Financieros y Seguros, VII EPF	20
11	Calendario de referencia respecto a las actividades a realizar en cada visita de entrevista durante una quincena, VII EPF	21
12	Comparación de puntajes asignados a cada cuestionario en la VII y VIII EPF	23
13	Gasto promedio mensual registrado en la Libreta de Gastos Individuales, según condición de administración del gasto*	28
14	Puntajes de la Libreta de Gastos Individuales según los distintos criterios de ponderación (Administrador)	29
15	Puntajes de la Libreta de Gastos Individuales según los distintos criterios de ponderación (No administrador)	30
16	Comparación de criterios entre los distintos escenarios de evaluación considerados .	32
17	Divisiones de la Clasificación de Consumo Individual por Finalidades (CCIF)	33
18	Porcentaje de hogares reclasificados como <i>break-off</i> según puntajes de corte	35
19	Gasto promedio mensual de los hogares*, según escenario de evaluación	41
20	Caracterización de los hogares, según puntaje de corte 71 puntos	43
21	Caracterización de los hogares, según puntaje de corte 81 puntos	43
22	Caracterización de los hogares reclasificados como <i>break-off</i> , según escenario E . . .	46
23	Caracterización de los hogares, según puntaje de corte 81 puntos	49

Índice de figuras

1	Tasa de respuesta a nivel hogares, según región	12
2	Participación de los cuestionarios en la captura del gasto promedio mensual de los hogares	24
3	Participación de módulos de la Libreta Gastos del Hogar en la captura del gasto promedio mensual de los hogares*	25
4	Participación de módulos de la Libreta Gastos en Servicios Financieros y Seguros en la captura del gasto promedio mensual de los hogares*	26
5	Participación relativa de los administradores al gasto registrado en la Libreta de Gastos Individuales*	28
6	Porcentaje de hogares <i>break-off</i> para los valores de corte de 50 a 100 puntos	36
7	Tasa de respuesta para escenarios A y E, según puntajes de corte	38
8	Estructura del gasto promedio mensual de los hogares* para escenarios A y E, según división CCIF	39
9	Estructura del gasto promedio mensual de los hogares* para escenarios A (71 puntos), E (81 puntos) y aproximación a la estructura de las Cuentas Nacionales (2015), según división CCIF	40
10	Gasto promedio mensual de los hogares*, según escenarios A, E y Total de hogares	41
11	Tipos de hogares para puntaje de corte 71 puntos, según escenarios A, E y Total de hogares	44
12	Tipos de hogares para puntaje de corte 81 puntos, según escenarios A, E y total de hogares	45
13	Gasto promedio mensual de los hogares reclasificados como <i>break-off</i> para escenario E, según puntaje de corte*	47
14	Tasa de respuesta en hogares para escenarios A y E, según puntajes de corte de 71 y 76 puntos	48
15	Estructura del gasto promedio mensual de los hogares* para escenarios A (71 puntos) y E (76 puntos), según división CCIF	49
16	Tasa de respuesta a nivel hogares de la VII y VIII EPF para escenarios A y E, según puntajes de corte de 71 y 76 puntos	50
17	Estructura del gasto promedio mensual de la VII EPF para escenarios A (71 puntos) y E (76 puntos), según división CCIF	52
18	Gasto promedio mensual de los hogares de la VII y VIII EPF para escenarios A y E, según puntajes de corte de 71 y 76 puntos	53

1 Introducción

La no respuesta es un fenómeno común en las encuestas a hogares y, en general, las oficinas estadísticas disponen de distintas estrategias para atenuar su efecto en las estimaciones.

Teóricamente, la literatura distingue entre la no respuesta al ítem (o también denominada no respuesta parcial) y la no respuesta total (Bethlehem & Schouten, 2004; Cobben, 2009; Dillman, Eltinge, Groves, & Little, 2002). La primera se relaciona con la omisión o falta de información en algunos ítems del cuestionario, mientras que la segunda implica una pérdida total de información de la unidad seleccionada. A pesar de que la no respuesta ha sido uno de los tópicos predilectos de las investigaciones en metodología de encuestas, trasladar a la producción estadística esta sencilla distinción no es una tarea fácil, pues en la mayoría de los casos las encuestas requieren que determinada cantidad de preguntas sean respondidas antes de clasificar a un hogar como entrevistado (Dillman et al., 2002). En general, las unidades que participan de un estudio presentan distintos grados de no respuesta al ítem, siendo necesario establecer un criterio para distinguir cuándo conviene tratar un determinado nivel de no respuesta parcial como no respuesta total.

Altos niveles de no respuesta al ítem pueden ser indicadores de una importante subdeclaración por parte de los informantes. Aunque normalmente las encuestas de ingreso y gasto deben lidiar con este tipo de sesgo, es deseable contar con estrategias para disminuir el impacto que ello puede tener en la calidad de las estimaciones. Al mismo tiempo, indicadores que se construyen como la suma agregada de distintos ítems resultan más sensibles a importantes niveles de omisión. En estos casos será preferible reclasificar unidades inicialmente codificadas como entrevistadas, a categorías de no respuesta, pues no completan un número de preguntas mínimo para ser considerado como suficiente (lo que se denomina en la literatura anglosajona como *break-off*).

El siguiente documento pone a disposición del público los criterios utilizados en la VIII Encuesta de Presupuestos Familiares (EPF) para establecer esta distinción. La VII versión de la encuesta fue la primera en implementar un sistema de evaluación que permitiera definir en la práctica cierto nivel de no respuesta parcial permitido. No obstante, debido a los ajustes metodológicos entre esta y su VIII versión, estos criterios debieron ser actualizados, para así responder a las nuevas características de la encuesta.

El ejercicio de actualización buscó, por un lado, analizar si era compatible introducir criterios más exigentes para evaluar la no respuesta parcial, cumpliendo al mismo tiempo con la muestra objetivo del diseño. Por otro lado, se buscó identificar si el cambio de criterios tendría un efecto en las estimaciones de gasto de los hogares. Esto, pues si la reclasificación excluye unidades que poseen características muy distintas a los hogares finalmente clasificados como entrevistados, probablemente se podría introducir un sesgo en las estimaciones. El documento analiza, por tanto, el efecto de la actualización de los criterios de evaluación en la tasa de respuesta de la encuesta, sus principales indicadores de gasto, así como las características de los hogares que son recodificados como parte del *break-off*.

2 La Encuesta de Presupuestos Familiares

La Encuesta de Presupuestos Familiares (EPF) es una encuesta económica aplicada a hogares, cuyo propósito es recopilar información sobre gastos e ingresos en que aquellos incurren durante un período determinado, constituyéndose en un estudio fundamental para la actualización del Índice de Precios al Consumidor (IPC).

La primera EPF data de los años 1956-1957. Luego, la encuesta fue realizada con una periodicidad aproximada de 10 años hasta su VI versión (2006-2007). La VII EPF, cuya recolección se ejecutó entre noviembre de 2011 y octubre de 2012, fue la primera en efectuarse con una periodicidad de cinco años. Dicho cambio buscó hacer posible la actualización de la canasta del IPC con esta misma frecuencia, conforme a las sugerencias efectuadas por la Organización para la Cooperación y el Desarrollo Económicos (OCDE), en el marco del ingreso de Chile a esta organización en 2010. La VIII EPF, que desarrolló su trabajo de campo entre julio de 2016 y junio de 2017, fue la segunda en responder a este compromiso de aplicación con periodicidad quinquenal.

El objetivo principal de la VIII EPF fue identificar la estructura y características del gasto en consumo final de los hogares urbanos en las capitales regionales y en algunas de sus zonas conurbadas, con un período de referencia de un año. Como objetivo secundario, la encuesta buscó identificar la estructura del ingreso total disponible de los hogares urbanos, para la misma área geográfica y período de referencia.

3 Falta de respuesta en las encuestas a hogares

3.1 Definición de la falta de respuesta

La calidad de las estimaciones que se obtienen a partir de una muestra es uno de los aspectos más sensibles de la producción estadística. Esta puede ser entendida como una función del error total de una encuesta, el que, en términos generales, se define como la diferencia entre el valor de una estimación estadística de una muestra respecto de ese mismo valor en la población (Biemer & Lyberg, 2003).

Desde una perspectiva que entiende que la calidad de los datos depende del ciclo completo de una encuesta (Groves, Fowler, Lepkowski, Singer, & Tourangeau, 2009), es posible distinguir entre distintos tipos de error, dependiendo del momento del ciclo en que se producen. Hay, por ejemplo, errores de muestreo, de no respuesta, de procesamiento y de medición, entre otros. El objetivo de toda planificación y diseño de una encuesta está enfocado en minimizar las sucesivas brechas que se producen entre el valor estimado muestral y el poblacional, en cada una de las etapas de la producción estadística.

Entre estos errores, el de no respuesta representa la brecha entre la muestra seleccionada y las unidades efectivamente entrevistadas (Groves et al., 2009). A pesar de los múltiples esfuerzos que se realizan al momento de recolectar los datos, en muchos países se observan tendencias decrecientes en las tasas de respuesta (Beullens, Loosveldt, Vandenplas, & Stoop, 2018). Cuando la no respuesta implica la pérdida total de información de la unidad muestral, se denomina **no respuesta a la unidad**. Por otro lado, si la información faltante solo involucra determinados ítems del cuestionario, los que pueden faltar porque la pregunta no fue formulada o debido a que el informante omite información o ignora la respuesta, se denomina **no respuesta al ítem**.

La no respuesta generalmente ocasiona una reducción de la muestra con la que inicialmente fue diseñada la encuesta. Esto hace que las estimaciones sean menos precisas, pero no por eso menos válidas (Bethlehem & Schouten, 2004). Un efecto más problemático es que la no respuesta puede introducir un sesgo en las estimaciones. Si las unidades entrevistadas presentan diferencias respecto a las no entrevistadas en características relevantes para el estudio, las estimaciones que se realicen a partir de la muestra efectiva pueden subestimar o sobreestimar los valores con los que se busca hacer inferencia estadística (Groves et al., 2009).

3.2 Falta de respuesta en la VIII EPF

Como ya se mencionó, la **no respuesta a la unidad** se produce cuando no se consigue la información requerida para una unidad muestral y, por lo tanto, se pierde dicha información en su totalidad. En principio, esto puede producirse por falta de contacto con las unidades muestrales, por rechazos a participar del estudio o por otros motivos (incapacidad física y/o mental, dificultades de comunicación, etc.).

En la VIII EPF el principal motivo de no respuesta guardó relación con el rechazo de los hogares a colaborar con el estudio. Lo anterior se produjo bajo dos tipos de situaciones. La primera fue el

denominado **rechazo inicial explícito**, es decir, cuando desde el comienzo la unidad seleccionada¹ se niega a colaborar con la encuesta. La segunda situación se refiere al **rechazo implícito**² y se produce cuando el hogar acepta colaborar, pero la información aportada por este es insuficiente para cumplir con los objetivos del estudio o bien el hogar abandona la encuesta antes de término.

La falta de respuesta a la unidad tuvo un seguimiento constante durante el transcurso de la VIII EPF, mediante los **indicadores de desempeño operativo**. Estos eran un conjunto de estadísticos que aportaban información sobre el trabajo realizado por los equipos de campo y que permitían detectar las dificultades a las que se enfrentaban, dando oportunidad de mejorar las deficiencias, pero, a su vez, identificar y reforzar las fortalezas de los equipos³ (INE, 2018).

La base para elaborar estos indicadores se encontraba en el instrumento Hoja de Ruta, el cual permitía consignar los intentos de contacto con el hogar por parte de los encuestadores y caracterizar el resultado final de cada unidad muestral, a través del **Código de Disposición Final**, registrado en el módulo Registro de Contacto con el Hogar (RC)⁴.

Las categorías de los códigos de disposición final utilizados por la VIII EPF se encuentran en línea con lo propuesto por la *American Association for Public Opinion Research* (AAPOR, 2016), correspondiendo a una adaptación de la clasificación de incidencias sugerida por esta asociación. Las categorías usadas fueron las siguientes:

- Entrevistadas (E)
- Entrevistas interrumpidas o *break-off*⁵ (BO)
- Rechazos (R)
- Otros contactos sin respuesta (O)
- No contacto (NC)
- Elegibilidad desconocida (ED)
- No elegibles (NE)

Entre los indicadores de desempeño operativo construidos con estas categorías se encontraban la tasa de contacto, tasa de cooperación, tasa de rechazo, tasa de respuesta, tasa de pérdida de aperturas

¹En la VIII EPF el diseño muestral constó de dos etapas: primero se seleccionaron las unidades primarias de muestreo que correspondían a las manzanas y luego las unidades secundarias de muestro correspondientes a las viviendas particulares ocupadas. Dentro de las viviendas se identificó a los hogares que las habitaban (unidad de análisis), quienes proporcionaron la información necesaria para el estudio (INE, 2018).

²Hogares donde se realizó al menos una entrevista, pero que su información no fue suficiente para ser incorporados en el análisis final de la encuesta. En otras palabras, se trata de hogares donde se interrumpió el proceso de entrevistas tras la primera visita, ya sea porque no fue posible reestablecer el contacto o porque sus integrantes rehusaron continuar con la colaboración del estudio (INE, 2014a).

³Para más información de los indicadores de desempeño operativo, revisar el apartado 8.1.5.2 de la *Metodología VIII EPF*.

⁴La Hoja de Ruta contenía además los módulos: Identificación del Encuestador y Encargado de Grupo, Visita de Supervisión del Hogar, Registro de Aplicación de Libretas, Control de Libretas de Gastos Individuales, Agenda de Citas con el Informante y Observaciones. Para más información del formulario Hoja de Ruta, revisar el manual de trabajo de campo para encuestadores tomo 1, VIII EPF, disponible en: <http://www.ine.cl/estadisticas/ingresos-y-gastos/epf>.

⁵Si bien AAPOR emplea una categoría conjunta para rechazos y *break-off*, dadas las características de la no respuesta en EPF, se clasifica como una categoría separada. Ello se debe a la relevancia de las dificultades de mantención de la colaboración y obtención de información suficiente como motivo de no respuesta. Pese a lo anterior, para el cálculo de indicadores (tasas de respuesta, rechazo, cooperación y contacto), las unidades *break-off* reciben el mismo tratamiento que las unidades rechazadas inicialmente.

y tasa de logro objetivo. Los cuatro primeros indicadores fueron la base para evaluar el trabajo realizado por el equipo de campo y los dos últimos, indicadores complementarios de la encuesta.

Dentro de estos indicadores, la tasa de respuesta fue el más relevante para medir el desempeño del trabajo de campo, ya que reportaba el porcentaje total de unidades donde se obtuvo información suficiente, constatando la efectividad del equipo para recopilar los datos necesarios para el estudio. Este indicador se obtiene dividiendo el total de unidades entrevistadas con información suficiente por la suma de las unidades elegibles y de elegibilidad desconocida⁶ (INE, 2019a):

$$TR = \frac{E}{E + BO + R + O + NC + ED} * 100$$

La tasa de respuesta se puede calcular en relación con las viviendas, es decir, la cantidad de viviendas donde se obtuvo información suficiente en al menos uno de sus hogares, y también en relación con los hogares residentes en esas viviendas. Es importante recordar que la EPF es una encuesta aplicada a hogares, pero que debido a la dificultad de contar con un registro que permita localizar de forma estable a los hogares para el muestreo, las **unidades secundarias de muestreo** corresponden a las viviendas particulares ocupadas. Con posterioridad, una vez iniciadas las labores en terreno, se identificaron al interior de las viviendas los hogares que las habitaban (**unidad de análisis**).

El cuadro 1 muestra que para la VIII EPF la tasa de respuesta⁷ a nivel de hogares para el total de capitales regionales (Gran Santiago + Resto de capitales regionales del país) fue de 62,1%, resultando un 37,9% de no respuesta a nivel nacional. Si se observa por separado, el Gran Santiago muestra una tasa de respuesta de un 55,1%, mientras que el Resto de capitales regionales presenta mejores niveles en este indicador, con un 72,1%. Por contraste, la no respuesta queda en 44,9% y 27,9%, respectivamente.

Cuadro 1: Tasa de respuesta a nivel de hogares según área de estimación

Área de estimación	Tasa de respuesta (%)	Tasa de no respuesta (%)
Gran Santiago	55,1	44,9
Resto capitales regionales	72,1	27,9
Total capitales regionales	62,1	37,9

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

A nivel regional (figura 1) Ñuble, Arica y Parinacota, Antofagasta y Metropolitana son las que presentan menores niveles de respuesta, siendo esta última la que tiene la tasa más baja, con un 55,1%, quedando por debajo de la tasa de respuesta total de capitales regionales (62,1% - representada por la línea punteada de la figura). En contraste, las regiones con mejor rendimiento corresponden a Tarapacá, Magallanes y Los Ríos, con tasas de respuesta de 90,6%, 87% y 83%, respectivamente.

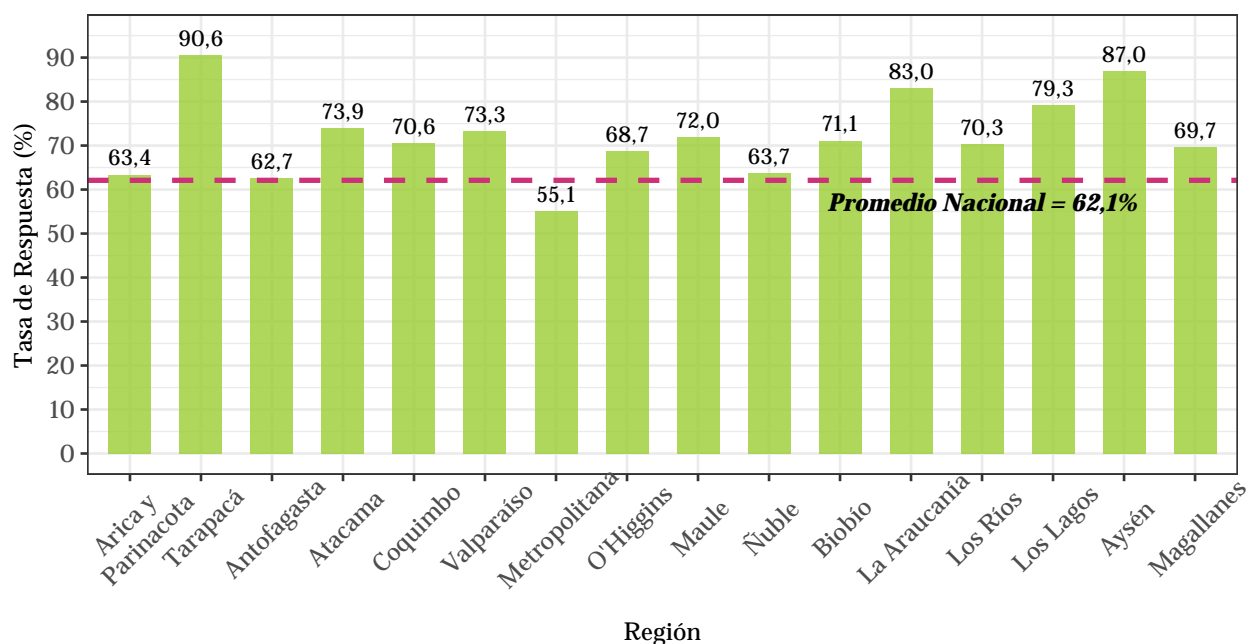
⁶Para más información sobre la tasa de respuesta y los otros indicadores de desempeño operativo, revisar el documento *Informe de Calidad VIII EPF*, disponible en: <http://www.ine.cl/estadisticas/ingresos-y-gastos/epf>.

⁷El cálculo de las tasas de respuesta a nivel hogar asume el supuesto de que en cada unidad sin respuesta o de elegibilidad desconocida existe un hogar. De este modo, la tasa de respuesta es calculada sobre un total de hogares estimados (elegibles o de elegibilidad desconocida).

Como se mencionó antes, la **no respuesta al ítem** se produce cuando existe respuesta por parte de la unidad seleccionada, pero falta información para algunas de sus preguntas.

En la VIII EPF se utilizaron seis instrumentos de recolección de datos o **libretas**; cada uno de ellos incluyó distintos apartados que agrupaban preguntas vinculadas por temáticas. Lo anterior permitió identificar la no respuesta en diferentes niveles: a nivel de pregunta, apartado o libreta. Por lo tanto, la falta de información en alguno de estos niveles se consideró como no respuesta al ítem, pudiendo existir desde la falta de respuesta a una pregunta hasta la falta de respuesta a alguna libreta completa.

Figura 1: Tasa de respuesta a nivel hogares, según región



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

4 Unidades entrevistadas y *break-off*

4.1 ¿Cuándo la no respuesta al ítem pasa a comprometer la información de una unidad completa?

Una de las estrategias más comunes para mitigar el efecto de la no respuesta al ítem corresponde al uso de métodos de imputación. Estos tienen el objetivo de substituir información faltante en los datos de una encuesta, para luego incluirla en las estimaciones que se realizan a partir de ella como valores observados. Entre los métodos de imputación más usados se encuentra el método *hot deck*, que fue utilizado para la imputación de ingresos y gastos en la VIII EPF⁸.

Sin embargo, existen ocasiones en que la no respuesta parcial no puede seguir siendo considerada como un problema que afecta únicamente a algunas preguntas de la encuesta. El nivel de no respuesta puede comprometer la calidad de las estimaciones, sobre todo cuando se trata de indicadores que se construyen a partir de la suma de varios ítems de un cuestionario. Así, por ejemplo, cuando varias preguntas de gasto no son contestadas, el nivel de subdeclaración de este indicador en un hogar puede ser bastante alto, comprometiendo la calidad de las estimaciones. En estos casos puede considerarse preferible reclasificar unidades inicialmente codificadas como entrevistadas, asignándolas a la categoría de unidades que rehusaron colaborar con la encuesta (*break-off*), dada la insuficiencia de la información reportada.

La Asociación Americana para la Investigación en Opinión Pública señala que es recomendable definir *a priori* criterios explícitos respecto a lo que distingue una entrevista completa de una parcial, y aquello que diferencia una entrevista parcial de un *break-off* (AAPOR, 2016). La AAPOR cita 3 criterios posibles para establecer esta distinción: 1) la proporción de todas las preguntas aplicables respondidas, 2) la proporción de preguntas cruciales o esenciales respondidas y 3) la proporción de todas las preguntas aplicables administradas. Estos criterios fueron propuestos por Frankel en 1983 (AAPOR, 2016). A pesar de la existencia de estas recomendaciones, desde la literatura también se señala que no existe una “regla de oro” para definir un límite exacto entre los problemas de la no respuesta al ítem y la no respuesta a la unidad (Dillman et al., 2002). La oficina de estadísticas laborales de Estados Unidos, por ejemplo, plantea que para el diario de gastos empleados en la Encuesta de Presupuestos Familiares de ese país (*Consumer Expenditure Diary Survey*), cuando existe una alta sospecha de una subdeclaración de gastos, unidades que inicialmente fueron clasificadas como entrevistadas son recodificadas como rechazo. *Bureau of Labor Statistics* destaca que esta proporción de casos es pequeña y que los criterios para identificar aquellos casos a reclasificar como *break-off* debido a esta subdeclaración son varios y dependen tanto del monto total que se declara en el diario como del monto declarado en el ítem de alimentación, entre otros criterios (Hubener, Knappenberger, Lee, & Tan, 2018).

Considerando que el objetivo principal de la EPF es identificar la estructura de gasto de los hogares, si un hogar muestra indicios de una importante subdeclaración de gastos, resulta razonable fijar criterios que permitan sopesar cuando la no respuesta al ítem amenaza con introducir un sesgo

⁸Para más información sobre la imputación, revisar el documento *Métodos de imputación VIII EPF: Gastos diarios e ingresos de la actividad laboral y jubilaciones*, disponible en: <https://www.ine.cl/inicio/documentos-de-trabajo/metodologicos>.

debido a la subdeclaración. Desde su VII versión, la EPF implementa un método para la evaluación del nivel de no respuesta al ítem asociado a cada hogar que participa en la encuesta. A través de una serie de criterios que asignan un puntaje en relación con la completitud de información de cada hogar, se establece un umbral para definir si la falta de respuesta debe ser abordada como un problema de no respuesta al ítem (y mitigado a través de los procesos de imputación) o no respuesta a la unidad (cuyo manejo se realiza a través de los ajustes a los factores de expansión). Este método de evaluación es denominado “grilla técnica de mínimos de calidad” (en adelante, grilla técnica).

La grilla técnica recodifica hogares que inicialmente se clasificaron como entrevistados, pero que por su nivel de no respuesta al ítem pasan a ser parte de la no respuesta total. De esta manera, un umbral exigente respecto a la no respuesta parcial “permitida” conlleva un mayor número de hogares reclasificados como rechazo, afectando la tasa de respuesta⁹ de la encuesta. A medida que se establece un criterio de corte más estricto, menor será la tasa de respuesta de la encuesta y el tamaño del grupo de unidades a partir del cual se realizan inferencias de la población. Por el contrario, si se establece un criterio más laxo, la reclasificación de casos será menor. Existe, por tanto, un *trade-off* entre el nivel de exigencia que se requiere del nivel de respuesta en cada hogar y el nivel de logro necesario para cumplir con el tamaño muestral definido como objetivo en el diseño muestral.

4.2 Criterios para la evaluación de la no respuesta al ítem en la VII EPF (2011-2012)

La séptima versión de la EPF incorporó criterios para reclasificar aquellos hogares que, por su nivel de no respuesta al ítem, en la práctica convenía considerarlos como *break-off*. A continuación se detalla la forma en que estos criterios fueron aplicados en relación con el método de evaluación, la ponderación de los cuestionarios y la definición de un puntaje de corte para distinguir entre los hogares entrevistados y los *break-off*.

4.2.1 Método de aplicación

El método de aplicación de los criterios en la VII EPF consistió en asignar un puntaje a ciertos módulos y preguntas de cada una de las libretas, en donde la suma total de los puntajes por hogar varió entre 1 y 100 puntos, siendo 1 el puntaje mínimo que un hogar podía obtener y 100, el máximo. En esta escala de puntajes, el valor 71 marcaba la diferencia entre una unidad entrevistada con cierto nivel de omisión parcial y una unidad clasificada como rechazo implícito (*break-off*). La selección de este puntaje como valor de corte tuvo relación con la ponderación de los registros diarios de gastos en la evaluación de la no respuesta parcial. Este aspecto será desarrollado con más detalle en un apartado posterior.

⁹Si bien la modificación de los criterios afecta a todos los indicadores de desempeño operativo, para fines del documento se aborda el impacto en el más relevante de estos, correspondiente a la tasa de respuesta.

Si bien en este sistema de evaluación cada hogar¹⁰ obtuvo un puntaje propio, considerando que los cuestionarios de la encuesta recopilan información a distintos niveles (hogar, persona, productos¹¹, etc.), aquellos ítems relacionados con una unidad distinta a la del hogar fueron ponderados por el número de personas o productos asociados a esa unidad. Así, por ejemplo, si en un hogar existió una omisión al momento de registrar el parentesco de alguno de sus integrantes, la penalización por esta falta de información fue ponderada por la cantidad de personas que debía responder esta información¹². En el cuadro 2 se presentan los distintos niveles respecto a los cuales cada instrumento de la encuesta recopilaba información.

Cuadro 2: Unidad de análisis por instrumento, VII EPF

Instrumento	Unidad de análisis
Registro de Personas en el Hogar	Vivienda
Registro de Personas en el Hogar	Hogar
Registro de Personas en el Hogar	Persona
Libreta de Gastos Individuales	Persona
Libreta de Gastos del Hogar	Hogar
Libreta de Gastos del Hogar	Persona
Libreta de gastos del Recuerdo	Hogar
Libreta de Ingresos	Persona
Libreta de Gastos en Servicios Financieros y Seguros	Persona
Libreta de Gastos en Servicios Financieros y Seguros	Producto financiero

Fuente: Instituto Nacional de Estadísticas (INE) - VII Encuesta de Presupuestos Familiares (EPF).

4.2.2 Ponderación por libretas

En la VII EPF, el valor de la penalización en el puntaje por la falta de respuesta en determinado ítem dependió de la ponderación relativa de cada libreta en la construcción del puntaje a nivel hogar. La ponderación de cada libreta fue asignada en función de la importancia relativa de la información que cada cuestionario aportó al cumplimiento de los objetivos de la encuesta. La libreta a la que se le asignó el mayor puntaje fue la Libreta de Gastos Individuales (LGI), pues era el cuestionario que contribuía proporcionalmente con la mayor cantidad de información sobre los gastos del hogar (cuadro 3).

¹⁰Un hogar está constituido por una o más personas que, unidas o no por vínculos de parentesco, residen en una misma vivienda o parte de ella y se benefician de un mismo presupuesto, principalmente para alimentación.

¹¹El uso de productos como unidad de análisis tiene lugar en el cuestionario de Gastos en Servicios Financieros y Seguros, el cual recopila información respecto a la tenencia de instrumentos financieros y recoge la información a nivel de cada instrumento utilizado por algún integrante del hogar.

¹²Ej.: Puntaje asociado a pregunta z/n , donde n corresponde a la cantidad de integrantes de un hogar.

Cuadro 3: Puntajes para la evaluación de la no respuesta al ítem, VII EPF

Libreta	Puntaje total
Registro de Personas en el Hogar	26
Libreta de Gastos Individuales	30
Libreta de Gastos del Hogar	15
Libreta de Gastos del Recuerdo	9
Libreta de Ingresos	10
Libreta de Gastos en Servicios Financieros y Seguros	10
Total	100

Fuente: Instituto Nacional de Estadísticas (INE) - VII Encuesta de Presupuestos Familiares (EPF).

4.2.2.1 Registro de Personas en el Hogar (RPH)

El objetivo del RPH es caracterizar sociodemográficamente a los integrantes de un hogar, para así establecer enlaces con el resto de los instrumentos de la encuesta. Estas características permiten no solo analizar los patrones de gasto de distintos tipos de hogares, sino que también su información es fundamental para los ejercicios de imputación y de cálculo de factores de expansión.

Si bien el RPH no captura gastos, considerando la importancia para el estudio de los datos consultados en este, en la grilla técnica de la VII EPF se le asignó el segundo mayor puntaje (26 puntos). Los ítems por los que se penalizó por falta de respuesta se presentan en el cuadro 4 y corresponden a aquellos utilizados tanto para la imputación como para el cálculo de factores.

Cuadro 4: Puntajes asignados a ítems prioritarios para la evaluación de la no respuesta en cuestionario de Registro de Personas en el Hogar, VII EPF

Ítems a considerar por módulo o pregunta	Puntaje	Ponderador
Parentesco con el jefe de hogar	5	Nº integrantes
Edad	5	Nº integrantes
Sexo	5	Nº integrantes
Estado civil	2	Nº integrantes
Nivel de educación e identificación de los estudiantes	5	Nº integrantes
Condición de actividad económica	4	Nº integrantes

Fuente: Instituto Nacional de Estadísticas (INE) - VII Encuesta de Presupuestos Familiares (EPF).

El puntaje asignado a cada pregunta se dividió por el número de integrantes que conformaban cada hogar, por lo que se podía obtener mediante la siguiente fórmula:

$$Puntaje\ RPH = \sum_{i=1}^6 \frac{P_i * x_i}{n}$$

Donde:

P_i = Puntaje total del ítem i

x_i = Número de integrantes que responden el ítem i de forma válida¹³

n = Número de integrantes del hogar

A modo de ejemplo, en el cuadro 5 se muestra cuál habría sido el puntaje de RPH en la VII EPF para un hogar de cuatro integrantes en que no se hubiese dispuesto información de la edad de uno de ellos y tampoco del nivel educacional de dos de sus integrantes. De un máximo de 26 puntos, en este caso la penalización por falta de respuesta habría hecho que este hogar obtuviese 23 puntos.

Cuadro 5: Ejemplo de puntajes en libreta Registro de Personas en el Hogar

Ítem	Puntaje teórico	Puntaje ejemplo
Parentesco con el jefe de hogar	5	5,0
Edad	5	5,0
Sexo	5	5,0
Estado civil	2	1,5
Nivel de educación	5	2,5
Condición de actividad económica	4	4,0
Puntaje total RPH	26	23,0

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

4.2.2.2 Libreta de Gastos Individuales (LGI)

El objetivo de este instrumento es servir de registro para todos los gastos en productos y servicios que realiza, con frecuencia diaria, cada integrante del hogar de 15 años o más. Este es el único instrumento autoadministrado de la encuesta, siendo utilizado por los informantes como un cuadernillo en el que día a día anotaron los gastos efectuados durante la quincena de colaboración con la EPF¹⁴.

La LGI fue la libreta de mayor importancia para los criterios de la grilla técnica de la VII EPF (30 puntos). A través de un análisis de sensibilidad de los datos obtenidos en la marcha blanca y del primer trimestre de trabajo de campo oficial de esa versión del estudio, se estableció como criterio que, para considerar como respondida esta libreta, debía contar con un mínimo de cuatro días de registro asociados al mejor informante¹⁵ del hogar, y que al menos un día debía corresponder al registro de los gastos de un día de fin de semana¹⁶ (cuadro 6).

¹³Se considera una respuesta como válida cuando el informante responde efectivamente la pregunta, o bien cuando por la secuencia en el flujo de preguntas, no corresponde responder esa pregunta y, por lo tanto, no aplica para ese informante o ese hogar.

¹⁴Cabe destacar que en el registro de gastos, si se señala que ese día no se realizó gasto alguno o en algún servicio el gasto de ese hogar es cero, esa respuesta también es considerada como válida.

¹⁵El mejor informante corresponde a la persona del hogar que tiene a lo menos cuatro días de registro y, entre ellos, al menos uno corresponde a un día de fin de semana.

¹⁶Los días con registros corresponden a días con gasto o con declaración de gasto cero (distinto al día sin registro que puede deberse a la omisión, olvido o rechazo a registrar).

Cuadro 6: Puntajes para la evaluación de la no respuesta al ítem en Libreta de Gastos Individuales, VII EPF

Ítems a considerar por módulo o pregunta	Puntaje	Ponderador
Libreta del mejor informante del hogar con un mínimo de 4 días de registro	25	Hogar
Al menos 1 día de registro en fin de semana	5	Hogar

Fuente: Instituto Nacional de Estadísticas (INE) - VII Encuesta de Presupuestos Familiares (EPF).

De esta forma, siguiendo con el ejemplo anterior, para el hogar de cuatro integrantes era suficiente con que uno de ellos hubiese cumplido con cuatro días de registro, correspondiendo uno de ellos a un día de fin de semana, para que el hogar hubiese obtenido los 30 puntos de esta libreta. Si ninguno hubiera cumplido el requerimiento, la no respuesta al ítem habría sido penalizada inmediatamente en 30 puntos, por lo que el hogar habría sido clasificado como *break-off*.

4.2.2.3 Libreta Gastos del Hogar (LGH)

El objetivo principal de la LGH es capturar todos los gastos que realiza un hogar cuyos pagos presentan algún tipo de periodicidad (mensual, semestral, anual, etc.). El ejemplo más claro son los gastos en servicios básicos de una vivienda, tales como agua, electricidad, entre otros.

La selección de los ítems para esta libreta se hizo utilizando la información de la marcha blanca y primer trimestre de trabajo de campo de la VII EPF, en función de su contribución al gasto total del hogar. Dado que en esta libreta se registra información a nivel hogar y a nivel persona, los ítems seleccionados para evaluar la presencia de no respuesta parcial se ponderaron de manera distinta, dependiendo del nivel al que se encontraba cada pregunta. El cuadro 7 resume los puntajes para cada uno de los módulos seleccionados para la evaluación de la grilla técnica.

Cuadro 7: Puntajes para la evaluación de la no respuesta al ítem en Libreta de Gastos del Hogar, VII EPF

Ítems a considerar por módulo o pregunta	Puntaje	Ponderador
Gastos básicos del hogar (agua y electricidad)	4	Hogar
Módulo de telecomunicaciones	7	Hogar
Módulo de educación	4	Nº estudiantes

Fuente: Instituto Nacional de Estadísticas (INE) - VII Encuesta de Presupuestos Familiares (EPF).

4.2.2.4 Libreta Gastos del Recuerdo (LGR)

El objetivo de esta libreta es servir de registro de los gastos que son poco frecuentes o no tienen, necesariamente, una periodicidad establecida para el hogar, pero pueden representar una suma importante de gasto en relación con su presupuesto. En la VII EPF, las preguntas de la LGR se encontraban organizadas en 21 tablas que consultaban por gastos en reparaciones menores del hogar, consultas médicas, artículos electrónicos del hogar, viajes realizados y compra de

automóviles, entre otros. Al comienzo de cada tabla existía una pregunta filtro que permitía distinguir si el hogar había o no realizado gastos en los ítems contenidos en ella.

En la VII EPF el encuestador debía registrar una “X” en la pregunta filtro en caso de que no se registraran gastos en alguna tabla, y dejarla en blanco en caso contrario. Esta diagramación inducía a error en el registro de la información, ya que cambiaba la lógica habitual de llenado de filtros frente al resto de los cuestionarios (donde los números 1 y 2 representaron el *sí* y el *no*, respectivamente). Por lo mismo, para esta libreta se evaluó que al menos 10 de las 21 tablas estuvieran contestadas de manera consistente (con o sin gastos). Cuando esta condición se cumplía, se asignaba la totalidad del puntaje (cuadro 8).

Cuadro 8: Puntajes para la evaluación de la no respuesta al ítem en Libreta de Gastos del Recuerdo, VII EPF

Ítems a considerar por módulo o pregunta	Puntaje	Ponderador
Libreta contestada en al menos 10 tablas	9	Hogar

Fuente: Instituto Nacional de Estadísticas (INE) - VII Encuesta de Presupuestos Familiares (EPF).

4.2.2.5 Libreta de Ingresos (LI)

La Libreta de Ingresos responde al objetivo secundario de la EPF, a saber, identificar la estructura del ingreso total disponible de los hogares. Los ingresos provenientes del trabajo, sea este dependiente o independiente, son los que más contribuyen a esta estructura. Por esta razón, los módulos donde se captura esta información fueron considerados prioritarios al momento de evaluar el nivel de no respuesta al ítem (cuadro 9). Al igual que en el RPH y en el módulo de educación de LGH, el puntaje de estos módulos fue ponderado por el número de personas que debía contestar estos módulos.

Cuadro 9: Puntajes para la evaluación de la no respuesta al ítem en Libreta de Ingresos, VII EPF

Ítems a considerar por módulo o pregunta	Puntaje	Ponderador
Asalariados (remuneración neta/monto bruto honorarios)	5	Nº ocupados dependientes
Cuenta Propia (monto líquido o promedio mensual)	5	Nº ocupados independientes

Fuente: Instituto Nacional de Estadísticas (INE) - VII Encuesta de Presupuestos Familiares (EPF).

4.2.2.6 Libreta de Gastos en Servicios Financieros y Seguros (LSF)

El objetivo principal de la LSF es capturar y medir de forma detallada el gasto asociado al uso o mantención de productos de servicios financieros, tales como intereses, comisiones y seguros¹⁷, entre

¹⁷La LSF incluye preguntas por primas de seguro asociadas a los productos financieros de los módulos: Créditos Hipotecarios, Tarjetas de Crédito de Establecimientos Comerciales, Créditos de Consumo y Automotriz, Cuentas Corrientes y Líneas de Crédito, y Tarjetas de Crédito Bancarias.

otros, así como también el gasto asociado a los pagos de las primas brutas de los seguros contratados por los integrantes del hogar que sean titulares de estos productos¹⁸.

Este instrumento sufrió importantes modificaciones en la VII versión de la encuesta en comparación con la VI EPF. Mientras que en la sexta versión se contaba con un número acotado de variables para caracterizar el gasto financiero de las personas, imputándose finalmente estos en su totalidad, en la VII EPF el cuestionario incluyó submódulos específicos en que los informantes debían responder utilizando, como información complementaria, los comprobantes de pago otorgados por las instituciones financieras, buscando así medir de manera más certera y confiable el gasto asociado a estos servicios.

La elección de los módulos para la VII EPF se realizó sobre la base de la información de la marcha blanca y primer trimestre de esta versión de la encuesta (cuadro 10). Si bien la información se encuentra a nivel de titulares de productos, la ponderación para cada módulo se efectuó a nivel hogar: si al menos uno de los instrumentos declarados por el hogar en el módulo se encontraba correctamente registrado, el hogar obtenía los puntos establecidos para ese módulo.

Cuadro 10: Puntajes para la evaluación de la no respuesta al ítem en Libreta de Gastos en Servicios Financieros y Seguros, VII EPF

Ítems a considerar por módulo o pregunta	Puntaje	Ponderador
Crédito hipotecario	2	Hogar
Crédito de consumo	2	Hogar
Créditos automotrices	1	Hogar
Créditos de educación	1	Hogar
Tarjetas no bancarias	3	Hogar
Seguros	1	Hogar

Fuente: Instituto Nacional de Estadísticas (INE) - VII Encuesta de Presupuestos Familiares (EPF).

4.2.3 Definición de un valor umbral para distinguir entre entrevistados y *break-off*

La LGI fue la libreta a la cual se le asignó un mayor puntaje de evaluación, contando con un total de 30 puntos. Debido a lo anterior, el umbral de corte se estableció en relación con el puntaje de LGI, fijándose en 71 puntos. De este modo, si un hogar no contaba con al menos una LGI respondida, el estado final de este se consignaba como rechazo (*break-off*), aun cuando el hogar hubiese contestado otra información consultada por la encuesta (INE, 2013).

Si bien la ponderación establecida para las libretas reconoce la importancia de la LGI para la estimación de gastos de un hogar, la falta de respuesta en otras libretas también podía gatillar que un hogar traspasara un umbral permitido de no respuesta al ítem, provocando su reclasificación. Por ejemplo, si un hogar cumplía con los requerimientos para la LGI, pero no reportaba información en la Libreta de Gastos del Hogar, Ingresos y Servicios Financieros y Seguros, el estado final de ese hogar era recodificado como *break-off*.

¹⁸Los seguros contratados de forma independiente (no asociados al uso o mantención de productos financieros) son consultados por separado en el módulo denominado Seguros, correspondientes a seguros de vivienda, seguros de automóvil, seguros privados de salud, entre otros.

La elaboración del sistema de penalización para evaluar la no respuesta parcial también consideró el orden en el cual se aplicaban los cuestionarios (cuadro 11). Teniendo en cuenta el puntaje umbral de corte de 71 puntos, si un hogar desistía de participar de la encuesta durante la segunda visita y devolvía las LGI con registro hasta ese día, podía alcanzar a completar los criterios mínimos para superar las exigencias de la grilla técnica ($71 \text{ puntos} = 26(RPH) + 15(LGH) + 30(LGI)$). En ese sentido, la asignación de los puntajes a nivel de libretas también se vinculaba a cómo se organizaron las visitas de entrevista y el orden de aplicación de los instrumentos.

Cuadro 11: Calendario de referencia respecto a las actividades a realizar en cada visita de entrevista durante una quincena, VII EPF

Día 1 al 5	Día 4 al 7	Día 8 al 12	Día 11 al 15
Primera visita	Segunda visita	Tercera visita	Cuarta visita
Entrevista RPH	Recuperación de información RPH	Recuperación de información LGH	Recuperación de información LI
Entrega LGI	Entrevista LGH	Segunda entrevista LGR	Seguimiento LGI
	Entrevista LGR	Entrevista LI	Seguimiento y retiro LGI
	Seguimiento LGI	Entrevista LSF	

Fuente: Instituto Nacional de Estadísticas (INE) - VII Encuesta de Presupuestos Familiares (EPF).

5 Actualización de los criterios para la VIII EPF

La evaluación del nivel de no respuesta parcial permite reducir posibles sesgos en las estimaciones, asociadas a la subdeclaración de los informantes. Debido a los ajustes metodológicos¹⁹ entre la VII y la VIII versión de la encuesta, se hizo necesario revisar la vigencia de los criterios utilizados para reclasificar la no respuesta, en función de las nuevas características del estudio. Al mismo tiempo, el objetivo de actualizar los criterios permitió analizar si resultaba posible aplicar una penalización más exigente al nivel de no respuesta parcial de las unidades.

La literatura recomienda establecer este tipo de criterios antes del término del trabajo de campo de una encuesta (AAPOR, 2016). En el caso de la VIII EPF, estos criterios fueron evaluados durante 2017, utilizando para esto información preliminar del primer semestre de recolección de la encuesta²⁰. Esta planificación permitió contar con un método definido para evaluar la no respuesta al ítem de los hogares seis meses antes de la publicación de la encuesta en junio de 2018. Nótese, sin embargo, que los resultados de los ejercicios que se presentan en la sección 7 y 8 de este documento se calcularon teniendo como base la información recolectada durante los 12 meses de trabajo de campo. El objetivo de esto es presentar a los usuarios los efectos que pudo haber tenido la actualización de estos criterios a los indicadores oficiales de la encuesta (que consideran un año como período de referencia). Cabe destacar que el equipo técnico de la EPF definió los criterios tomando en cuenta solo el primer semestre, pues era necesario contar con estos antes de la publicación.

Para la actualización de los criterios se definieron *a priori* 3 dimensiones a evaluar de la grilla técnica utilizada en la VII EPF. Estas dimensiones fueron las más relevantes en la construcción del puntaje para la evaluación de la no respuesta parcial, y se listan a continuación:

- Ponderación de las libretas en el cálculo del puntaje.
- Ponderación de los ítems evaluados al interior de cada libreta.
- Criterio de evaluación de la Libreta de Gastos Individuales (LGI).

La evaluación de los cambios en estas dimensiones se realizó a partir de la conformación de escenarios alternativos, los que se desarrollan en detalle en la sección 6 del presente documento. En lo que sigue de este apartado se revisará en detalle en qué consistieron las modificaciones propuestas.

5.1 Ponderación de los cuestionarios

Tal como se mencionó con anterioridad, en la VII versión de la encuesta cada cuestionario tenía un puntaje, asignado en función de su importancia relativa en el registro de gastos en el hogar.

¹⁹Entre la VII y la VIII versión de la encuesta se realizaron modificaciones a los cuestionarios, que incluyeron cambios en el fraseo y flujo de algunas preguntas y categorías de respuesta, eliminación /inclusión de preguntas, entre otros. Además, para mejorar los tiempos de entrevista por cuestionario, se modificó el orden de aplicación del flujo de entrevistas. Estas modificaciones buscaban mejorar la declaración de gastos e ingresos y reducir el tiempo de entrevista por cada cuestionario, para obtener una mayor y mejor captura de información y, a su vez, disminuir el rechazo de los hogares.

²⁰Desde aquí en adelante, cada vez que se mencione el primer semestre se hará referencia a los datos calculados de forma preliminar durante 2017.

En el cuadro 12 se observa la reasignación de los puntajes por cuestionario entre la VII y la VIII versión de la encuesta. La nueva ponderación fue establecida de acuerdo con la contribución de los cuestionarios en la captura del gasto promedio²¹, calculado para el primer semestre de trabajo de campo de la VIII versión de la encuesta, estableciendo un puntaje mínimo para los cuestionarios que no contribuyen al gasto (Registro de Personas en el Hogar) o lo hacen en menor medida (Libreta de Ingresos y Libreta de Gastos en Servicios Financieros y Seguros).

Cuadro 12: Comparación de puntajes asignados a cada cuestionario en la VII y VIII EPF

Cuestionario	Puntaje VII EPF	Puntaje VIII EPF
Registro de Personas en el Hogar	26	15
Libreta de Gastos Individuales	30	29
Libreta de Gastos del Hogar	15	25
Libreta de gastos del Recuerdo	9	16
Libreta de Ingresos	10	10
Libreta de Gastos en Servicios Financieros y Seguros	10	5

Fuente: Instituto Nacional de Estadísticas (INE) - VII y VIII Encuesta de Presupuestos Familiares (EPF).

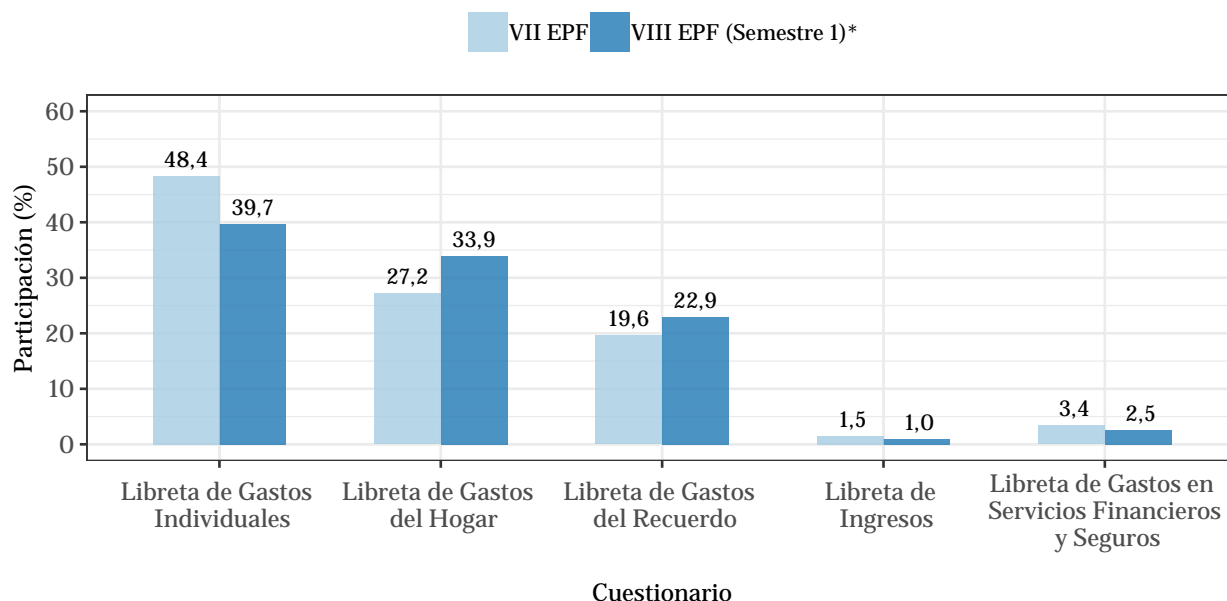
En el caso del RPH se estableció una ponderación de 15% en el total del puntaje, mientras que para la LI, un 10% y para la LSF, un 5%. Estos puntajes representan una evaluación normativa realizada por el equipo técnico respecto a la participación de estas libretas para el cumplimiento de los objetivos de la EPF. Para el resto de las libretas se repartió el peso restante (70%), de acuerdo con su contribución al gasto en los resultados del primer semestre de la VIII EPF (figura 2).

La libreta a la que se le asignó el mayor puntaje fue la Libreta de Gastos Individuales (LGI), debido a que corresponde al cuestionario que aporta la mayor cantidad de información sobre los gastos del hogar (39,7% sobre el gasto total del primer semestre de la VIII EPF).

Con esta información, la ponderación de la LGI se define en 29% respecto del total del puntaje de la grilla técnica, mientras que LGH y LGR actualizan su ponderación a un 25% y 16%, respectivamente.

²¹Correspondiente al gasto promedio mensual de los hogares, excluyendo el arriendo imputado.

Figura 2: Participación de los cuestionarios en la captura del gasto promedio mensual de los hogares



Fuente: Instituto Nacional de Estadísticas (INE) – VII y VIII Encuesta de Presupuestos Familiares (EPF).

*Las estimaciones de la VIII EPF fueron calculadas con datos preliminares del primer semestre de la encuesta.

La modificación de la ponderación por libreta se realizó en función del aumento de la participación del gasto de las libretas de gastos del hogar y del recuerdo, las que incrementaron su puntaje en 10 y 7 puntos, respectivamente, en relación con el puntaje de la VII EPF.

5.2 Ponderación de los ítems de cada cuestionario

La nueva ponderación de las preguntas seleccionadas en cada libreta se construyó a partir de la importancia de los módulos que contenían estos ítems. Esta importancia fue medida como la contribución del módulo al gasto promedio total de los hogares, a excepción de los puntajes asignados a las preguntas del cuestionario Registro de Personas en el Hogar, los cuales se asignaron de acuerdo con su relevancia para caracterizar el gasto de los hogares y su aporte para el proceso de imputación. A continuación se revisarán los cambios asociados a cada ítem por libreta de la encuesta.

5.2.1 Registro de Personas en el Hogar (RPH)

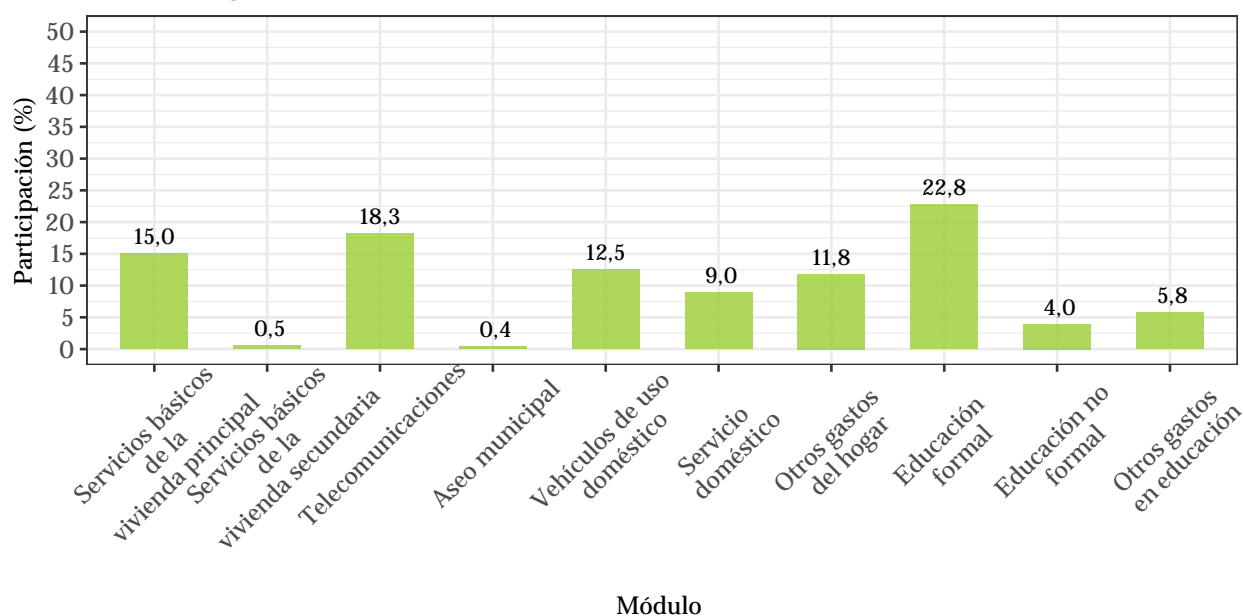
La única modificación respecto a la séptima versión corresponde a la penalización asociada a la respuesta en la pregunta por el estado civil. Se decide excluir esta pregunta del método de evaluación, redistribuyendo su puntaje a los ítems de condición de actividad económica y edad. El estado civil de los integrantes de un hogar fue considerado inicialmente en la VII EPF como insumo para el proceso de imputación de LGI, la cual posteriormente no se incluyó, debido a su baja correlación con el gasto

individual (INE, 2014b). La asignación de puntaje para el resto de las preguntas se mantuvo respecto a la VII EPF.

5.2.2 Libreta de Gastos del Hogar (LGH)

Para este cuestionario se mantuvieron los módulos evaluados en la grilla técnica de la VII EPF. Esto, dado que, de acuerdo con los datos del primer semestre de su octava versión, los gastos asociados a los servicios básicos del hogar, telecomunicaciones y educación formal siguieron siendo los módulos más importantes en esta libreta en términos de su contribución al gasto (figura 3).

Figura 3: Participación de módulos de la Libreta Gastos del Hogar en la captura del gasto promedio mensual de los hogares*



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

*Las estimaciones de la VIII EPF fueron calculadas con datos preliminares del primer semestre de la encuesta.

5.2.3 Libreta de Gastos del Recuerdo (LGR)

Para esta libreta se adaptó la evaluación, otorgando puntaje a todas las tablas de la libreta, relevando la importancia de cada ítem contenido en LGR y disminuyendo de esta forma la posibilidad de subdeclaración de algunos gastos.

De esta manera, mientras que en la VII EPF bastaba con que al menos 10 tablas estuviesen correctamente contestadas para que un hogar tuviese todo el puntaje de esta libreta, en la VIII versión se buscó evaluar todas las tablas del cuestionario, ponderando cada una lo mismo²². De este

²²Con excepción de la tabla 22, en la que se registran gastos fuera del período de referencia y otros gastos que no han sido especificados en las tablas anteriores.

modo, para cada tabla se revisó: a) la consistencia del filtro con lo que se registra al interior de esta, y b) la existencia de respuesta válida en las variables de descripción del gasto, valor y establecimiento para cada una de las filas de la tabla. En caso de que existiera al menos una omisión en una fila de la tabla, esta era evaluada como sin respuesta en su conjunto.

Esta modificación suaviza los puntajes del cuestionario respecto a la evaluación realizada en la VII EPF, donde los cuestionarios podían tener el puntaje máximo (9 puntos) o bien ser evaluadas con el mínimo.

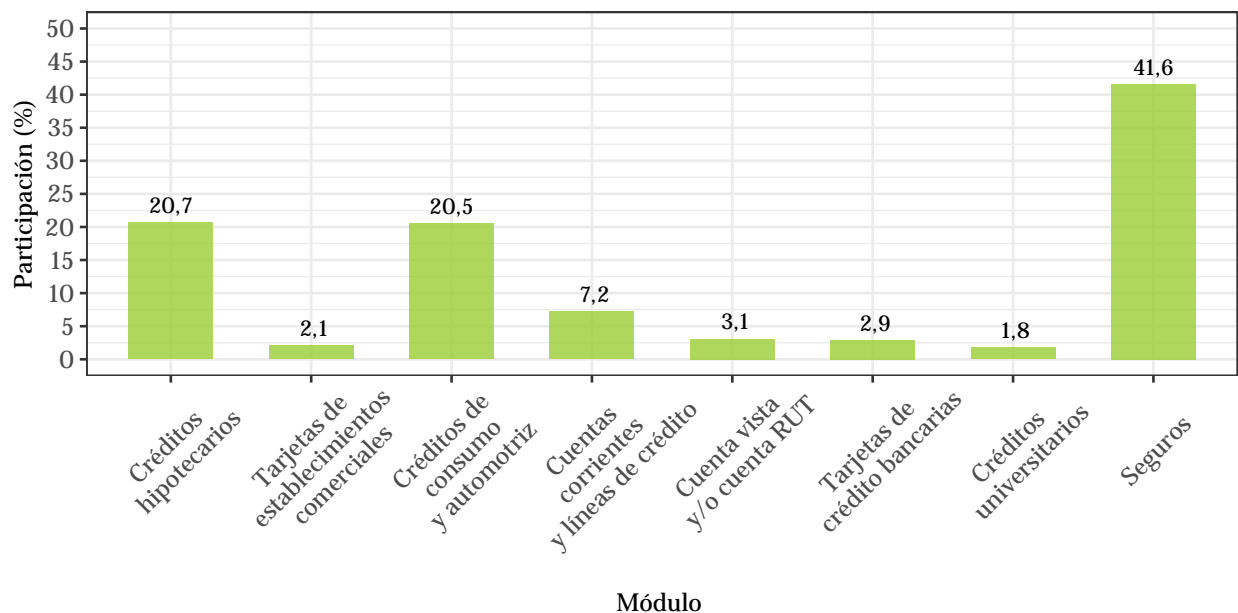
5.2.4 Libreta de Ingresos (LI)

Esta libreta no presenta cambios respecto a la evaluación de la VII EPF. Tanto los ingresos del trabajo asalariado como los ingresos del trabajo independiente son los más significativos en cuanto a la estructura de ingresos de los hogares, considerando los datos del primer semestre de la VIII EPF.

5.2.5 Libreta de Gastos en Servicios Financieros y Seguros (LSF)

En este cuestionario se reasignó el puntaje de los módulos de la libreta, de acuerdo con la participación del gasto en el primer semestre de la VIII EPF (figura 4), es decir, a la contribución de cada módulo en el gasto total de los hogares.

Figura 4: Participación de módulos de la Libreta Gastos en Servicios Financieros y Seguros en la captura del gasto promedio mensual de los hogares*



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

*Las estimaciones de la VIII EPF fueron calculadas con datos preliminares del primer semestre de la encuesta.

De esta manera, parte del puntaje quedó reasignado a los módulos Cuenta corriente y líneas de crédito, y Tarjetas de crédito bancarias, los que anteriormente no eran evaluados en la grilla de la VII EPF. Se mantuvieron los módulos Crédito hipotecario, Crédito de consumo o automotrices y Seguros. Con esta decisión, no obstante, quedaron fuera los módulos de Tarjetas de establecimientos comerciales y Créditos universitarios.

5.3 Criterios de evaluación del registro diario de gastos

En la VII EPF la evaluación del registro de gastos en la LGI no consideraba el número de integrantes del hogar mayores de 15 años²³. De esta forma, si al menos uno de los integrantes contestaba la LGI registrando como mínimo cuatro días, y al menos uno de estos correspondía a un día de fin de semana, la información del hogar era calificada con 30 puntos (puntaje máximo para este cuestionario). Como consecuencia, los hogares unipersonales eran evaluados con mayor exigencia que aquellos compuestos por más de una persona elegible. Por ejemplo, si un hogar unipersonal entregaba su libreta con los gastos registrados todos los días, obtenía el mismo puntaje que un hogar compuesto por dos personas en donde solo uno de los integrantes había anotado sus gastos.

Como medida para contrarrestar este desequilibrio, se decidió introducir dos nuevos criterios en la evaluación de la respuesta en la LGI:

- a) Considerar para el cálculo del puntaje la cantidad de integrantes que efectivamente contestaron el instrumento respecto al total elegible de cada hogar;
- b) Asignar distinta ponderación a los integrantes de acuerdo con su contribución en el gasto total del hogar registrado en la LGI.

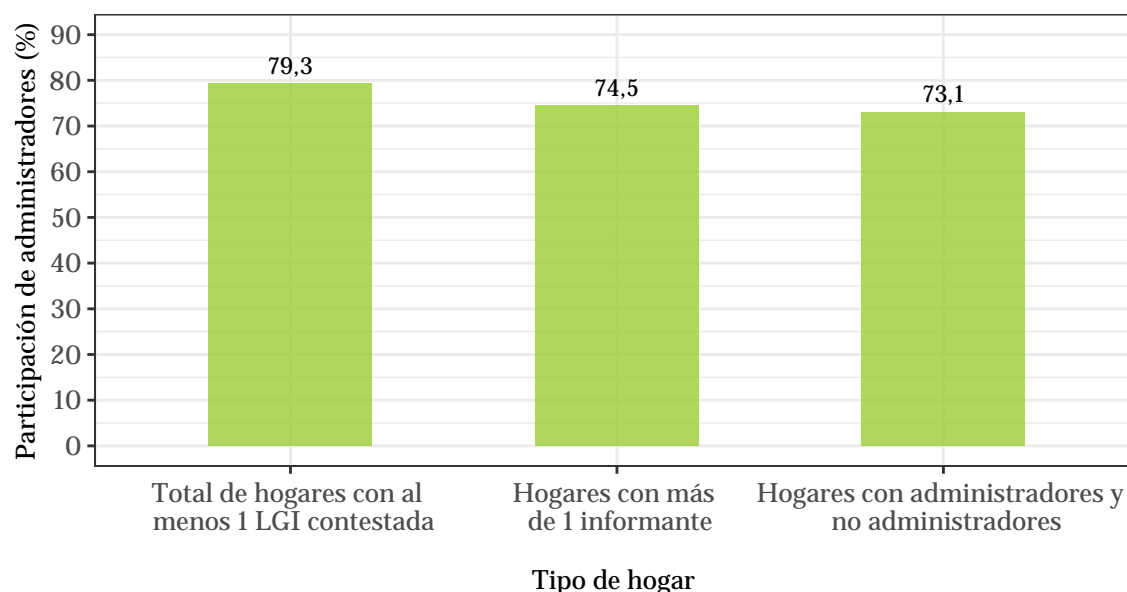
Para abordar el punto a) bastó identificar la cantidad de personas elegibles en el RPH. Para el punto b) se decidió utilizar la variable que indica quién o quiénes son administradores de gasto. Esta pregunta corresponde a una declaración que los propios informantes realizan respecto del manejo de los gastos en su hogar²⁴, siendo el único requisito para ser considerado como tal es que sean de 15 años o más.

En la VII EPF, la figura de administrador de gastos ya aparecía como relevante en su contribución al gasto capturado en la LGI, gastando en promedio un poco más del doble que quienes no eran administradores. A su vez, de acuerdo con los datos del primer semestre de la VIII EPF, el aporte relativo del gasto de los administradores al gasto total del hogar registrado en esa libreta los ubica como los principales contribuyentes. Considerando todos los hogares entrevistados en donde al menos existe una LGI respondida, su aporte representa en promedio el 79,3% del total del gasto del hogar registrado en la libreta (figura 5).

²³De acuerdo con el *Manual de Trabajo de Campo para Encuestadores de la VIII EPF*, los encuestadores deben entregar una LGI a “aquellos miembros del hogar de 15 años o más, que se encuentren en condiciones físicas y/o mentales de registrar sus gastos” (INE, 2016, pág. 94). Este mismo criterio se utilizó para la VII EPF.

²⁴La pregunta se encuentra en el cuestionario de RPH y se encuentra fraseada como sigue: “¿Quién es la persona que habitualmente realiza las compras del hogar?”.

Figura 5: Participación relativa de los administradores al gasto registrado en la Libreta de Gastos Individuales*



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

*El número total de hogares para cada barra es de 7.711, 6.247 y 4.889, respectivamente.

Para el cálculo de estos porcentajes se consideró el gasto registrado en la LGI con ajuste temporal e imputado para la corrección de la no respuesta.

Las estimaciones de la VIII EPF fueron calculadas con datos preliminares del primer semestre de la encuesta.

Respecto al gasto promedio, los datos del primer semestre de la VIII EPF señalan que la relación entre el gasto de administradores y no administradores es aproximadamente de una razón de 5:2 (cuadro 13).

Cuadro 13: Gasto promedio mensual registrado en la Libreta de Gastos Individuales, según condición de administración del gasto*

Tipo de informante	Total personas que respondieron LGI	Total personas en hogares con más de un informante elegible	Total personas en hogares con administradores y no administradores
Es administrador	115.029	117.051	121.510
No es administrador	45.997	45.980	54.595
Total personas	83.897	82.132	88.328

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

* Las estimaciones de la VIII EPF fueron calculadas con datos preliminares del primer semestre de la encuesta.

Resulta evidente con estos datos la necesidad de ponderar de manera distinta la falta de respuesta de aquellos que son administradores versus aquellos que no lo son. Para la VIII EPF se analizaron dos opciones de ponderación: a) establecer una ponderación fija al administrador del gasto o b) una ponderación variable que se adapte a la cantidad de integrantes que son administradores y los

que no. Considerando la contribución relativa al gasto de este tipo de informantes, se analizó una ponderación de 73% para las libretas contestadas por los administradores y 27% para el resto.

La forma de cálculo para cada una de estas alternativas puede ser expresada como sigue:

a) Ponderación fija

$$\text{Puntaje final LGI} = \frac{(73\% * Pa)}{x} + \frac{(27\% * Pna)}{y}$$

Donde:

x = Número de Administradores de gasto del hogar

y = Número de No Administradores de gasto del hogar

Pa = Puntaje del Administrador de gasto del hogar (o la suma de estos)

Pna = Puntaje del No Administrador de gasto del hogar (o la suma de estos)

b) Ponderación relativa al número de elegibles

$$\text{Puntaje final LGI} = \frac{\left[\left(\frac{2x}{2x+y}\right) * Pa\right]}{x} + \frac{\left[\left(\frac{y}{2x+y}\right) * Pna\right]}{y}$$

Donde:

x = Número de Administradores de gasto del hogar

y = Número de No Administradores de gasto del hogar

Pa = Puntaje del Administrador de gasto del hogar (o la suma de estos)

Pna = Puntaje del No Administrador de gasto del hogar (o la suma de estos)

Para ejemplificar ambas alternativas, se puede considerar el siguiente caso hipotético. En un hogar de cuatro integrantes, de los cuales dos de ellos son elegibles para contestar la LGI y solo uno de ellos la responde de forma válida²⁵, afortunadamente quien responde es el administrador de gastos del hogar²⁶. Mientras que en la VII EPF este hogar tendría el puntaje completo, el criterio de ponderación fija penalizaría en 8 puntos a este hogar por la falta de respuesta. En el caso de la ponderación variable (o relativa), esta penalización sería de 10 puntos en total. El cuadro 14 resume el puntaje en la LGI para cada uno de estos criterios.

Cuadro 14: Puntajes de la Libreta de Gastos Individuales según los distintos criterios de ponderación (Administrador)

Criterios	Puntaje LGI Administrador	Puntaje LGI No Administrador	Puntaje Final
Criterio VII EPF	30	0	30
Criterio VIII EPF Ponderación fija	30	0	22
Criterio VIII EPF Ponderación variable	30	0	20

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

²⁵Es decir, tiene 4 días de registro válido (con o sin gasto) y al menos uno de esos días corresponde a un fin de semana.

²⁶En la VIII EPF, el administrador del gasto corresponde a la persona que habitualmente realiza las compras del hogar, pudiendo existir más de uno por hogar. En el caso de este ejemplo existe solo un administrador del hogar.

En el cuadro 15 se observa qué ocurre con el ejemplo anterior: si el integrante del hogar que responde la LGI no es el administrador del gasto, el puntaje cambia radicalmente en relación con lo observado en el criterio de la antigua versión de la encuesta.

Cuadro 15: Puntajes de la Libreta de Gastos Individuales según los distintos criterios de ponderación (No administrador)

Criterios	Puntaje LGI Administrador	Puntaje LGI No Administrador	Puntaje Final
Criterio VII EPF	0	30	30
Criterio VIII EPF Ponderación fija	0	30	8
Criterio VIII EPF Ponderación variable	0	30	10

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

6 Plan de análisis

El análisis del efecto que tiene para la encuesta el cambio en los criterios de evaluación puede ser desagregado en tres motivaciones analíticas. Por un lado, se busca evaluar si resulta compatible con la muestra objetivo esperada establecer criterios más estrictos para la evaluación del nivel de no respuesta al ítem de la encuesta. En este sentido, resulta necesario analizar cómo varía la tasa de respuesta bajo distintos niveles de exigencia.

Por otro lado, interesa analizar si la modificación de los criterios de reclasificación de unidades entrevistadas a rechazo tiene algún efecto en los indicadores de gasto de los hogares. Si la recodificación excluye unidades que poseen características muy distintas a los hogares entrevistados, esto podría introducir un sesgo en las estimaciones. De esta manera, resulta fundamental comparar no solo si es que se observan cambios en las estimaciones de gasto dependiendo del nivel de exigencia, sino que también las características sociodemográficas de los hogares entrevistados versus aquellos que son recodificados como *break-off*.

Finalmente, para una encuesta periódica como la EPF²⁷, interesa analizar si es que la aplicación de estos cambios habría afectado los resultados de sus versiones anteriores. Tanto para los indicadores de logro como para los de gasto, interesa conocer cuáles habrían sido los resultados de la VII EPF con los nuevos criterios propuestos.

Para cubrir las dos primeras motivaciones, se definieron distintos escenarios que combinan los nuevos criterios de evaluación con los utilizados en la VII EPF. En total se evaluaron 5 escenarios con 5 metodologías de evaluación distintas, utilizando para esto la totalidad de los hogares que respondieron al menos el Registro de Personas en el Hogar. Para la comparación se definió un escenario control, que utiliza los mismos criterios de la VII versión de la encuesta. El cuadro 16 detalla los distintos criterios para cada escenario.

En términos generales, si bien todos los escenarios alternativos proponen cambios respecto a la penalización por el nivel de no respuesta al ítem para la calificación de cada unidad, los escenarios C y E son los que incorporan mayores ajustes respecto a los criterios utilizados en la VII EPF, modificando no solo la ponderación de las variables al interior de cada libreta, sino que también ajustando los pesos de cada una, de acuerdo con su contribución en la captura del gasto.

Todos los escenarios tienen en común que incorporan un mayor nivel de exigencia respecto a la evaluación de la LGI. Como ya se ha dicho, esto responde a la necesidad de vincular la evaluación de este instrumento al número de integrantes del hogar que debía reportar información de gastos.

²⁷Tal como se mencionó anteriormente, hasta la VI EPF (2006-2007) la periodicidad de la encuesta fue de 10 años. Posterior a eso, la VII EPF (2011-2012) fue la primera en levantarse con una periodicidad de cinco años. A su vez, la VIII EPF (2016-2017) fue la segunda en responder al compromiso OCDE del levantamiento quinquenal.

Cuadro 16: Comparación de criterios entre los distintos escenarios de evaluación considerados

Escenario	Ponderación por libreta	Ponderación de c/variable	Criterio evaluación LGI
Escenario A (control)	Utiliza los criterios aplicados en la VII EPF.	Utiliza los criterios aplicados en la VII EPF.	Utiliza los criterios aplicados en la VII EPF.
Escenario B	Utiliza los criterios aplicados en la VII EPF.	Redistribuye los puntajes al interior de cada libreta en función de la contribución del gasto en la VIII EPF.	El puntaje obtenido se pondera de acuerdo con el n° de integrantes elegibles, otorgando un 73% a los administradores de gasto y un 27% a quienes no lo son*.
Escenario C	Reasigna los puntajes en función de la contribución del gasto en la VIII EPF.	Redistribuye los puntajes al interior de cada libreta en función de la contribución del gasto en la VIII EPF.	El puntaje obtenido se pondera de acuerdo con el n° de integrantes elegibles, otorgando un 73% a los administradores de gasto y un 27% a quienes no lo son*.
Escenario D	Utiliza los criterios aplicados en la VII EPF.	Redistribuye los puntajes al interior de cada libreta en función de la contribución del gasto en la VIII EPF.	El puntaje obtenido se pondera de acuerdo con el n° de integrantes elegibles, donde el puntaje de un administrador de gastos pondera el doble que un integrante que no lo es†.
Escenario E	Reasigna los puntajes en función de la contribución del gasto en la VIII EPF.	Redistribuye los puntajes al interior de cada libreta en función de la contribución del gasto en la VIII EPF.	El puntaje obtenido se pondera de acuerdo con el n° de integrantes elegibles, donde el puntaje de un administrador de gastos pondera el doble que un integrante que no lo es†.

Fuente: Instituto Nacional de Estadísticas (INE) - VII y VIII Encuesta de Presupuestos Familiares (EPF).

* Esta distinción de puntaje se establece sobre la base del monto de gasto promedio que representaron los administradores de gastos y los no administradores respecto al gasto total de la VIII EPF.

† Esta distinción de puntaje se establece sobre la base que el monto de gasto promedio de los administradores de gastos representó el doble de gasto que el monto promedio de los no administradores, de acuerdo con los datos de la VIII EPF.

Se realizaron 4 ejercicios de comparación con los escenarios definidos. En relación con el efecto en los indicadores de respuesta de la encuesta, se realizó un análisis de dominancia para comparar el porcentaje de hogares reclasificados como *break-off* respecto del total de hogares entrevistados parcial y completamente²⁸. Con este análisis se buscó revisar si en la elección de criterios para la evaluación de la no respuesta al ítem, un escenario será siempre más exigente respecto a otro para cualquier puntaje de corte que se elija, evaluando así la sensibilidad de las metodologías de evaluación.

Con este primer ejercicio se identificó cuál de los escenarios resultaba ser el más exigente para distintos puntajes de corte, seleccionando un escenario que permitiera la comparación con el escenario control para tenerlo como referencia en el resto de los ejercicios.

Un segundo ejercicio consistió en analizar el cambio en la tasa de respuesta de la encuesta al compararla con el escenario base. Aunque la decisión de los criterios de exigencia no pasa por determinado valor de este indicador, sí es necesario observar cómo este se ve afectado por un cambio en los valores de corte.

Un tercer ejercicio consistió en la comparación de los indicadores de gasto de la encuesta. Al respecto, uno de los aspectos más sensibles de la EPF corresponde a la estructura de gasto de los hogares, teniendo en cuenta que esta se utiliza como base para la elaboración de los ponderadores de la canasta

²⁸La proporción se calculó como la división entre el número de hogares calificado como *break-off* para determinado puntaje de corte respecto del total de hogares en los que al menos se contestó la libreta de RPH.

del Índice de Precios al Consumidor. Para este caso, se comparó la estructura de gasto por divisiones de la CCIF²⁹, las que se especifican en el cuadro 17.

Cuadro 17: Divisiones de la Clasificación de Consumo Individual por Finalidades (CCIF)

Divisiones CCIF	
División 1	Alimentos y bebidas no alcohólicas
División 2	Bebidas alcohólicas, tabaco y estupefacientes
División 3	Prendas de vestir y calzado
División 4	Vivienda, agua, electricidad, gas y otros combustibles
División 5	Muebles. Equipamiento y conservación ordinaria del hogar
División 6	Salud
División 7	Transporte
División 8	Comunicaciones
División 9	Recreación y cultura
División 10	Educación
División 11	Restaurantes y hoteles
División 12	Bienes y servicios diversos
División 13	Gasto de consumo individual de las instituciones sin fines de lucro que sirven a los hogares
División 14	Gasto de consumo individual del gobierno general

Fuente: Instituto Nacional de Estadísticas (INE) - VII y VIII Encuesta de Presupuestos Familiares (EPF).

Como cuarto ejercicio, se realizó una caracterización sociodemográfica de los hogares que, aumentando el nivel de exigencia, eran reclasificados como *break-off*. Con esto, se buscó analizar si es que los hogares excluidos presentaban respuestas muy distintas a aquellos hogares entrevistados. De manera indirecta, este ejercicio buscó evaluar la existencia de sesgo por no respuesta. Este ejercicio solo incluye la comparación entre las unidades entrevistadas y aquellas que cambiaron su estado final de entrevistadas a rechazo (*break-off*). Quedan fuera de esta comparación aquellas unidades que rechazaron explícitamente, desde un inicio, participar de la encuesta³⁰.

Para la selección de un puntaje de corte se evaluaron los distintos resultados de los ejercicios realizados, con lo que se escogió un nivel de exigencia para el nivel de no respuesta parcial permitido para la VIII EPF.

²⁹La EPF utiliza la Clasificación del Consumo Individual por Finalidades (CCIF) para clasificar el consumo final de los hogares. La CCIF es una clasificación funcional que cubre los gastos de consumo individual en línea con el Sistema de Cuentas Nacionales y consta de 14 divisiones. Para efectos de la VIII EPF se utilizaron solo las primeras doce divisiones, correspondientes al consumo de los hogares (las divisiones 13 y 14 se refieren al gasto de consumo individual de las Instituciones Sin Fines de Lucro que Sirven a los Hogares y el Gobierno). Los mayores niveles de agregación de la CCIF son las 12 divisiones, las que se forman de grupos. Los grupos están compuestos por clases y estas, a su vez, por subclases. Las subclases son el resultado de la agregación de los productos. Los índices de nivel mayor de agregación (división, grupo y clase) permiten la comparación internacional, en cambio, a nivel de subclase y producto quedan sujetos a la definición de cada país, según sus características propias (INE, 2013).

³⁰Los resultados de este ejercicio no permiten asegurar que no exista sesgo por no respuesta, ya que no se dispone de información respecto a los hogares que rechazaron sin entregar algún tipo de información.

En relación con el tercer interés analítico de este documento, para analizar cómo hubiese afectado este cambio de criterios en los resultados de la VII EPF, se plantean dos ejercicios que comparan las tasas de respuesta y los indicadores de gasto entre la VII EPF y la VIII EPF. Considerando los ajustes metodológicos entre una y otra versión, para resguardar la validez de esta comparación los resultados de ambas versiones han sido estimados considerando: a) el método de imputación de la VII EPF (Factor de no respuesta³¹) y b) el cálculo de los factores de expansión considerando la metodología empleada en la séptima versión de la encuesta³².

³¹El factor de no respuesta es un método diseñado para mitigar el problema de la no respuesta parcial de la LGI. Se basa en el supuesto de que los días con registro de una libreta pueden representar a los días que no tienen registro. El método consiste en reponderar la información de gastos que tiene una libreta, para así completar la quincena. Para más información de este método revisar INE (2019b) o INE (2014b).

³²Para más información sobre el cálculo de los factores de expansión revisar la *Metodología de la VII EFP* (INE, 2013).

7 Evaluación de escenarios

7.1 Indicadores de respuesta

7.1.1 Análisis de dominancia

El análisis de dominancia (figura 6) permite identificar cuál sería el porcentaje de hogares que pasaría a ser parte del *break-off* de la encuesta si se modifica el puntaje de corte entre los 50 y 100 puntos en todos los escenarios evaluados. Nótese que, a medida que aumenta el valor al que se realiza este corte, de manera independiente al escenario evaluado, aumenta la proporción de hogares que serían reclasificados como *break-off*. Por ejemplo, respecto del total de hogares que contestaron al menos el RPH, si el puntaje de corte se estableciera en 50 puntos, el porcentaje de hogares recodificado como *break-off* estaría alrededor de un 11%. Por otro lado, si se indicara como estándar que para ser considerado como entrevistado los hogares debiesen tener respuesta en todos los ítems evaluados (exigencia de 100 puntos), la cantidad de hogares reclasificado como *break-off* alcanzaría valores cercanos al 60% (cuadro 18).

Cuadro 18: Porcentaje de hogares reclasificados como *break-off* según puntajes de corte

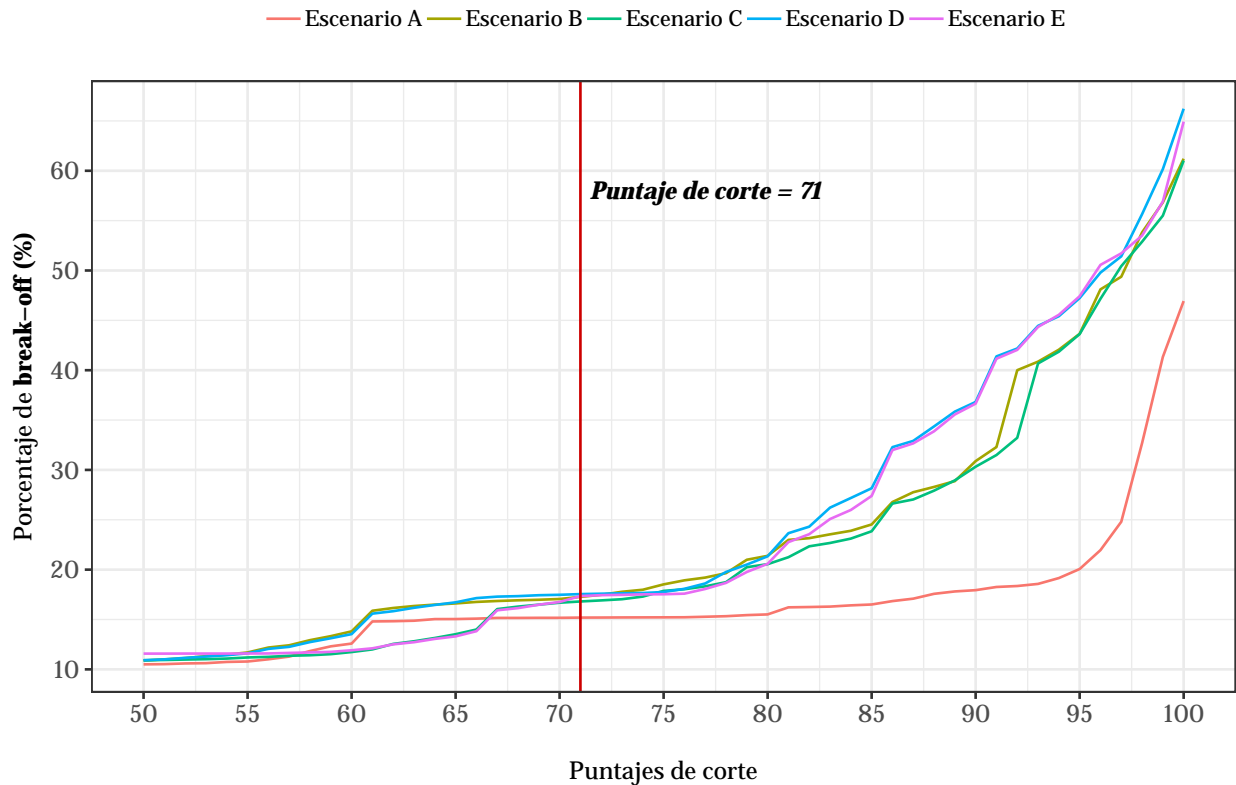
Escenarios/Puntajes	50 ptos.	60 ptos.	70 ptos.	80 ptos.	90 ptos.	100 ptos.
Escenario A	10,5	12,6	15,2	15,5	17,9	46,9
Escenario B	10,9	13,8	17,1	21,4	30,9	61,2
Escenario C	10,9	11,7	16,7	20,5	30,3	61,0
Escenario D	10,9	13,5	17,5	21,3	36,8	66,2
Escenario E	11,6	11,9	16,8	20,6	36,6	64,9

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

La figura 6 ilustra que el escenario A es el menos exigente en comparación con el resto de los escenarios. La línea vertical corresponde al puntaje de corte de 71 puntos, utilizado en la VII EPF. En este corte, se observa que los escenarios alternativos tienen un porcentaje de *break-off* mayor al escenario de control, pero entre estos la diferencia en la proporción de hogares reclasificados es mucho menor.

Si se observa ahora la pendiente de las curvas de la figura 6, se notará que la variación del *break-off* para el escenario control describe una curva más convexa que el resto de los escenarios. Esta forma refleja la dicotomía con la cual se evaluaba la Libreta de Gastos Individuales bajo los criterios empleados en la VII EPF. Si ninguno de los integrantes contestaba este cuestionario, inmediatamente el hogar era reclasificado como *break-off*.

Figura 6: Porcentaje de hogares *break-off* para los valores de corte de 50 a 100 puntos



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

Una situación distinta ocurre con los escenarios alternativos cuyas pendientes son mucho más pronunciadas que la del escenario A. Un criterio que considere al número de integrantes elegibles en el cálculo del puntaje de la LGI suaviza la curva de incremento del *break-off* al aumentar el valor de corte.

Al comparar ambos criterios de evaluación para la LGI, el más exigente es el que combina la ponderación de la cantidad de integrantes elegibles del hogar con el prorrateo del puntaje del administrador del gasto versus aquellos que no lo son. Los escenarios que utilizan este criterio de evaluación (escenarios D y E) son los que presentan los mayores porcentajes de *break-off* en los puntajes de corte sobre 80 puntos.

La diferencia entre estos dos escenarios (D y E) es leve y la cercanía de las curvas indica que no es posible establecer una dominancia absoluta, como en el caso de los escenarios alternativos frente al escenario de control. En este sentido, la preferencia de un escenario por sobre otro no impactaría de manera sustantiva a la proporción de hogares reclasificados, como sí ocurriría en caso de mantener los criterios de evaluación de la VII EPF.

El objetivo de buscar una nueva forma de evaluación de los datos es aplicar criterios más exigentes para penalizar por el nivel de no respuesta parcial. Teniendo esto en cuenta, los escenarios D y E aparecen como los mejores candidatos, siendo los que presentan una mayor recodificación en los

puntajes más altos de evaluación (sobre 80 puntos).

De manera preliminar, es posible concluir que la evaluación aplicada en la VII EPF es menos estricta que las alternativas propuestas. A su vez, estas presentan importantes similitudes, por lo que la elección de una por sobre las demás dependerá más de un criterio normativo y de que sea preferible de aplicar. En este sentido, el escenario elegido para continuar con el resto de los ejercicios de evaluación es el escenario E. Este escenario incluye como parte de sus criterios de evaluación actualizaciones relevantes, tales como la modificación de las ponderaciones por cuestionarios e indicador y la innovación en la forma de evaluar la LGI, considerando la estructura del hogar.

Tomando en cuenta que el propósito del documento es evaluar en qué medida es posible utilizar una evaluación más exigente de los datos y un puntaje de corte más estricto (más alto), en los siguientes apartados se trabajará sobre la selección de un puntaje de corte por sobre el valor umbral de la VII EPF (71 puntos). El rango de puntajes elegido para este análisis va desde 71 a 81 puntos. Este intervalo fue escogido en función del puntaje de corte aplicado en la VII EPF, es decir, se utiliza como mínimo los 71 puntos y como máximo, los 81 puntos. Esto, debido a que se ajusta a los objetivos de diseño muestral de la encuesta para cumplir un nivel adecuado de respuesta en relación al tamaño muestral objetivo del estudio.

7.1.2 Tasa de respuesta

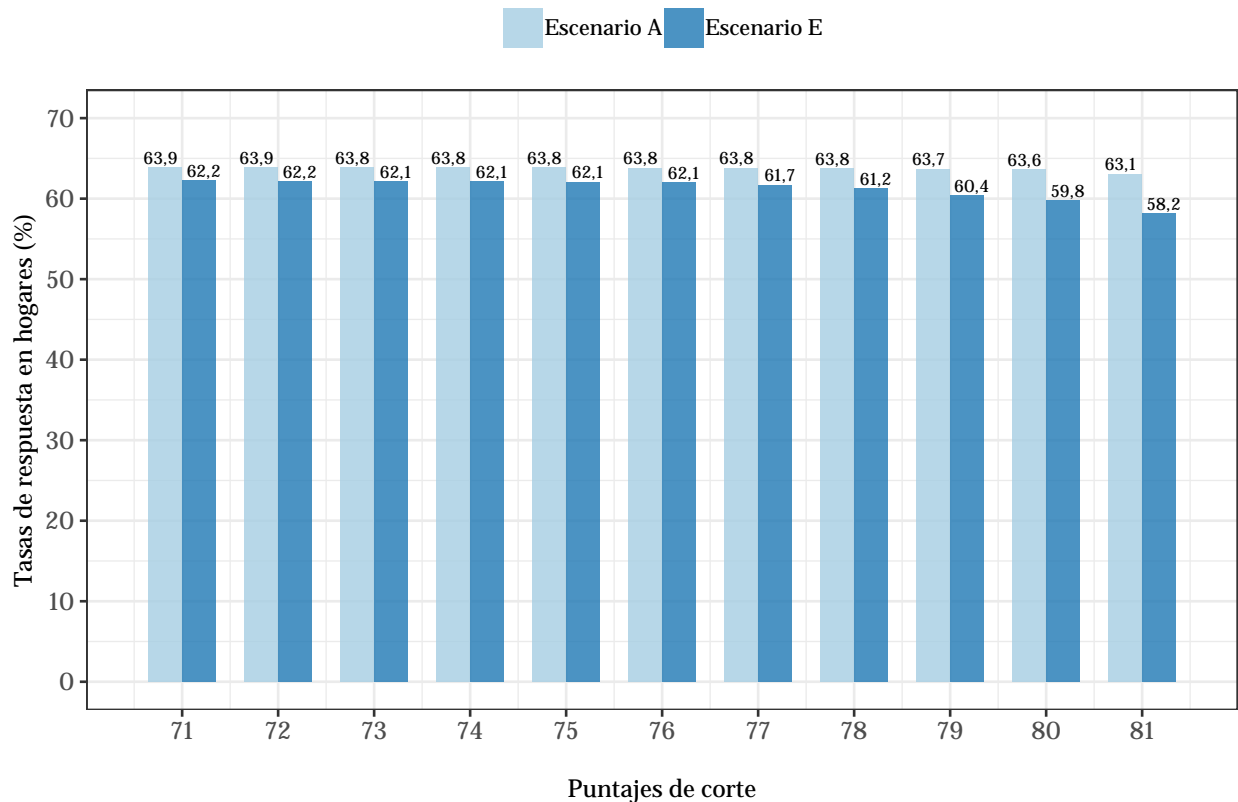
La tasa de respuesta se ve afectada por criterios de evaluación más exigentes (figura 7). A medida que se aumenta el puntaje de corte, la tasa de respuesta disminuye tanto en el escenario de control (A) como en el escenario E. Mientras que en el primero, para un puntaje de corte de 71 puntos, alcanza a un 63,9%, en el escenario E, para un puntaje de corte de 81 puntos, la tasa de respuesta baja a 58,2%.

Así, el criterio que más afecta a la tasa de respuesta es la elección de un puntaje de corte, observándose esto más claramente en el escenario E. En el escenario A se aprecia una pequeña diferencia de 0,7 puntos porcentuales (pp.) entre los 71 y 81 puntos, en tanto que en el escenario E la diferencia alcanza 4,2 puntos en el mismo rango de puntajes.

Cuando la comparación se realiza entre escenarios para un mismo puntaje de corte, si se considera un umbral de 71 puntos, la distancia entre el escenario A y E es de solo 1,7 puntos. La diferencia aumenta a medida que se incrementa el puntaje de corte, hasta llegar a los 81 puntos, en que la diferencia entre ambos escenarios alcanza 4,9 pp.

Lo anterior se relaciona con lo presentado en la figura 6, asociado al porcentaje de *break-off* para distintos escenarios y puntos de corte. En el escenario A se observa que el porcentaje de hogares reclasificados como *break-off* no presenta variaciones importantes entre el rango de los 60 y 90 puntos. Por otro lado, el escenario E, así como el resto de los escenarios evaluados, presenta una curva con menor pendiente en la variación del porcentaje de *break-off*.

Figura 7: Tasa de respuesta para escenarios A y E, según puntajes de corte



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

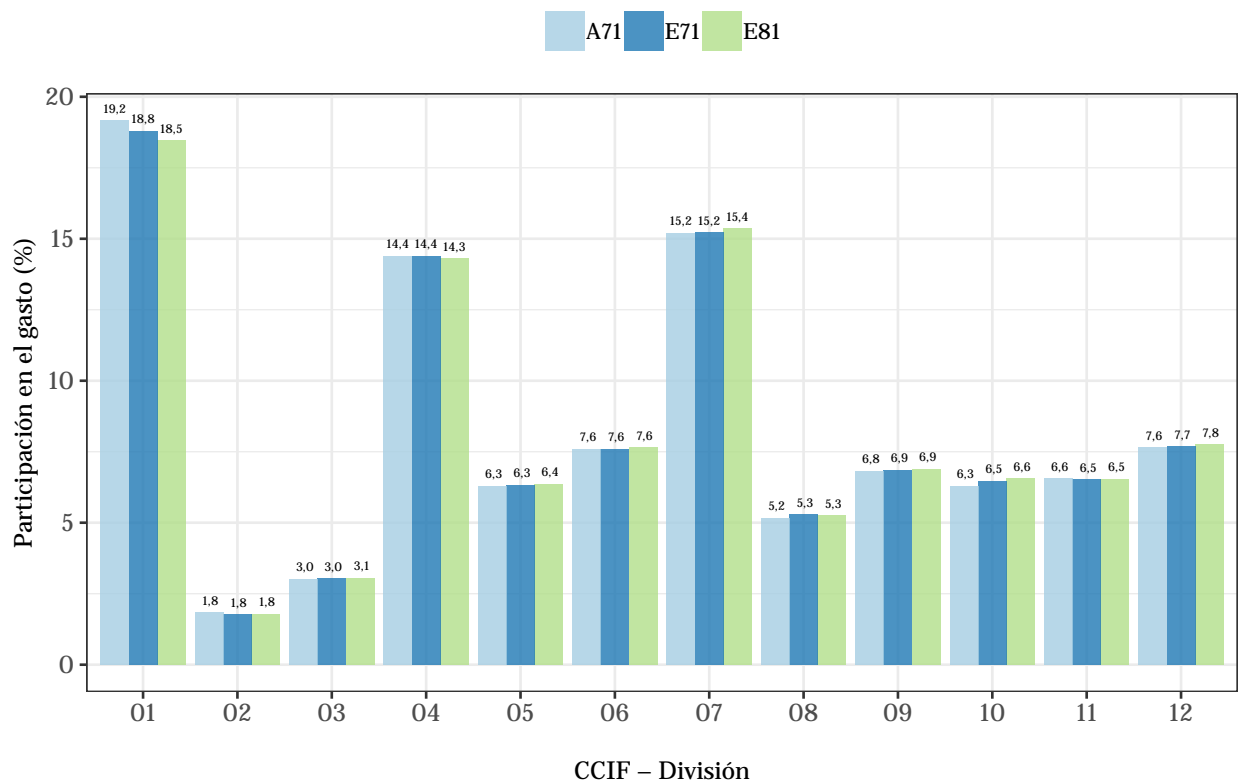
Mientras que en el escenario A la elección de un puntaje de corte más exigente no modifica de manera sustantiva el porcentaje de hogares clasificado como *break-off*, en el caso del escenario E esta elección se torna relevante por la cantidad de hogares que pasan a ser reclasificados dentro de la categoría de *break-off*. En este sentido, se estima que tanto este escenario como el resto de los escenarios propuestos son más sensibles al nivel de no respuesta parcial presente en las libretas, siendo considerados, por tanto, como métodos de evaluación más precisos que el aplicado en la VII EPF.

Una evaluación más exigente y también más acorde a las características de la encuesta afecta de manera negativa las tasas de respuesta. No obstante, para todo el rango de puntajes evaluado, el nivel de exigencia que se plantea permite alcanzar una tasa de respuesta adecuada a los tamaños muestrales objetivos del diseño de la encuesta. En este sentido, dado el mejor desempeño de los equipos de trabajo de campo respecto a la VII EPF, es posible adoptar un alto nivel de exigencia sin sacrificar precisión para las estimaciones de gasto.

7.2 Efectos en los indicadores de gasto de la encuesta

En contraste con la tasa de respuesta, tanto el gasto promedio de los hogares como su estructura no presentan diferencias importantes al comparar el escenario control con el escenario E. Este resultado sugiere que no hay diferencias en términos de las estimaciones de gasto al evaluar de forma más exigente el nivel de falta de respuesta al ítem. Como se observa en la figura 8, la estructura de gasto promedio mensual de los hogares clasificados como entrevistados es casi idéntica entre los escenarios, independiente del puntaje de corte elegido. Cabe destacar que las estimaciones de gasto que se realizan en este apartado utilizan las metodologías de imputación (*hot deck*) y de factores de expansión aplicadas en la VIII EPF, con el objetivo de evaluar únicamente el efecto en el cambio de la grilla en la estimación del gasto y su estructura.

Figura 8: Estructura del gasto promedio mensual de los hogares* para escenarios A y E, según división CCIF



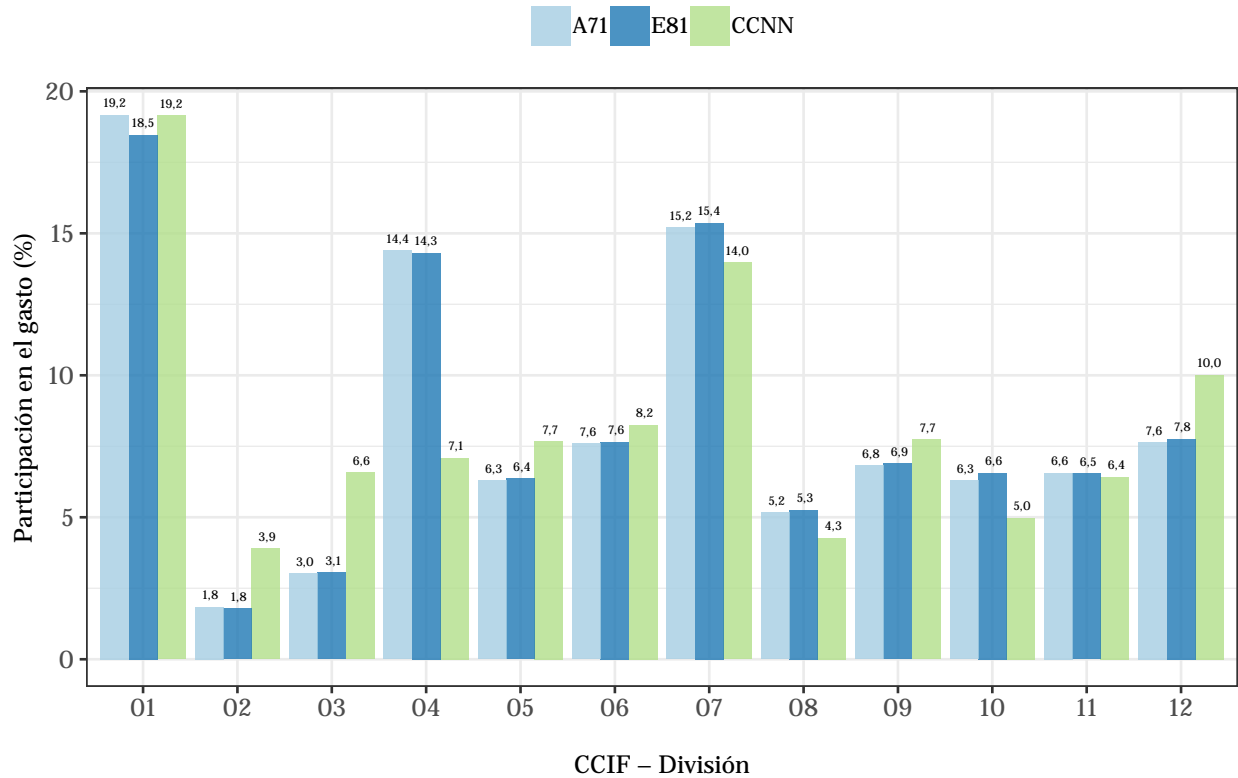
Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).
* Las estimaciones excluyen el arriendo imputado.

En vista de las similitudes observadas entre las estructuras de gasto comparadas en la figura anterior, en la figura 9 se incluye la estructura de gastos estimada a partir de las Cuentas Nacionales (CCNN). Es necesario advertir que los datos de las CCNN no son directamente comparables con los entregados por la VIII EPF, pues difieren en áreas de representatividad y la clasificación de algunos productos no permite una comparación estricta³³. Teniendo en cuenta lo anterior, se observa que la estructura

³³Para realizar la comparación se realizaron los siguientes ajustes a los datos de la VIII EPF: 1) Se eliminaron gastos

de gasto estimada a partir de las Cuentas Nacionales se encuentra en sintonía con las estructuras de gasto de los escenarios propuestos.

Figura 9: Estructura del gasto promedio mensual de los hogares* para escenarios A (71 puntos), E (81 puntos) y aproximación a la estructura de las Cuentas Nacionales (2015), según división CCIF



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).
* Las estimaciones excluyen el arriendo imputado.

Para comparar el gasto promedio mensual por hogar, adicionalmente se consideró un escenario en el que todos los hogares son incluidos en el análisis, independiente del nivel de no respuesta (o en otras palabras, un escenario donde no se aplicaron criterios de reclasificación).

El cuadro 19 muestra cuál sería el gasto promedio mensual por hogar al modificar el escenario y el puntaje de corte; además, señala la cantidad de hogares utilizados para hacer el análisis. Se observa que el promedio de gasto de los hogares clasificados como entrevistados en el escenario E es solo levemente mayor al promedio del escenario control. A su vez, cuando se estima el gasto con el total de hogares (última columna), se registra un gasto promedio menor que el resto de los escenarios, no obstante tiene un mayor número de hogares. Esta diferencia podría sugerir que la inclusión de

de la división 4 relacionadas con pequeñas reparaciones y mantenimiento del hogar, porque las CCNN los considera como formación bruta de capital fijo; 2) De la división 4 también se descontó un estimativo del monto total y arriendo imputado, debido a que la metodología de cálculo difiere; 3) Se eliminaron gastos de la división 7 asociados a automóviles y motocicletas usadas; 4) Se consideró que el gasto en educación (división 10) corresponde a un 80% de lo entregado por las CCNN, y 5) De la división 12 se eliminan los servicios asociados a la intermediación financiera y seguros de vida.

hogares con un alto nivel de no respuesta al ítem introduce un sesgo en las estimaciones de gasto, el que podría estar ocasionado por una subdeclaración de los informantes. Es decir, la estimación de gasto promedio de los hogares cambia considerablemente cuando se compara el gasto del total de hogares con el resto de los escenarios en donde sí existen criterios de reclasificación. Esto se refleja en la figura 10, donde no se observan diferencias significativas entre los escenarios, pero sí cuando se compara con el gasto promedio del total de hogares.

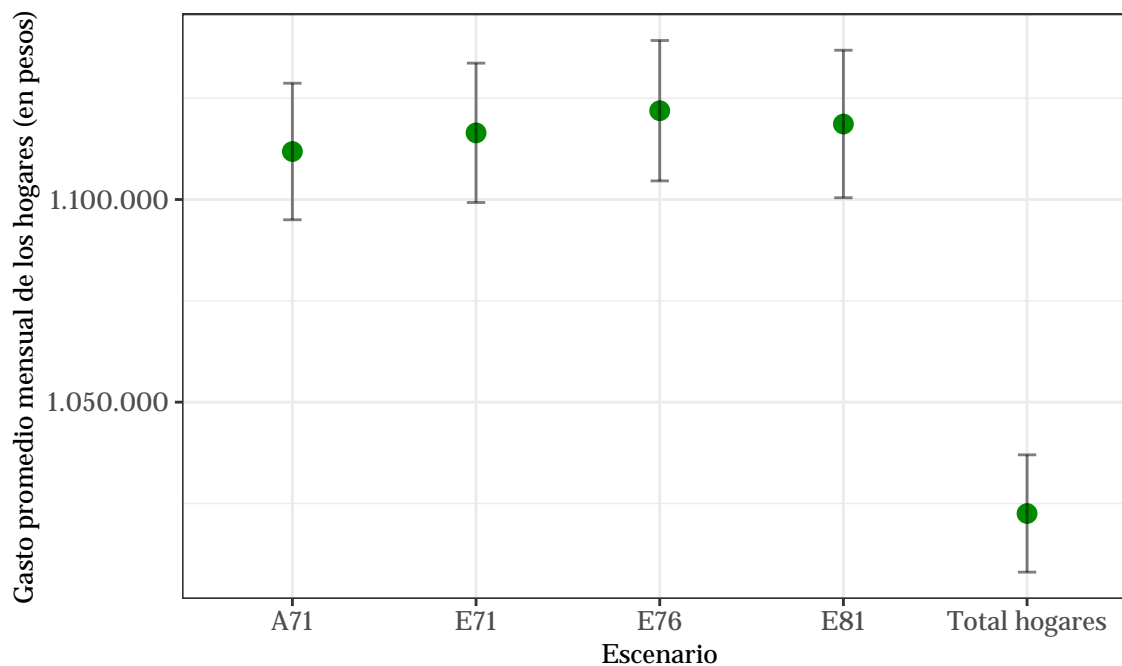
Cuadro 19: Gasto promedio mensual de los hogares*, según escenario de evaluación

Características	Escenario A71	Escenario E71	Escenario E76	Escenario E81	Total hogares
Gasto promedio mensual hogar	1.111.852	1.116.459	1.121.925	1.118.640	1.022.525
Número de hogares	15.683	15.677	15.239	14.283	18.492

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

* Las estimaciones excluyen el arriendo imputado.

Figura 10: Gasto promedio mensual de los hogares*, según escenarios A, E y Total de hogares



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

* Las estimaciones excluyen el arriendo imputado.

Del análisis de los distintos escenarios se concluye que es posible establecer criterios más exigentes sin afectar las estimaciones de gasto de los hogares, en términos del promedio mensual de gasto y de su estructura a nivel de divisiones. Las estimaciones del escenario control en comparación con las estimaciones del escenario E son casi idénticas y no presentan diferencias significativas. En tanto, al comparar los resultados entre el total de hogares y los distintos escenarios, se observan diferencias

significativas, lo que justifica utilizar un método que permita establecer una diferencia entre hogares cuya información resulta ser insuficiente y hogares que cuentan con un mínimo de datos, con el objetivo de no sesgar las estimaciones de la encuesta.

7.3 Caracterización sociodemográfica de los hogares por nivel de colaboración

Para evaluar de manera indirecta la existencia de un sesgo por no respuesta, en el siguiente apartado se identifica si existen diferencias significativas entre las características de aquellos hogares entrevistados versus aquellos que fueron reclasificados en la categoría de *break-off*. Para esto, se compara la información de un conjunto de variables sociodemográficas relacionadas con el nivel de gasto entre estos grupos de hogares³⁴. Lo que se espera del ejercicio es que los hogares entrevistados sean similares a los hogares que forman parte del *break-off* en el conjunto de características seleccionadas. Si bien observar esto no hace posible afirmar que no existe sesgo por no respuesta, es una aproximación para evaluar la magnitud del sesgo en los hogares excluidos.

Los cuadros 20 y 23 presentan la comparación de resultados para los distintos niveles de colaboración. Estos, en cada cuadro, se ordenan de la siguiente forma: las primeras dos columnas muestran los resultados para los hogares clasificados como entrevistados en los escenarios A y E. Las siguientes dos columnas exponen los resultados para los hogares que fueron reclasificados como rechazo en ambos escenarios. En la última columna se encuentran los resultados para el total de hogares, a los que no se les aplicó ninguna metodología de evaluación y fueron imputados con el método de *hot deck* para la corrección de la no respuesta parcial.

El cuadro 20 muestra el resultado de las variables sociodemográficas para ambos escenarios cuando el puntaje de corte es de 71 puntos. En términos del gasto, las unidades que son reclasificadas como *break-off* presentan un promedio considerablemente menor al de las entrevistadas. El ingreso este también es menor, no obstante, la brecha se reduce respecto a lo observado con el gasto.

Si se analizan otras características, como el nivel de escolaridad, la edad y el sexo de la jefatura del hogar, no se aprecian diferencias importantes entre los hogares entrevistados, aquellos reclasificados y el total de hogares sin utilizar criterios de reclasificación. Así, por ejemplo, en todos los grupos los hogares presentan en mayor proporción jefaturas masculinas. En relación con los años de escolaridad, el grupo que considera a todos los hogares sin reclasificación tiene un nivel levemente menor al de educación media completa (12 años).

Esta misma tendencia se ve cuando se analizan estas diferencias para un puntaje de corte de 81 puntos (cuadro 23). La diferencia más relevante se observa al comparar el gasto de las unidades reclasificadas como *break-off* en este puntaje de corte y las unidades que sufren esta recodificación cuando el umbral es de 71 puntos. Estas últimas presentan un promedio de gasto menor al de los hogares reclasificados como *break-off* cuando se establece un valor de corte de 81 puntos. Evidentemente, cuando se utiliza un criterio más estricto para distinguir entre la no respuesta el ítem y la no respuesta total, se incorporan a la categoría de *break-off* hogares que tienen un nivel

³⁴ Este ejercicio no incluye aquellas unidades que rechazaron explícitamente desde un inicio a participar de la encuesta.

de no respuesta parcial menor que aquellos que componen el grupo conformado por criterios menos exigentes.

Cuadro 20: Caracterización de los hogares, según puntaje de corte 71 puntos

Características	A71 - Entrevistada	E71 - Entrevistada	A71 - Break-off	E71 - Break-off	Total Hogares
Ingreso disponible promedio mensual de los hogares*	1.163.258,2	1.160.665,3	1.046.921,8	1.075.181,0	1.145.933,9
Gasto promedio mensual de los hogares*	1.089.339,1	1.092.920,5	640.667,2	684.425,8	1.022.525,1
Hombre JH (%)	58,5	58,5	57,2	57,4	58,3
Edad JH	51,6	51,4	49,5	50,6	51,3
Años de escolaridad JH	12,0	12,0	11,8	11,7	11,9
Hombre SP (%)	62,1	61,9	60,8	61,8	61,9
Edad SP	50,8	50,7	47,8	48,7	50,4
Años de escolaridad SP	12,4	12,4	12,3	12,2	12,4
Total de hogares	15.683,0	15.284,0	2.809,0	3.208,0	18.492,0

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

* Las estimaciones excluyen el arriendo imputado.

Cuadro 21: Caracterización de los hogares, según puntaje de corte 81 puntos

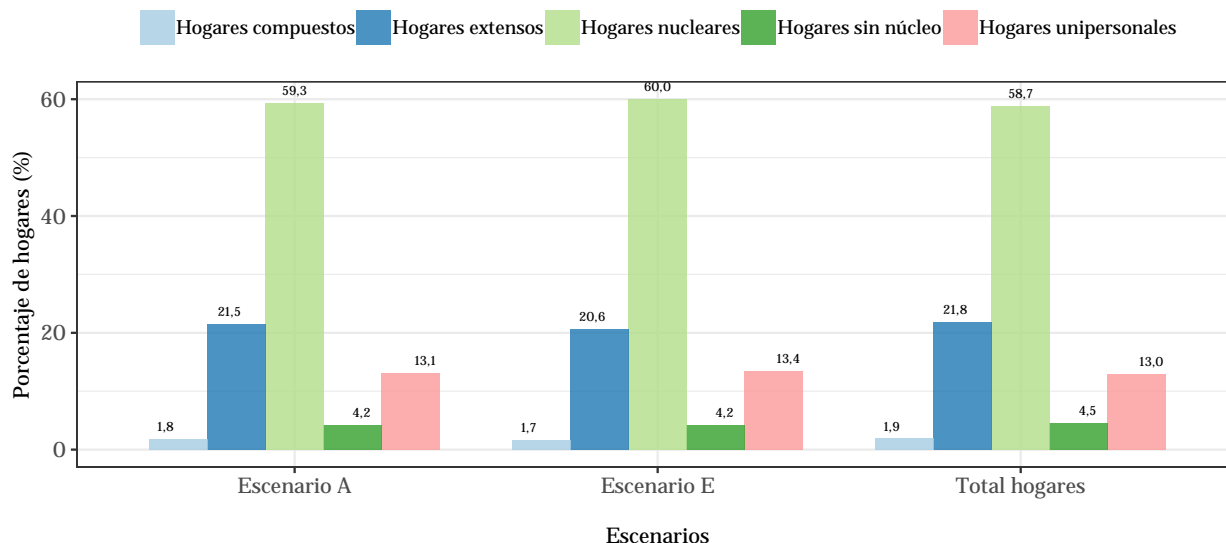
Características	A81 - Entrevistada	E81 - Entrevistada	A81 - Break-off	E81 - Break-off	Total Hogares
Ingreso disponible promedio mensual de los hogares*	1.161.390,8	1.153.807,3	1.064.482,1	1.119.285,4	1.145.933,9
Gasto promedio mensual de los hogares*	1.088.357,5	1.087.867,5	675.613,9	801.364,6	1.022.525,1
Hombre JH (%)	58,5	58,4	57,4	58,2	58,3
Edad JH	51,6	51,3	49,4	51,1	51,3
Años de escolaridad JH	11,9	12,0	11,9	11,6	11,9
Hombre SP (%)	62,1	61,6	61,0	62,9	61,9
Edad SP	50,9	50,7	47,8	49,4	50,4
Años de escolaridad SP	12,4	12,5	12,4	12,1	12,4
Total de hogares	15.493,0	14.283,0	2.999,0	4.209,0	18.492,0

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

* Las estimaciones excluyen el arriendo imputado.

La figura 11 muestra la distribución del tipo de hogares³⁵ para el total de los hogares y aquellos clasificados como entrevistados en los escenarios A y E. Para ambos escenarios, los hogares entrevistados no difieren de manera significativa en la distribución de esta variable. El tipo de hogar nuclear es el más frecuente, lo que se condice con la tendencia en la región hace más de 10 años (Ullman, Maldonado, & Nieves, 2014). Le siguen los hogares extensos, con valores cercanos al 21%, y los hogares unipersonales, con 13%.

Figura 11: Tipos de hogares para puntaje de corte 71 puntos, según escenarios A, E y Total de hogares



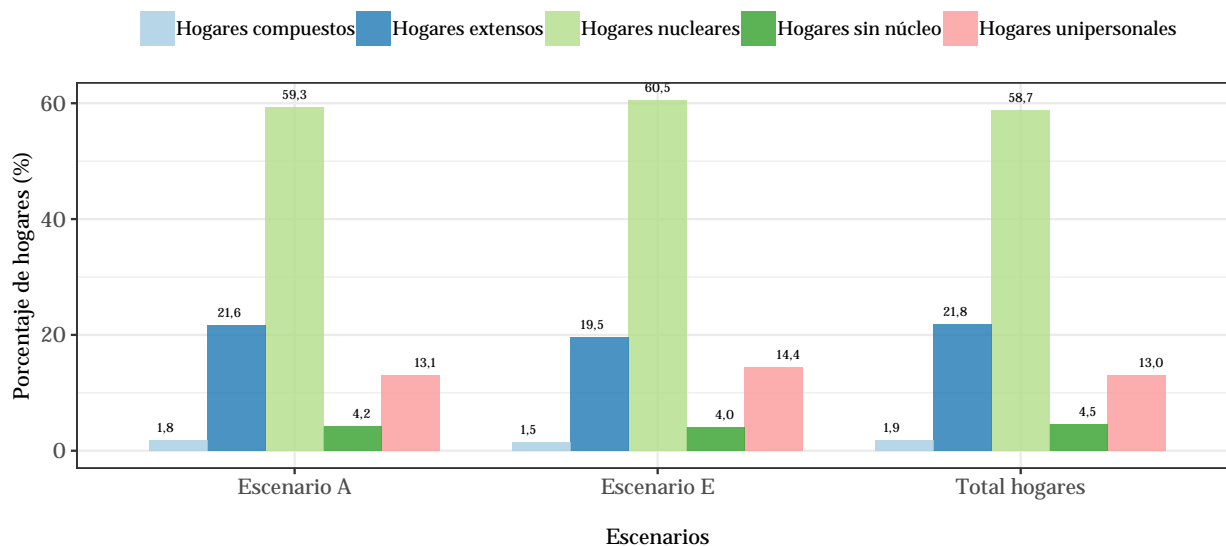
Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

Para el puntaje de corte de 81 puntos se observa una leve diferencia en la presencia de hogares extensos si se compara el total de hogares respecto de los entrevistados (figura 12). Mientras que los hogares extensos representan el 21,6% en el caso de las unidades entrevistadas del escenario A, este grupo baja a un 19,5% en el escenario E. Por otro lado, los hogares nucleares aumentan su participación en las unidades entrevistadas, llegando a 59,3% en el escenario control y a 60,5% en el escenario E, en tanto que en el total de las unidades representa el 58,7% de los casos.

Si bien estas diferencias son menores, logran reflejar en parte la dificultad que implica para la encuesta el mantener el contacto con ciertos tipos de hogares. Cabe recordar que la fase de aplicación de la EPF se encuentra organizada en 4 visitas en las que el hogar debiese responder un total de 6 cuestionarios. Es probable que en los hogares extensos sea más difícil para los encuestadores lograr la participación de todos los integrantes elegibles para responder la libreta de registros diarios de gasto (LGI) que en hogares nucleares, donde la cantidad de personas en promedio es menor y la dinámica de relaciones al interior de esos hogares también es distinta.

³⁵Según la definición utilizada por la Cepal los hogares pueden ser clasificados en 5 categorías de acuerdo con su estructura: 1) Hogares unipersonales: constituidos por una sola persona; 2) Hogares nucleares: hogar conformado por un núcleo conyugal primario (pareja sin hijos, pareja con hijos o persona con hijos, exclusivamente); 3) Hogar extenso: conformado por un núcleo conyugal, más otros parientes no nucleares, exclusivamente; 4) Hogar compuesto: conformado por un núcleo conyugal o un hogar extenso más otros no parientes, y 5) Hogares sin núcleo: jefatura de hogar sin pareja ni hijos, aunque pueden existir otras relaciones de parentesco (Cepal, 1997).

Figura 12: Tipos de hogares para puntaje de corte 81 puntos, según escenarios A, E y total de hogares



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

7.3.1 Características de los hogares reclasificados como *break-off*

Además de analizar las características de los hogares de acuerdo con su nivel de colaboración, es necesario analizar también las características de aquellos hogares que cambian su clasificación a medida que aumenta el puntaje de corte. Así, a diferencia del apartado anterior, en que el objetivo era analizar qué caracterizaba a los hogares clasificados como entrevistados de aquellos reclasificados como *break-off* para un determinado punto de corte, el objetivo aquí es analizar a los hogares que, para un criterio más flexible, son clasificados como entrevistados, pero al aumentar el valor de corte son reclasificadas como *break-off*. Para realizar este ejercicio se decidió analizar a los hogares que son recodificados como rechazo, a medida que se aumenta en dos puntos el puntaje de corte del escenario E.

El cuadro 22 resume las características sociodemográficas de los hogares que se van recodificando como parte de la categoría de rechazo en el escenario E. En la primera columna se observa el total de hogares que son reclasificados como *break-off* cuando el puntaje de corte es de 71 puntos. Las siguientes columnas presentan las características de los hogares que se suman a este *break-off* inicial (71 puntos) cuando se aumenta de dos en dos el puntaje de corte. Dado lo anterior, desde la segunda columna el número total de hogares es menor y aumenta a medida que se incrementa el puntaje³⁶.

³⁶Por ejemplo, el total de hogares que se reclasificarían como *break-off* en los 72 puntos sería de 3.225 hogares, en el corte de 74 puntos sería de 3.240 hogares, y así sucesivamente. Sin embargo, el ejercicio aquí realizado solo analiza a los hogares que se incorporan al rechazo implícito una vez que se aumenta el puntaje de corte.

Cuadro 22: Caracterización de los hogares reclasificados como break-off, según escenario E

Características	71 pts.	72 pts.	74 pts.	76 pts.	78 pts.	80 pts.
Ingreso disponible promedio mensual de los hogares*	1.075.181,0	1.315.964,6	1.084.843,3	1.383.265,7	1.497.139,6	1.244.740,2
Gasto promedio mensual de los hogares*	684.425,8	1.464.175,0	926.903,5	1.142.991,9	1.378.699,2	1.122.931,1
Hombre JH (%)	57,4	46,2	69,6	80,9	61,1	60,0
Edad JH	50,6	45,9	49,5	43,6	57,4	50,9
Años de escolaridad JH	11,6	14,3	11,7	13,2	11,4	11,4
Hombre SP (%)	61,8	46,4	85,9	60,9	62,6	66,4
Edad SP	49,6	45,6	49,9	42,0	55,4	50,8
Años de escolaridad SP	12,1	15,0	11,4	14,9	12,3	11,9
Total de hogares	3.208,0	17,0	15,0	13,0	200,0	353,0

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

* Las estimaciones excluyen el arriendo imputado.

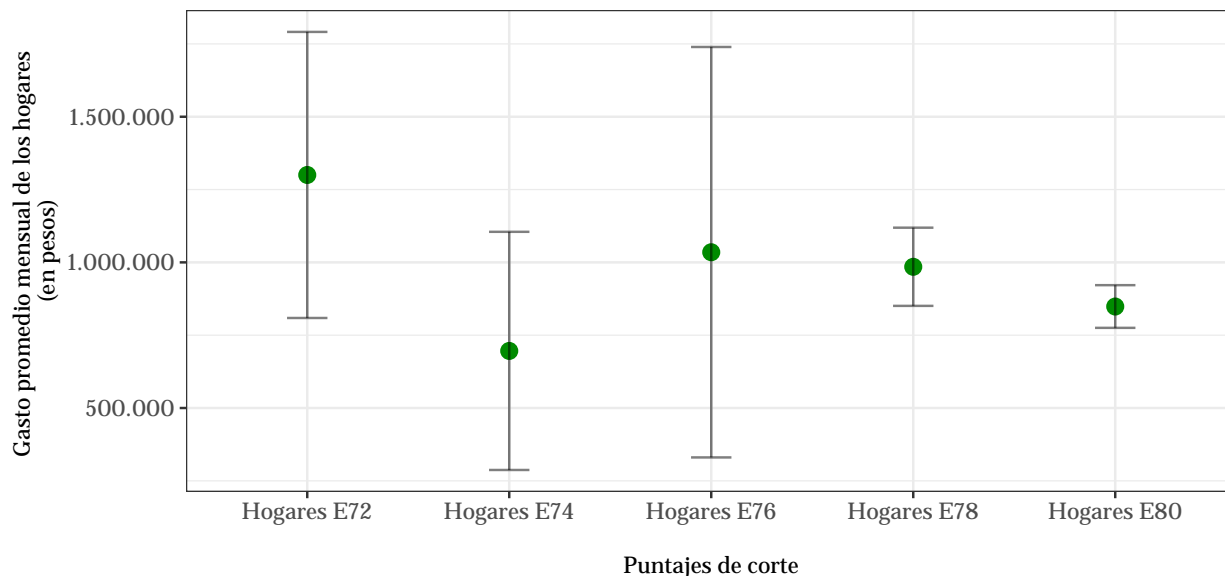
Se observa que, al incrementar el nivel de exigencia, el promedio del gasto e ingreso de los hogares siempre es mayor. Esta situación vuelve a dejar en evidencia que, al utilizar criterios más estrictos, los hogares que son reclasificados como *break-off* tienen en promedio un mayor nivel de información del gasto del hogar que aquellos incluidos previamente en esa categoría.

En relación con las características sociodemográficas, la proporción de jefaturas de hogar masculinas aumenta a medida que se incrementa el puntaje de corte. La mayor proporción de hombres como jefes de hogar se presentan en el *break-off* asociado a 76 puntos de corte, alcanzando un 80,9%. Con respecto a la edad del jefe de hogar, y en el mismo valor de corte de 76 puntos, se observa que el promedio de la jefatura está más cerca de los 50 años. En tanto, los años de escolaridad del jefe de hogar no registran una tendencia clara respecto a los años de educación media completa (12 años), siendo en promedio 13,2 años en el puntaje de corte de 76 puntos.

La figura 13 muestra el gasto promedio para los hogares reclasificados como *break-off* en los distintos puntajes de corte y sus respectivos intervalos de confianza, con el objetivo de identificar si existen diferencias significativas entre los distintos niveles de exigencias³⁷.

³⁷Para el cálculo del gasto promedio mensual de los hogares fue necesario calcular factores de expansión para cada escenario y punto de corte seleccionado. Sin embargo, no se utilizó el diseño complejo en dos etapas de la encuesta, por lo tanto, la varianza para cada escenario podría estar sobrestimada.

Figura 13: Gasto promedio mensual de los hogares reclasificados como *break-off* para escenario E, según puntaje de corte*



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).
* Las estimaciones excluyen el arriendo imputado.

Se observa que cuando se aumenta de 74 a 76 puntos, el gasto total de los hogares que se recodifican como *break-off* es de \$1.034.916 y presenta un intervalo de confianza mucho mayor que el resto de los puntajes de corte. Esta diferencia se produce por la baja cantidad de hogares que se incorporan a la categoría de *break-off* cuando se cambia de un punto de corte de 74 a 76 (en total 13 hogares).

7.4 Selección de un nivel de exigencia

El análisis comparativo realizado previamente permitió distinguir cómo se comportaban los indicadores principales de la encuesta en el escenario de control y el escenario E, dentro de un rango entre 71 y 81 puntos de corte.

En función de los resultados aquí presentados, la elección de un puntaje de corte para el escenario E implica tomar una decisión apoyada en criterios normativos más que relacionada con lo que se observa en los datos. Tal como se vio anteriormente, tanto la estructura de gasto como el promedio de este indicador para los hogares no presenta diferencias significativas cuando se compara los hogares entrevistados para los extremos del rango de puntajes elegidos.

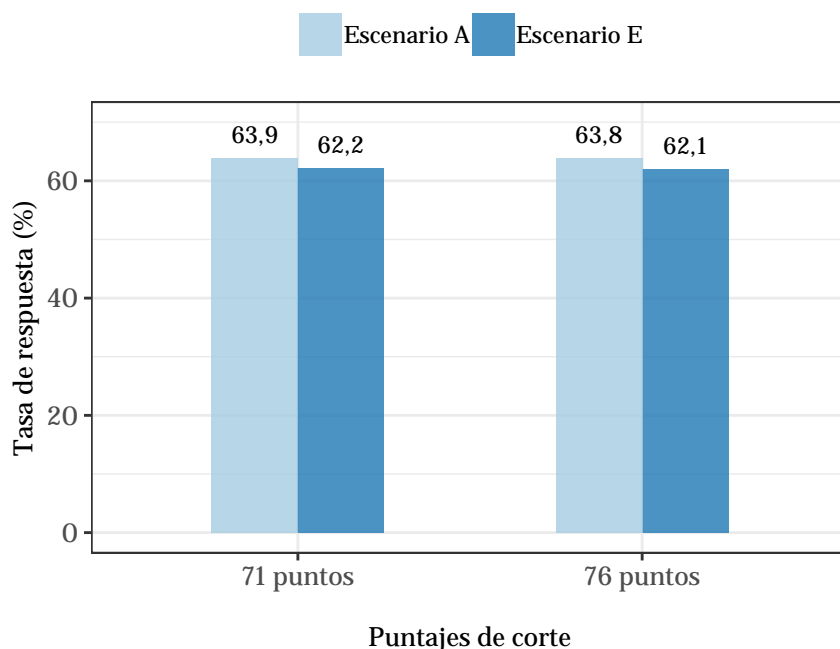
El puntaje a seleccionar debe estar en línea con el cumplimiento de la muestra objetivo planteada en el diseño muestral de la encuesta, permitiendo alcanzar una tasa de respuesta adecuada, sin sacrificar precisión en las estimaciones de gasto. Para la VIII EPF el puntaje escogido fue de 76 puntos. El principal argumento que sustentó esta elección corresponde a la modificación de la contribución de los cuestionarios al gasto, en donde la Libreta de Gastos del Hogar incrementa su participación en aquel y, por ende, se aproxima a la contribución realizada por la LGI. En este sentido, como la

penalización asociada a la falta de respuesta en la LGH es de 25 puntos³⁸, la no respuesta en este cuestionario gatillará la reclasificación del hogar como parte del *break-off*.

La elección del escenario E también se apoya en la implementación de nuevos criterios para evaluar la falta de respuesta en la LGI, a diferencia de la VII EPF, en que la penalización podía ser 0 o 30 puntos, sin permitir rangos intermedios. Al utilizar este criterio, la penalización por la falta de respuesta parcial depende en menor medida de la evaluación de un cuestionario en particular, por lo tanto, la reclasificación a *break-off* depende de la combinación de respuestas válidas y no solo si se respondió una libreta.

La comparación del escenario de control y el escenario E con un puntaje de corte de 76 puntos muestra que los principales indicadores de la encuesta mantienen la tendencia que se revisó anteriormente. El aumento del nivel de exigencia del método y un mayor puntaje de corte generan un crecimiento de los hogares que son reclasificados como *break-off* y, por ende, se registra una menor tasa de respuesta de la encuesta (figura 14). No obstante, el gasto promedio no presenta diferencias significativas cuando se modifica el método y su puntaje de corte (figura 15).

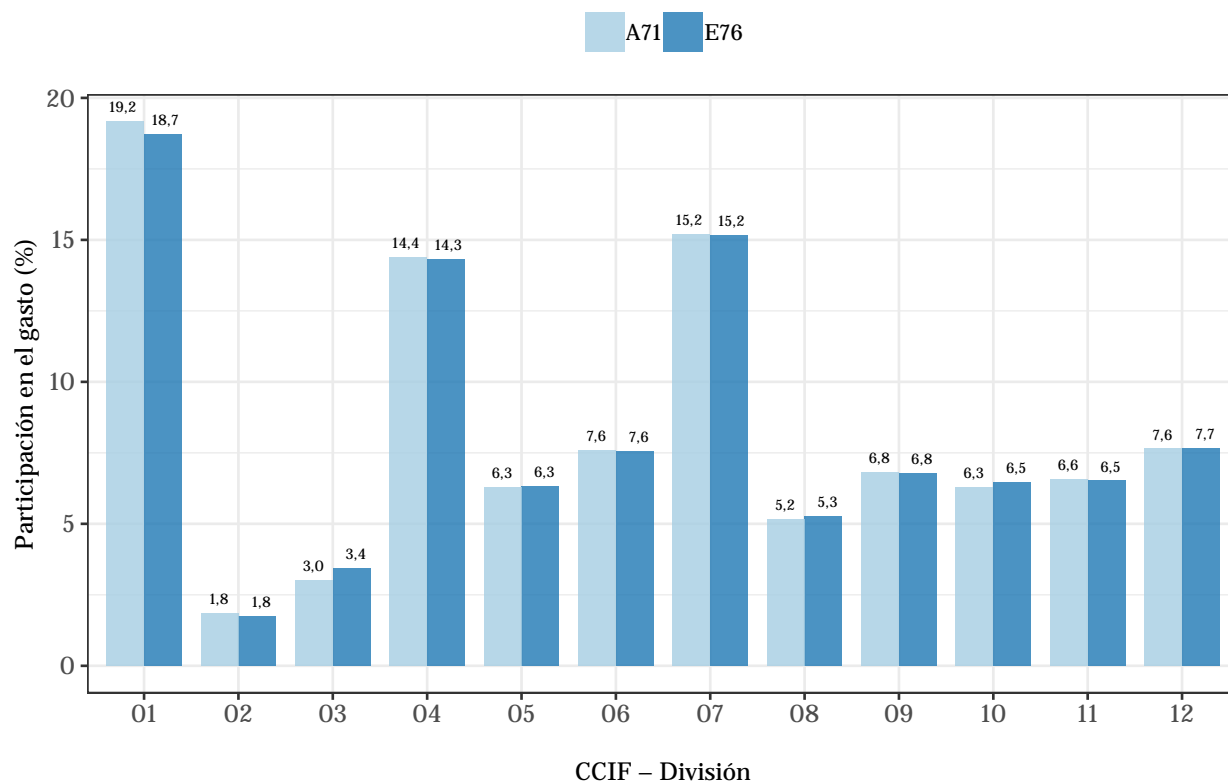
Figura 14: Tasa de respuesta en hogares para escenarios A y E, según puntajes de corte de 71 y 76 puntos



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

³⁸Los puntajes por libreta del escenario E corresponden a: RPH (15 puntos), LGI (29 puntos), LGH (25 puntos), LGR (16 puntos), LI (10 puntos) y LSF (5 puntos).

Figura 15: Estructura del gasto promedio mensual de los hogares* para escenarios A (71 puntos) y E (76 puntos), según división CCIF



Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF).

* Las estimaciones excluyen el arriendo imputado.

Cuadro 23: Caracterización de los hogares, según puntaje de corte 81 puntos

Características	A71 - entrevistada	E76 - entrevistada	A71 - break-off	E76 - break-off	Total hogares
Ingreso disponible promedio mensual de los hogares*	1.163.258,2	1.160.426,0	1.046.921,8	1.077.383,8	1.145.933,9
Gasto promedio mensual de los hogares*	1.089.339,1	1.092.745,6	640.667,2	690.370,0	1.022.525,1
Hombre JH (%)	58,5	58,5	57,2	57,6	58,3
Edad JH	51,6	51,4	49,5	50,5	51,3
Años de escolaridad JH	12,0	12,0	11,8	11,7	11,9
Hombre SP (%)	62,1	61,9	60,8	61,9	61,9
Edad SP	50,8	50,8	47,8	48,6	50,4
Años de escolaridad SP	12,4	12,4	12,3	12,2	12,4
Total de hogares	15.683,0	15.239,0	2.809,0	3.253,0	18.492,0

Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF).

* Las estimaciones excluyen el arriendo imputado.

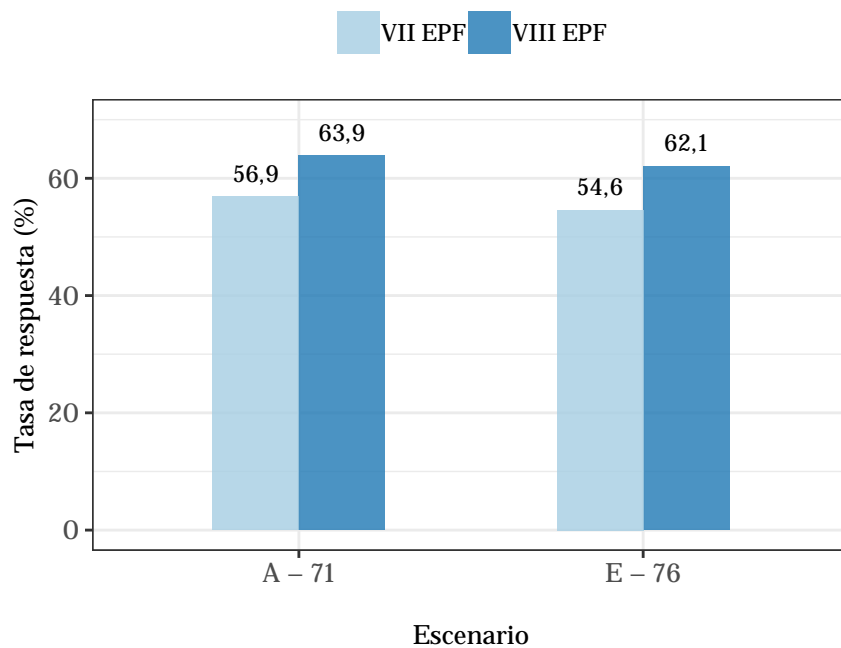
8 Efecto de la actualización en los resultados de la VII EPF

Para analizar cuáles habrían sido los efectos de este cambio en los resultados de la VII EPF -en términos de su tasa de respuesta y de las estimaciones de gasto- se presenta a continuación una comparación de los resultados ya publicados de esta versión y los resultados considerando el nivel de exigencia planteado por el escenario E para un valor de corte de 76 puntos. Para complementar esta comparación se presentan, además, los resultados de la VIII EPF tomando en cuenta el escenario control (criterio VII EPF) y el nivel de exigencia escogido.

8.1 Tasa de respuesta

En relación con la VII EPF, la tasa de respuesta de la VIII EPF muestra una mejora importante, tanto si se compara los resultados de ambas versiones utilizando el criterio control (A71) o el criterio escogido (E76). La figura 16 muestra los datos de la VII y VIII EPF a nivel nacional, en donde se observa que, para el criterio control, la VIII EPF presenta una tasa de respuesta de 63,9%, 7 pp. más que la tasa de la VII EPF (56,9%). Al realizar la comparación considerando el criterio de reclasificación escogido para la VIII EPF, se mantiene la tendencia en la que se aprecia una mejora de este indicador para su última versión, con un 62,1%, aumentando incluso su distancia con la VII EPF, con 54,6% (7,5 pp. de diferencia).

Figura 16: Tasa de respuesta a nivel hogares de la VII y VIII EPF para escenarios A y E, según puntajes de corte de 71 y 76 puntos



Fuente: Instituto Nacional de Estadísticas (INE) – VII y VIII Encuesta de Presupuestos Familiares (EPF).

El impacto de un cambio de criterios pareciera afectar más fuertemente a la tasa de respuesta de la VII EPF que a los resultados de la VIII. En el caso de la primera, el uso de los nuevos criterios hace caer este indicador en 2,3 pp. al aplicar el escenario E76 (más exigente que el escenario control), mientras que para la VIII EPF la diferencia es de 1,8 pp.

8.2 Indicadores de gasto promedio

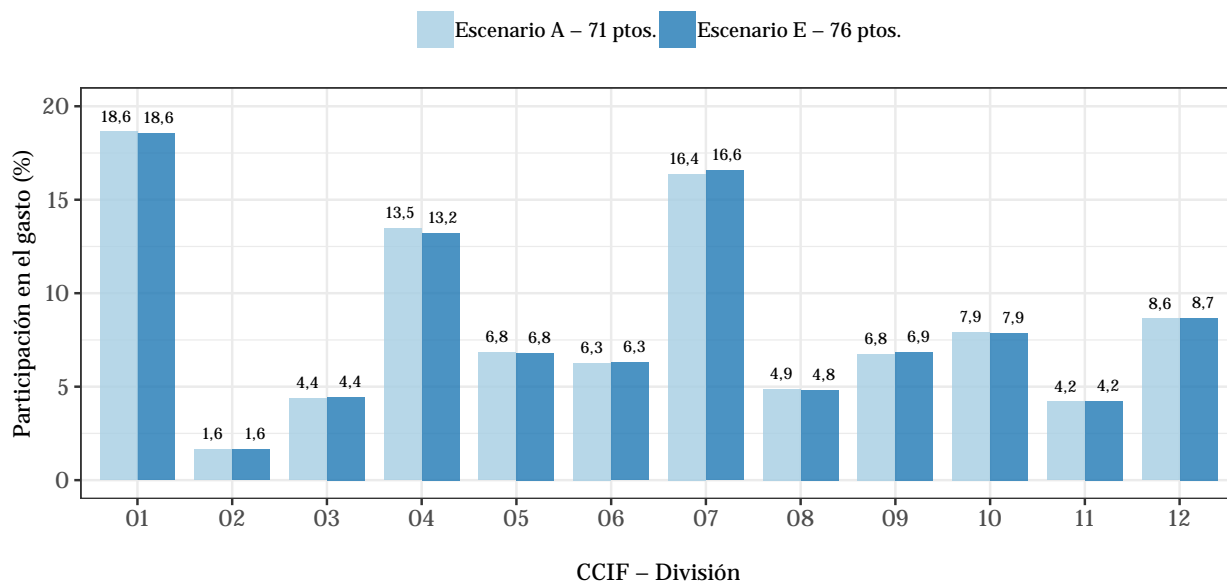
El análisis del efecto en los indicadores de gasto se divide, por un lado, en la revisión de la estructura de gasto de los hogares y, por otro, en el gasto promedio mensual para ambas versiones de la encuesta³⁹.

Cabe destacar que, para realizar este análisis, los resultados de la VII y VIII EPF fueron estimados considerando las metodologías para el cálculo de factores y para la imputación de gastos diarios empleados en la VII EPF. Para la VIII EPF se evaluaron 3 ajustes metodológicos, entre los que se incluye la actualización de los criterios de reclasificación tratados en este documento. Además, se efectuaron cambios a la metodología de cálculo de factores de expansión, así como también se evaluó la implementación del método *hot deck* para la imputación del registro de gastos diarios. Mientras que para todas las secciones anteriores del documento las estimaciones fueron hechas considerando estos nuevos ajustes, los datos que se muestran a continuación presentan un escenario hipotético en el cual el único cambio entre la VII y VIII EPF es la actualización de los criterios de la grilla técnica.

La figura 17 presenta la estructura de gasto de los hogares clasificados como entrevistados en la VII versión de la encuesta, tanto para el escenario control como para el escenario E76. Tal como se ilustra, la estructura de gastos es casi idéntica en ambos escenarios, no existiendo diferencias relevantes a nivel de divisiones. Esta tendencia está en línea con los datos presentados por la VIII EPF.

³⁹La actualización de los valores de la VII EPF se realizó con la variación del IPC de abril 2012 a diciembre 2016, que corresponde a 16,5412289%.

Figura 17: Estructura del gasto promedio mensual de la VII EPF para escenarios A (71 puntos) y E (76 puntos), según división CCIF

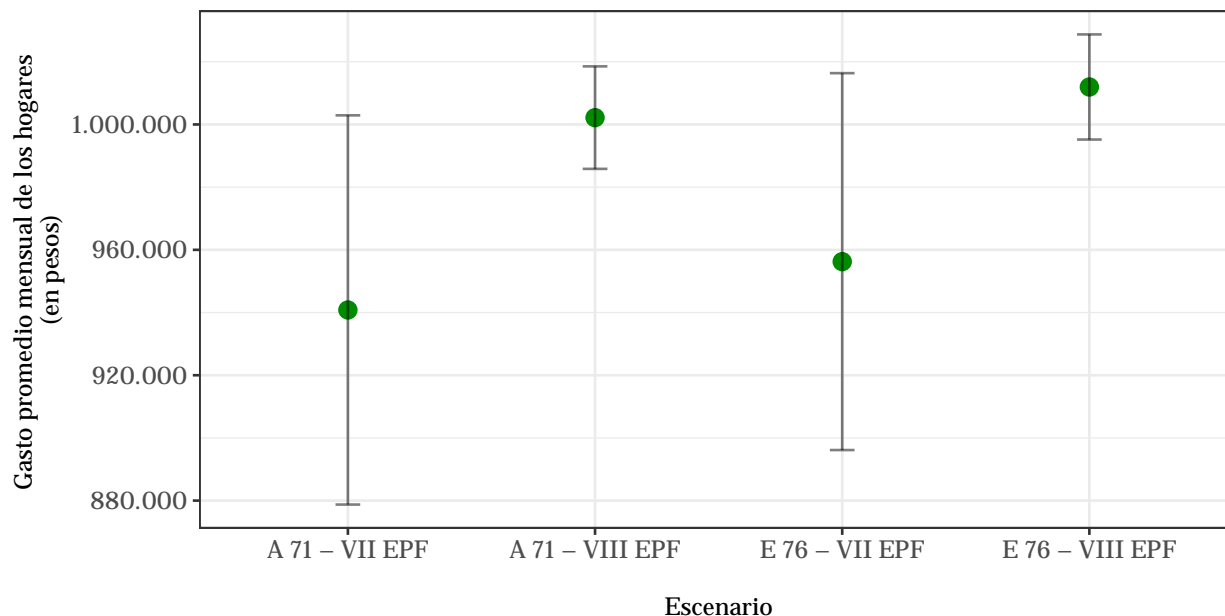


Fuente: Instituto Nacional de Estadísticas (INE) – VII Encuesta de Presupuestos Familiares (EPF).

En relación con el gasto promedio mensual de los hogares, la figura 18 muestra el gasto tanto de la VII como de la VIII EPF en los distintos escenarios. La primera registra un gasto promedio mensual⁴⁰ de \$940.823 cuando se aplica la grilla técnica del escenario de control, en tanto, si se utiliza el escenario propuesto (E76), el promedio aumenta a \$956.248. La diferencia en el gasto para los distintos niveles de exigencia es de \$15.425. No obstante, dicha estimación no presenta una diferencia significativa a nivel de gasto promedio por hogar, por lo que un aumento en la exigencia de la evaluación de los datos no hubiese afectado las estimaciones del gasto promedio en la VII EPF.

⁴⁰El dato oficial del gasto promedio mensual por hogar (ajustado por factor de no respuesta sin arriendo imputado) publicado en la VII EPF corresponde a \$807.406 pesos. Dicho monto fue ajustado por IPC desde abril de 2012 (mes central de la VII EPF) a diciembre de 2016 (mes central de la VIII EPF).

Figura 18: Gasto promedio mensual de los hogares de la VII y VIII EPF para escenarios A y E, según puntajes de corte de 71 y 76 puntos



Fuente: Instituto Nacional de Estadísticas (INE) – VII y VIII Encuesta de Presupuestos Familiares (EPF).
* Las estimaciones excluyen el arriendo imputado.

Algo muy similar se observa con los datos de la VIII EPF, en que el gasto promedio mensual de los hogares es de \$1.002.180 cuando se aplica el escenario de control. Dicho gasto aumenta en \$9.777 cuando el método de evaluación aplicado es el escenario E, con 76 puntos, alcanzando un gasto promedio mensual de \$1.011.957. Sin embargo, al igual que las estimaciones con los datos de la VII EPF, la VIII no presenta diferencias estadísticamente significativas, por lo tanto, se rechaza que un método más estricto de evaluación de los datos tenga un efecto sustantivo en el gasto promedio mensual de los hogares en la VIII EPF.

La comparación realizada entre los datos de la VII y la VIII EPF para ambos escenarios (A71 y E76) refleja que no hay diferencias sustantivas en los indicadores de gasto promedio mensual de los hogares y su estructura. Es decir, al aplicar el escenario E con puntaje de corte de 76 puntos en la VII EPF, se hubiesen obtenido resultados similares a los publicados.

9 Conclusiones

Aunque la distinción entre no respuesta parcial y no respuesta a la unidad es clara desde un punto de vista teórico, en la práctica pueden existir diversos criterios para establecer esta diferencia. Dependiendo de los objetivos de cada producto estadístico y de las características del estimador, en general será necesario definir una distinción acorde con las características de cada producto, que permita establecer una separación concreta entre ambos tipos de no respuesta.

El presente documento tiene por objetivo poner a disposición de los usuarios de la EPF los criterios utilizados en su VIII versión para realizar esta distinción. A través de un método de evaluación, denominado grilla técnica, se propuso una definición concreta para separar a aquellos hogares entrevistados con información parcial, de aquellos hogares que deben ser considerados al interior de la categoría de *break-off*. Este método permitió medir el nivel de no respuesta parcial de cada unidad, permitiendo excluir hogares en los que se tenía sospecha de un importante nivel de subdeclaración y reduciendo con esto posibles sesgos en las estimaciones asociados a la no respuesta.

La VII EPF fue la primera versión de la encuesta en que se utilizó un método que evaluó de manera sistemática el nivel de no respuesta parcial en cada uno de los hogares que participaron en el estudio. Producto de las modificaciones realizadas entre la VII y VIII versión, estos criterios debieron ser actualizados, para así responder a las nuevas características de la encuesta. Con esta actualización se buscó evaluar la posibilidad de utilizar criterios de evaluación más exigentes, así como también analizar el posible efecto en las estimaciones de gasto de la encuesta.

Los nuevos criterios evaluados consideraron:

- a) Cambios en la ponderación de los cuestionarios, para la penalización por la falta de información en estos, reasignando los puntajes en función de la contribución del gasto en el primer semestre de la VIII EPF.
- b) Cambios en la ponderación de las preguntas al interior de cada cuestionario, redistribuyendo los puntajes al interior de cada libreta en función de la contribución del gasto en el primer semestre de la VIII EPF.
- c) Modificaciones en el criterio de evaluación de la Libreta de Gastos Individuales. En la VII EPF esta evaluación no consideró el número de integrantes elegibles que debía responder esta libreta, sino que solo se evaluó la “mejor libreta” en términos del registro de gastos. En la VIII EPF se evaluó penalizar por la falta de respuesta de cada informante elegible que debía contestar esta libreta, ponderando cada individuo según quienes ejercían o no el rol de administradores de gasto.

Para el desarrollo de los análisis se definieron 4 escenarios alternativos y un escenario control, que aplicó los criterios utilizados en la VII EPF. Los análisis muestran que todos los escenarios propuestos son más estrictos que el escenario control. Sin embargo, el escenario que utilizó las nuevas ponderaciones por cuestionario y pregunta ajustadas a la participación en el gasto, y que además incluyó un criterio para evaluar la respuesta en la LGI, que prorrateó el peso de los administradores en la construcción del puntaje de acuerdo al número de elegibles, resultó ser el más exigente en los puntajes de corte más altos.

El método de evaluación se basa en el cálculo de puntajes de una escala de 1 a 100, en donde mayores puntajes representan un menor nivel de omisión de información. En la VII EPF el puntaje de corte que establecía la diferencia entre hogares entrevistados y *break-off* se realizó utilizando como referencia la libreta de mayor ponderación en la construcción del puntaje (LGI, con 30 puntos) y también del gasto. Para la VIII versión se eligió como referencia el aporte de la Libreta de Gastos del Hogar (25 puntos), debido al aumento de su participación en la construcción del gasto para esta versión de la encuesta.

En términos generales, al comparar los resultados de aplicar los criterios de evaluación de la VII EPF (escenario A71) y el nuevo criterio para la VIII versión de la encuesta (escenario E76), no se observan diferencias significativas en los indicadores de gasto. Tanto la diferencia en el monto del gasto promedio mensual, así como la estructura del gasto promedio del hogar según divisiones no son estadísticamente distintas de cero.

La selección del método de evaluación debe equilibrar, por una parte, el cumplimiento de la muestra objetivo planteada en el diseño muestral de la encuesta (reflejado indirectamente en la tasa de respuesta) y, por otra, asegurar un nivel de exigencia que permita obtener precisión en las estimaciones (*trade-off*). En relación con esto, cuando se analizan los resultados en términos de la tasa de respuesta, para ambas versiones de la encuesta, se observa que un aumento en el nivel de exigencia disminuye las tasas de respuesta. No obstante, esta disminución no pone en riesgo el cumplimiento de la muestra objetivo, por lo que es posible adoptar un nivel de exigencia más alto sin sacrificar los indicadores de respuesta de la encuesta.

10 Referencias

- AAPOR. (2016). Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys. American Association for Public Opinion Research. Recuperado de https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf
- Bethlehem, J., & Schouten, B. (2004). Nonresponse Adjustment in Household Surveys. Statistics Netherlands.
- Beullens, K., Loosveldt, G., Vandenplas, C., & Stoop, I. (2018). Response Rates in the European Social Survey: Increasing, Decreasing, or a Matter of Fieldwork Efforts? *Survey Methods: Insights from the Field*. Recuperado de <https://surveyinsights.org/?p=9673>
- Biemer, P., & Lyberg, L. (2003). *Introduction to Survey Quality*. New Jersey: John Wiley & Sons.
- Cepal. (1997). Políticas sociales, familia y trabajo en la América Latina de fin de siglo. *Cepal. División de Desarrollo Social*.
- Cobben, F. (2009). *Nonresponse in Sample Surveys: Methods for Analysis and Adjustment*. Statistics Netherlands. Recuperado de <https://www.cbs.nl/-/media/imported/documents/2009/36/2009-x11-cobben.pdf>
- Dillman, D., Eltinge, J., Groves, R., & Little, J. (2002). Survey Nonresponse in Design, Data Collection, and Analysis. En *Survey Nonresponse*. Canada: John Wiley & Sons.
- Groves, R., Fowler, F., Lepkowski, J., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*. New York: John Wiley & Sons.
- Hubener, E., Knappenberger, C., Lee, Y., & Tan, L. (2018). The Consumer Expenditure 2017 Data Quality Profile. U.S. Bureau of Labor Statistics. Recuperado de <https://www.bls.gov/cex/dqreport2017.pdf>
- INE. (2013). Metodología VII Encuesta de Presupuestos Familiares. Instituto Nacional de Estadísticas. Recuperado de <https://www.ine.cl/estadisticas/ingresos-y-gastos/epf>
- INE. (2014a). Medición y análisis de la falta de respuesta total en la VII EPF. Instituto Nacional de Estadísticas. Recuperado de <https://www.ine.cl/estadisticas/ingresos-y-gastos/epf>
- INE. (2014b). Métodos de imputación VII EPF: Gastos diarios e ingresos de la actividad laboral principal y jubilaciones. VII Encuesta de Presupuestos Familiares. Instituto Nacional de Estadísticas. Recuperado de <https://www.ine.cl/estadisticas/ingresos-y-gastos/epf>
- INE. (2016). Manual de trabajo de campo para encuestadores, VIII EPF, Tomo I. Instituto Nacional de Estadísticas.
- INE. (2018). Metodología VIII Encuesta de Presupuestos Familiares. Instituto Nacional de Estadísticas. Recuperado de <https://www.ine.cl/estadisticas/ingresos-y-gastos/epf>
- INE. (2019a). Medición y análisis de la falta de respuesta total en la VIII EPF. Instituto Nacional de Estadísticas. Recuperado de <https://www.ine.cl/inicio/documentos-de-trabajo/metodologicos>
- INE. (2019b). Métodos de Imputación VIII EPF: Gastos diarios e ingresos de la actividad laboral y

jubilaciones. Instituto Nacional de Estadísticas. Recuperado de <https://www.ine.cl/inicio/documentos-de-trabajo/metodologicos>

Ullman, H., Maldonado, C., & Nieves, M. (2014). La evolución de las estructuras familiares en América Latina, 1990-2010: Los retos de la pobreza, la vulnerabilidad y el cuidado. *División de Desarrollo Social, Cepal*.