

Documentos de trabajo

**Cálculo de medidas de precisión
para medianas de ingreso y
gasto de la VIII EPF**

Autores:

Klaus Lehmann Meléndez
Camila Mena Carvacho



INSTITUTO NACIONAL DE ESTADÍSTICAS

Morandé 801, Santiago de Chile

Teléfono: 562 3246 1000

Correo: ine@ine.cl

Facebook: @ChileINE

Twitter: @INE_Chile

www.ine.cl

Klaus Lehmann Meléndez

Camila Mena Carvacho

Subdepartamento de Estadísticas Socioeconómicas

Departamento de Estadísticas Demográficas y Sociales

Subdirección Técnica

Los autores agradecen la colaboración de Iván Touron Romero, jefe del Subdepartamento de Diseño de Marcos y Muestras, por su revisión metodológica y observaciones las que fueron esenciales para desarrollar el estudio. También agradecen los comentarios y aportes de Leonardo González Allendes, jefe de Departamento de Estadísticas Demográficas y Sociales, Agustín Arce Garcés coordinador del equipo técnico del Subdepartamento de Estadísticas Socioeconómicas y Juan Urbina Manzanares, profesional del Departamento de Estadísticas Demográficas y Sociales.

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.

Cálculo de medidas de precisión para medianas de ingreso y gasto de la VIII Encuesta de Presupuestos Familiares (EPF)

Resumen

El documento describe los procedimientos y resultados de distintas metodologías para la construcción de medidas de precisión para medianas de ingresos y gastos en el contexto de la VIII Encuesta de Presupuestos Familiares (EPF). En primer lugar, se expone la importancia de considerar la mediana en distribuciones de ingresos y gastos como medida de tendencia central. Luego, se presentan las metodologías probadas para la construcción de medidas de precisión para percentiles. Los resultados dan cuenta de que a nivel agregado no existe una diferencia importante entre las diferentes metodologías, pero al hacer estimaciones en dominios pequeños, los métodos comienzan a mostrar diferencias. El documento termina concluyendo que el método más adecuado para los datos de la EPF es un remuestreo tipo *subbootstrap*.

Palabras clave: *mediana, precisión, remuestreo, bootstrap, subbootstrap, ingresos, gastos.*

Abstract

The document describes the procedures and results of different methodologies for the construction of precision measures for medians of expenditures and income in the context of the VIII Household Budget Survey (EPF). In the first place, the importance of considering the median in distributions of expenses and income is exposed as measure of central tendency. Then, the proven methodologies for the construction of precision measurements for percentiles are presented. The results show that at the aggregate level there is no important difference between the different methodologies, but when making estimates in small domains, the methods begin to show differences. The document concludes that the most appropriate method for the data of the EPF is a resampling type *subbootstrap*.

Keywords: *median, precision, resampling, bootstrap, subbootstrap, income, expenditures.*

Índice general

1	Introducción	7
2	Motivación del análisis de la mediana	9
3	Métodos para construir medidas de precisión	13
3.1	Métodos basados en réplicas	13
3.2	Métodos analíticos	16
3.3	Resumen de los métodos	17
4	Resultados de las pruebas	18
4.1	Resultados a nivel agregado	18
4.1.1	Ingreso disponible mensual del hogar	18
4.1.2	Gasto mensual del hogar	20
4.2	Resultados desagregados	25
4.2.1	Relación entre la precisión y características de la muestra	29
4.3	<i>Bootstrap</i> versus <i>subbootstrap</i>	33
5	Conclusiones	38
	Referencias	41

Índice de cuadros

1	Muestreo de conglomerados mediante <i>bootstrap</i>	14
2	Muestreo de conglomerados mediante <i>subbootstrap</i>	15
3	Coefficientes de variación del Ingreso disponible mediano mensual del hogar, según área de estimación	19
4	Amplitud del intervalo de confianza del Ingreso disponible mediano mensual del hogar, según área de estimación	20
5	Coefficientes de variación del Gasto mediano mensual del hogar, según área de estimación	21
6	Amplitud del intervalo de confianza del Gasto mediano mensual del hogar, según área de estimación	22
7	Coefficientes de variación del Gasto mediano mensual del hogar, según división de la CCIF	22
8	Amplitud del intervalo de confianza del Gasto mediano mensual del hogar, según división de la CCIF	24
9	Grados de libertad y amplitud del intervalo de confianza para cada una de las categorías de CIUO (mujeres)	27
10	Grados de libertad y coeficiente de variación para cada una de las categorías de CIUO (mujeres)	27
11	Grados de libertad y amplitud del intervalo de confianza según situación en el empleo (mujeres)	29
12	Grados de libertad y coeficiente de variación según situación en el empleo (mujeres)	29

Índice de figuras

1	Función de densidad de <i>kernel</i> del Ingreso disponible y Gasto mensual del hogar . . .	9
2	Ingreso disponible mensual del hogar, según área de estimación VIII EPF	10
3	Gasto mensual del hogar, según área de estimación VIII EPF	11
4	Gasto mensual del hogar, según divisiones de la CCIF	12
5	Intervalos de confianza para la mediana del Ingreso disponible mensual del hogar a nivel Total de capitales regionales	20
6	Intervalos de confianza para la mediana del Gasto mensual del hogar a nivel Total de capitales regionales	21
7	Intervalos de confianza para la mediana del Gasto mensual del hogar, según división de la CCIF	23
8	Estimación de la mediana e intervalos de confianza, según CIUO para mujeres a nivel de Total de capitales regionales	26
9	Estimación de la mediana e intervalos de confianza, según situación en el empleo (CISE) para mujeres a nivel Total de capitales regionales	28
10	Amplitud del intervalo de confianza según el número de observaciones (Mujeres) . .	30
11	Coefficiente de variación según el número de observaciones de los seis métodos (CISE, CIUO, Edad, Educación)	31
12	Coefficiente de variación según grados de libertad	32
13	Diferencia porcentual respecto al intervalo de confianza del diseño complejo - Clasificación internacional uniforme de ocupaciones (CIUO-08)	34
14	Diferencia porcentual respecto al intervalo de confianza del diseño complejo - Clasificación internacional de la situación en el empleo (CISE)	34
15	Diferencia porcentual respecto al intervalo de confianza del diseño complejo - Nivel educativo	35
16	Intervalos de confianza del diseño complejo, <i>bootstrap</i> y <i>subbootstrap</i> - Clasificación internacional uniforme de ocupaciones (CIUO-08)	36
17	Diferencia porcentual respecto al intervalo de confianza del diseño complejo - Tramos de edad	36
18	Diferencia porcentual respecto al intervalo de confianza del diseño complejo	37

1 Introducción

El presente documento tiene por objetivo dar cuenta de las pruebas realizadas por el Subdepartamento de Estadísticas Socioeconómicas, destinadas a evaluar diferentes métodos para la construcción de medidas de precisión para las medianas de ingresos y gastos de la VIII Encuesta de Presupuestos Familiares (EPF). En ese sentido, todas las pruebas aquí contenidas tienen la finalidad de respaldar la decisión de qué método utilizar para construir medidas de precisión para estimaciones de medianas. Con ello se presenta la información necesaria para comprender la metodología de construcción de intervalos de confianza y coeficientes de variación utilizada por el Subdepartamento.

Tanto en la VII como en la VIII EPF la publicación oficial de tabulados estuvo constituida principalmente por medias de ingreso y gasto. La necesidad de publicar información de medianas se relaciona con el hecho de que, para distribuciones fuertemente inclinadas hacia uno de los dos extremos, tal como ocurre con el ingreso y el gasto, la mediana, como estadístico de tendencia central, entrega información de gran relevancia para comprender el fenómeno de estudio, lo cual contribuye con importante información complementaria para realizar análisis más acabados de los ingresos y gastos de los hogares.

Ahora bien, utilizar la mediana implica algunas complejidades, pues el cálculo de sus medidas de precisión no es tan sencillo como en el caso de la media. Dado que la distribución de la mediana es desconocida, es necesario recurrir a estrategias no paramétricas para el cálculo de intervalos de confianza y coeficientes de variación. En este documento se revisan dos enfoques para construir medidas de precisión de percentiles. Por un lado, se revisan las metodologías que implican procedimientos de remuestreo, y por otro, se abordan métodos que no utilizan réplica. El objetivo es comparar los resultados obtenidos en cada caso e identificar la capacidad que tienen los métodos de generar indicadores de precisión a niveles más desagregados. Esto último es fundamental, puesto que la metodología seleccionada requiere que sea posible de ser aplicada a dominios de estimación más pequeños.

En términos generales, se identificó que ambos enfoques arrojan resultados similares, sin embargo, los métodos de remuestreo tienden a generar una estimación de la varianza ligeramente más alta que los métodos no replicados, lo cual se traduce en intervalos de confianza más amplios y coeficientes de variación más elevados. Esto último inclinó la decisión hacia un enfoque de remuestreo, ya que, considerando que los métodos no generan resultados demasiado disímiles, se tomó la determinación de que es preferible considerar un enfoque conservador a la hora de realizar estimaciones sobre la varianza (coeficientes de variación más altos e intervalos más amplios), que eventuales subestimaciones de la misma (coeficientes de variación más pequeños e intervalos más estrechos). La razón es que el error tipo I¹ tiende a aumentar si se subestima la varianza, lo cual incrementa la probabilidad que el intervalo no contenga el parámetro de interés.

Este documento busca entregar mayores antecedentes respecto a las decisiones implicadas en la elaboración de las estadísticas oficiales, poniendo a disposición de los usuarios el detalle del método utilizado para construir las medidas de precisión para las medianas de ingresos y gastos considerando los datos de la VIII EPF. Se espera que la documentación entregada sea de utilidad, no solo para las

¹Este tipo de error se comete cuando el investigador rechaza la hipótesis nula (H_0) siendo esta verdadera.

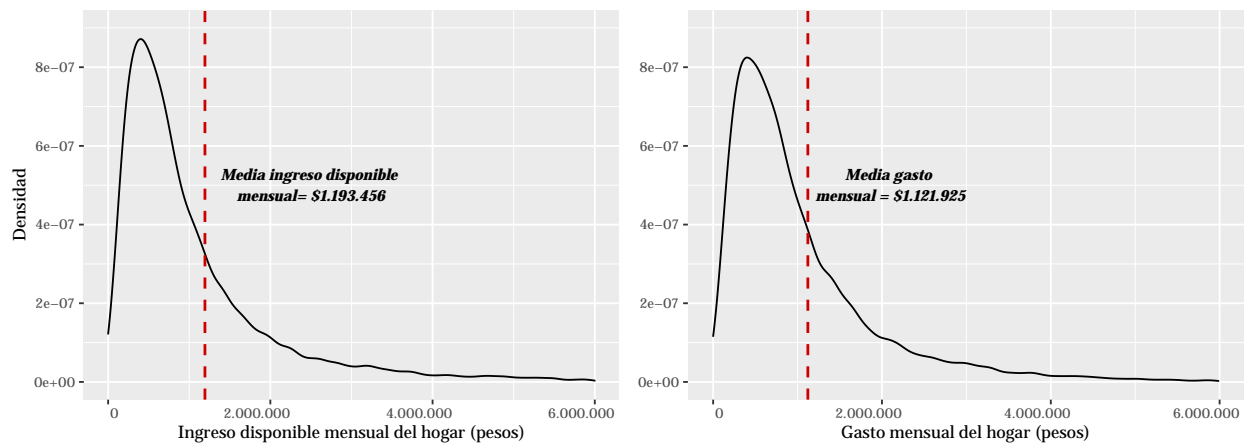
siguientes versiones de la encuesta, sino que también para otras encuestas de hogares y oficinas estadísticas que enfrenten desafíos similares. Asimismo, se espera que a partir de esta publicación surjan nuevas propuestas para calcular medidas de precisión de la mediana.

2 Motivación del análisis de la mediana

Tradicionalmente, los tabulados oficiales de la EPF se han publicado con la información de la media de ingresos y gastos, una medida de tendencia central que permite resumir en una cifra el comportamiento de las distribuciones analizadas y comparar distintas subpoblaciones. No obstante, la media se convierte en un indicador insuficiente cuando existe una alta dispersión de los datos. Una distribución altamente inclinada hacia una de las colas provoca que la media se vea afectada por los valores extremos y se aleje de los valores centrales de la distribución que muchas veces es la realidad que se quiere describir.

Los datos de la VIII EPF muestran que las distribuciones de ingresos y gastos de los hogares son asimétricas, pues están fuertemente inclinadas hacia la izquierda, la razón de esto es que existen algunos valores que, por ser altos, empujan la media hacia la derecha. En la figura 1 se ilustra cómo la media se aleja del centro de la distribución dada la influencia de altos valores. En este caso, utilizar únicamente la media sin consideraciones adicionales podría conducir a una evaluación que no refleje adecuadamente los ingresos y gastos de la mayor parte de la población².

Figura 1: Función de densidad de *kernel* del Ingreso disponible y Gasto mensual del hogar



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

Una forma de lidiar con la sensibilidad de la media frente a la dispersión de los datos es incorporar información adicional, en este sentido la mediana surge como un complemento de la media. En términos generales, la mediana busca representar la ubicación del individuo medio en la distribución, lo que significa una ventaja en relación con la media, debido a que su cálculo se ve afectado en menor medida por valores atípicos. Por lo tanto, en comparación con la media, esta entrega una cifra más cercana a lo observado en la ubicación del individuo medio (percentil 50) de la población cuando la distribución es muy dispersa.

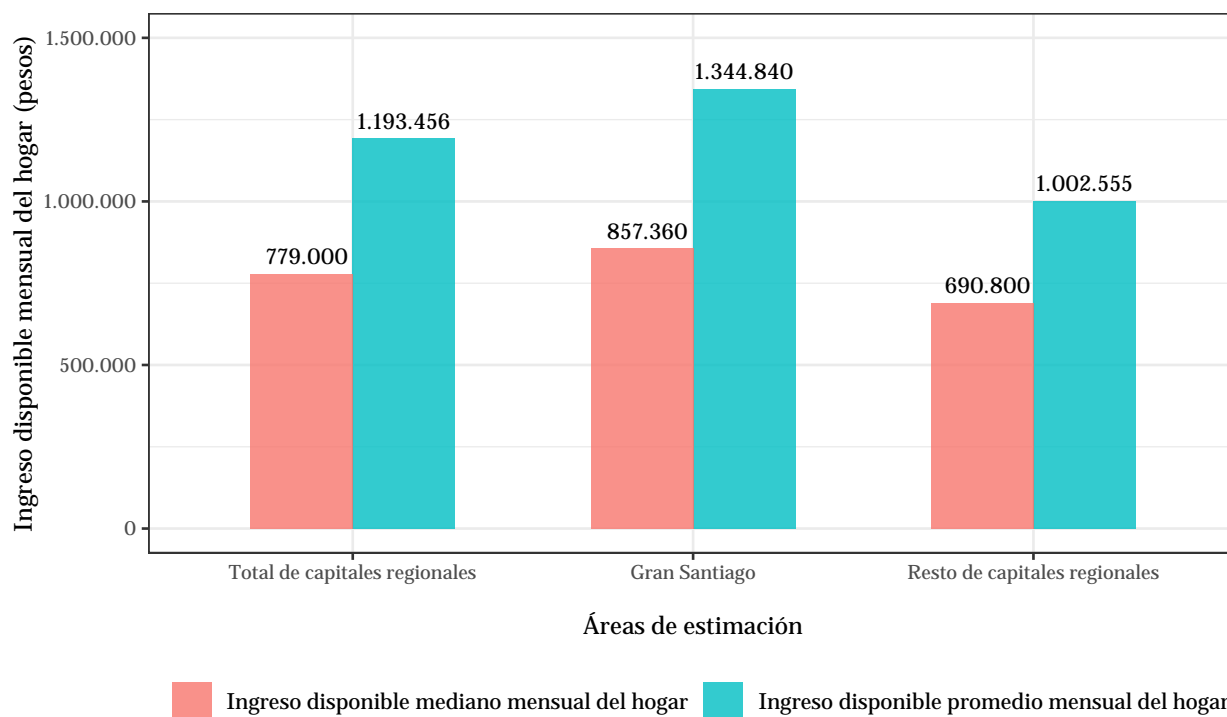
²En la figura 1 se truncaron los datos a 6 millones de pesos tanto para el ingreso disponible como para el gasto. Lo que genera que no se ilustre la información de 305 hogares para el ingreso disponible y 94 hogares en el caso del gasto del hogar. No obstante, el cálculo de ambas medias se realiza con el total de hogares. El objetivo del truncamiento de los datos es apreciar de forma más clara la asimetría de la distribución.

En ese sentido, la propuesta de entregar estimaciones de la mediana en los tabulados oficiales de la encuesta no tiene por objetivo reemplazar a la media, sino complementar la información sobre ingresos y gastos que ya se entrega con los resultados de la VIII EPF. La mediana permite contar con información adicional respecto de la dispersión de los datos y observar qué tan influenciada se encuentra la media por los valores extremos de la distribución.

En las figuras 2 y 3 se muestran comparaciones entre la media y la mediana del ingreso disponible y gasto mensual, según área de estimación, divididas en Total de capitales regionales, Gran Santiago y Resto de capitales regionales. En ambas figuras, se advierte que en todos los casos la media se aleja de manera importante de la mediana.

Así, por ejemplo, el promedio del ingreso disponible mensual a nivel del Total de capitales regionales es \$1.193.456, mientras que la mediana es \$779.000, es decir, existe una diferencia de más de \$400.000. En términos aún más concretos, al identificar en qué percentil de la distribución se ubica la media, se observa que dicho valor está en el percentil 69 de la distribución del ingreso disponible mensual del Total de capitales regionales, distanciándose de la tendencia central de la distribución. Este ejemplo ilustra como la información de la media y la mediana, en conjunto, contribuyen a un análisis de los resultados más completo.

Figura 2: Ingreso disponible mensual del hogar, según área de estimación VIII EPF

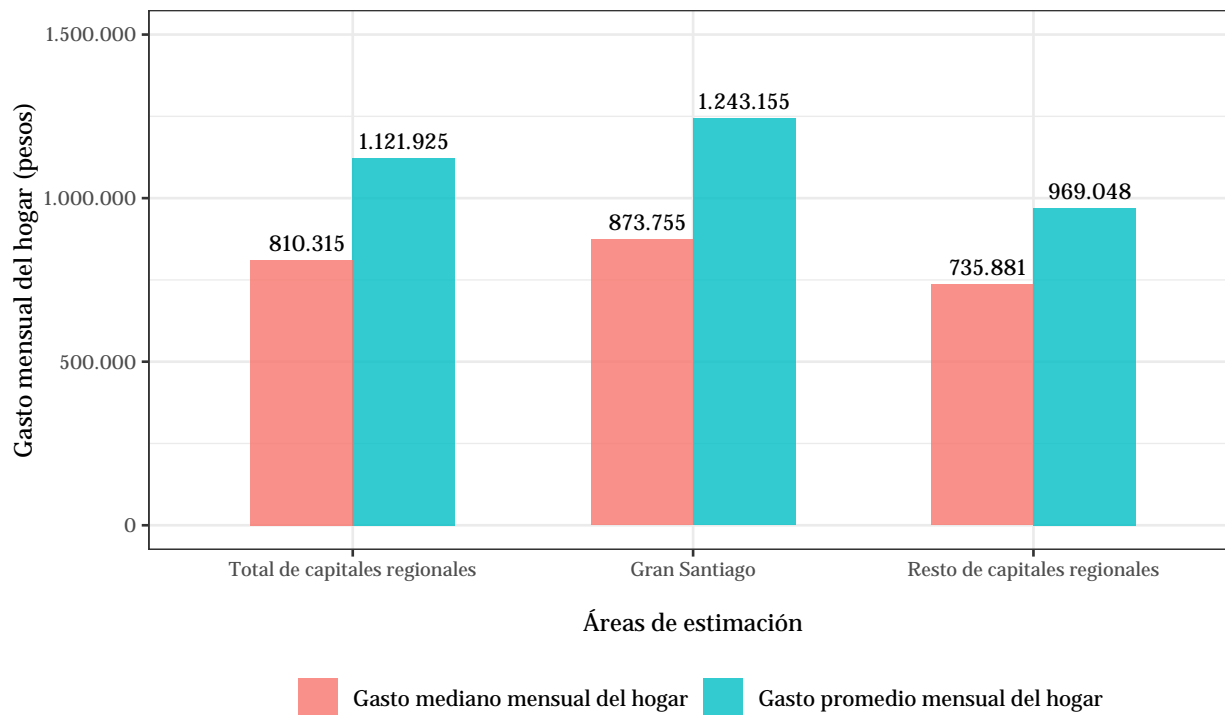


Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF)

En la figura 3 se presenta la información sobre el gasto mensual de los hogares. Nuevamente, se observa algo similar a lo ocurrido con el ingreso disponible, en donde por ejemplo a nivel de Total de capitales regionales, la media es marcadamente más alta que la mediana, con valores de \$1.121.925

y \$810.315, respectivamente, diferencia que se explica por la presencia de valores extremos en la distribución.

Figura 3: Gasto mensual del hogar, según área de estimación VIII EPF



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

Por último, si se compara la media y la mediana del gasto mensual del hogar según las divisiones de la Clasificación de Consumo Individual por Finalidades³ (CCIF), nuevamente se advierten diferencias importantes entre estas dos mediciones (figura 4). El caso más ilustrativo es el de la división de educación, donde la media es 174 veces mayor que la mediana, con valores de 72.596 y 417 pesos, respectivamente. Al contar con la información de la media y la mediana se deduce que una parte importante de los hogares no tiene gastos en educación o estos son muy cercanos a cero, mientras que por otro lado existe una fracción minoritaria de los hogares que realiza gastos muy elevados, lo cual afecta sustantivamente la media y, por ende, las conclusiones que de ella se puedan extraer, en lo que respecta al gasto en educación.

La información mostrada da cuenta de que las distribuciones de ingreso y gasto presentan una alta heterogeneidad, lo que provoca que la media se vuelva insuficiente como información resumen de todos los hogares. Tal como se ha señalado, la publicación de medianas busca ser un complemento

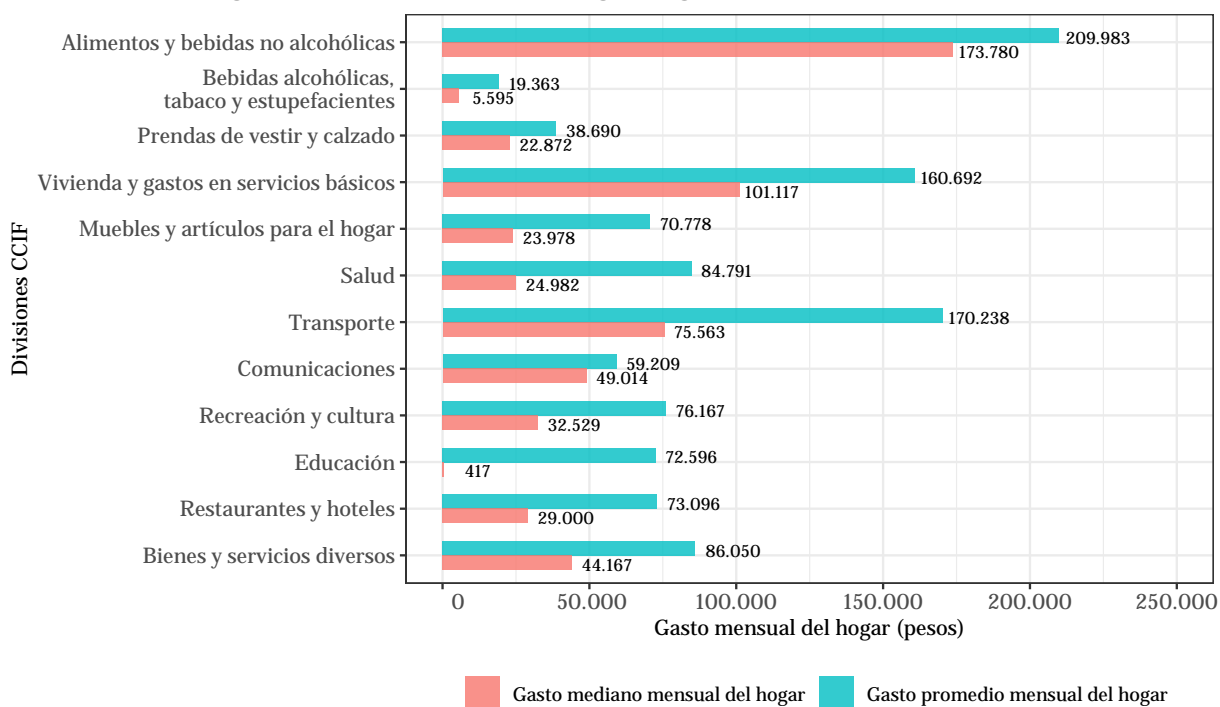
³La EPF utiliza la Clasificación del Consumo Individual por Finalidades (CCIF) para clasificar los gastos en consumo final de los hogares. La CCIF utilizada por la encuesta corresponde a una adaptación nacional de la versión de Naciones Unidas de 1999, la cual es una clasificación funcional que cubre 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 12 divisiones, correspondientes al gasto en consumo de los hogares, donde las restantes dos divisiones (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.

para las cifras ya entregadas de medias de ingresos y gastos de la VIII EPF. En este sentido, se ofrece información complementaria, para llevar a cabo análisis más completos.

El mayor nivel de agregación de la CCIF nacional lo presentan 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, 2018).

Existe una nueva versión de la clasificación publicada por Naciones Unidas en 2018 que contempla 13 divisiones de gasto en consumo de los hogares en vez de 12. Dicha versión del clasificador será adaptado para la IX EPF (2021-2022).

Figura 4: Gasto mensual del hogar, según divisiones de la CCIF



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

3 Métodos para construir medidas de precisión

La mediana aporta información relevante para describir las distribuciones de ingreso y gasto, tal como se revisó en el apartado anterior. No obstante, contar solo con su estimación puntual no es suficiente. Adicionalmente, se requiere calcular alguna medida de precisión, debido a que los datos de la EPF provienen de una muestra⁴. Por ejemplo, podría ser de interés conocer el intervalo de confianza, el error relativo o el coeficiente de variación. Dichas medidas de precisión, cuyo cálculo es relativamente sencillo y conocido para el caso de la media, no lo es tanto para la mediana. La razón es que la distribución de esta no es conocida (como sí ocurre con la media), por lo tanto, calcular su varianza y generar medidas de precisión requiere la utilización de métodos no paramétricos.

En términos generales, es posible distinguir dos grandes enfoques para calcular medidas de precisión de mediana. La primera de las estrategias se basa en la generación de réplicas. Con el avance de la computación, se ha hecho cada vez más sencillo la utilización de estrategias de remuestreo. De hecho, el poder computacional actual hace posible la generación de miles de muestras en un tiempo aceptable, con equipos que no exceden la capacidad de un computador doméstico. La idea en este enfoque es generar múltiples muestras, a partir de la muestra efectiva que fue seleccionada inicialmente. A partir de estas muestras se puede calcular una gran cantidad de veces el estadístico de interés, lo cual permite observar la imprecisión generada por la selección aleatoria de una sola muestra.

La segunda estrategia para construir medidas de precisión para percentiles se basa en métodos “analíticos”, que comenzaron a utilizarse en la década de 1950 a partir del trabajo inicial de Woodruff (1952), quien propuso por primera vez un algoritmo para construir intervalos de confianza para percentiles sin realizar supuestos respecto a su distribución.

Dentro del contexto de muestras complejas es posible utilizar ambas estrategias, por lo que en el presente documento se abordan tanto los métodos que utilizan remuestreo como los métodos analíticos para probar la construcción de medidas de precisión para la mediana.

3.1 Métodos basados en réplicas

Los métodos de remuestreo parten de la siguiente pregunta: ¿qué hubiese ocurrido si en lugar de contar con la muestra efectiva se hubiese seleccionado otra? Esta pregunta se puede abordar a través de la creación de nuevas muestras, a partir de la originalmente seleccionada. El propósito de este enfoque es calcular el estadístico de interés muchas veces, para así registrar los posibles valores que se obtienen a partir de muestras alternativas.

Cabe mencionar que dentro de este enfoque existe un abanico de posibilidades para llevar a cabo el remuestreo. Dentro de los algoritmos más utilizados están *bootstrap*, *jackknife*, *subbootstrap* y *BRR* (*Balanced repeated replication*). En el presente documento se prueban dos de estos mecanismos de remuestreo: *bootstrap* y *subbootstrap*. En términos generales, este tipo de remuestreo permite

⁴Es importante tener en consideración que la mediana, al igual que cualquier otro estadístico de tendencia central calculado con datos de una muestra, corresponde a un estimador de un valor poblacional, por ende, está sujeto a un error muestral.

realizar inferencias a partir de datos muestrales, el que a través de distintas iteraciones va seleccionando una nueva muestra, distinta a la efectiva, y permite calcular una gran cantidad de veces el estadístico de interés (Burn, 2006; Carpenter, 1999; Efron, 1979; Lumley, 2010).

Es importante señalar que en el contexto de encuestas de hogares no es común utilizar un muestreo aleatorio simple. Usualmente, se recurre a diseños complejos, en los cuales por lo general no todas las unidades tienen la misma probabilidad de ser seleccionadas. Considerando esta situación, el remuestreo debiese respetar las características del diseño construido inicialmente. En el caso específico de la VIII EPF se utiliza un muestreo estratificado en dos etapas. Esto implica que se seleccionan manzanas al interior de estratos, para luego seleccionar viviendas dentro de cada manzana⁵.

Con el propósito de no pasar por alto el diseño complejo, el remuestreo selecciona unidades de primera etapa (UPM) al interior de cada estrato. Dicho de otro modo, la unidad de selección del remuestreo no son las viviendas, sino los conglomerados. Así, un remuestreo realizado mediante *bootstrap* para el estrato h podría generar un resultado como el del cuadro 1. Se observa que en el primer remuestreo se seleccionan todos los conglomerados del estrato h . En la segunda iteración, dado que el muestreo es con reemplazo, se selecciona 2 veces al primero, 2 veces al segundo y ninguna vez a los restantes. En el tercero se selecciona una vez al primero y 3 veces al cuarto, y así hasta llegar a la muestra s . Se advierte que el tamaño de muestra en el caso de *bootstrap* es igual a n y debido a que el total de conglomerados en la VIII EPF es de 304, cada uno de los remuestreos generados será de tamaño 304.

Cuadro 1: Muestreo de conglomerados mediante *bootstrap*

		Muestra 1	Muestra 2	Muestra 3		Muestra s
h	Conglomerado 1	1	2	1	...	1
	Conglomerado 2	1	2	0		2
	Conglomerado 3	1	0	0		1
	Conglomerado 4	1	0	3		0

Una vez que se hace la selección de UPM, es necesario reponderar los factores de expansión del diseño, ya que se debe distribuir el peso de los conglomerados no seleccionados dentro de los que sí fueron seleccionados. Esto equivale a multiplicar el factor de expansión de cada unidad muestral del diseño (vivienda)⁶ por la cantidad de veces que fue seleccionado el conglomerado que la contiene. Una vez hecha esta operación, es posible calcular las medidas de precisión para la mediana (o cualquier otro percentil) utilizando los datos del remuestreo.

Es importante mencionar que, en el marco de este documento, los cálculos realizados con un muestreo tipo *bootstrap* se utilizan para construir dos tipos de intervalos de confianza. En el primer caso, siguiendo las recomendaciones de Lumley (2010), se calcula la desviación estándar respecto a la mediana empírica de cada remuestreo, este cálculo valida que la varianza no se vea afectada por remuestreos atípicos.

⁵Para más detalles respecto al diseño muestral de la VIII EPF, ver el documento Metodología VIII EPF, disponible en la página web del INE (www.ine.cl).

⁶En el diseño bietápico de la encuesta, una vez que se seleccionan las manzanas, son seleccionadas las viviendas particulares ocupadas al interior de estas. Dichas unidades corresponden a las unidades de segunda etapa (USM).

Es decir, según Lumley (2010, *pág. 240*) la estimación de la varianza se calcula como⁷:

$$\hat{var}[\hat{T}] = \text{scale} * \sum_{i=1}^m (T_i^* - \bar{T}^*)^2 * \text{rscales}_i \quad (1)$$

Donde m corresponden a la cantidad de réplicas, T_i^* su estimación y \bar{T}^* su mediana.

Una vez calculado dicho valor⁸, el intervalo de confianza se construye restando y sumando a la mediana (observada en la muestra) el producto entre un valor crítico y el error estándar, de la siguiente manera:

$$\text{lim} = \text{mediana} \pm Z_\alpha * se \quad (2)$$

Donde se es el error estándar de la distribución de medianas replicadas, Z_α es el valor crítico y α es el nivel de significancia.

La segunda forma de construir el intervalo de confianza mediante un remuestreo tipo *bootstrap* sigue la propuesta de Burn (2006). El autor propone construir intervalos de confianza a partir de los valores ubicados en los percentiles 0,025 y 0,975 de la distribución de medianas obtenidas a partir del remuestreo. Esto quiere decir que no necesariamente se obtendrán intervalos simétricos.

En el caso de un remuestreo *subbootstrap*, el tamaño de cada muestra es de $n - 1$. Esto implica que en cada estrato es necesario distribuir el peso de un conglomerado entre los que sí fueron seleccionados, multiplicando el factor de expansión de cada unidad muestral por $n/(n - 1)$. A modo de ejemplo, podría darse una situación en la que son seleccionados los primeros dos conglomerados de un estrato compuesto por tres, lo cual implica que es necesario distribuir el peso del conglomerado restante entre los 2 primeros. El resultado se traduce en que el peso de cada uno sea el equivalente a 1,5 veces el peso inicial, tal como se muestra en el cuadro 2⁹. Una vez calculados los pesos de cada conglomerado, se lleva a cabo la misma operación que en el caso de *bootstrap*, es decir, los pesos se multiplican por el factor de expansión del diseño.

Cuadro 2: Muestreo de conglomerados mediante *subbootstrap*

		Muestra 1
h	Conglomerado 1	1,5
	Conglomerado 2	1,5
	Conglomerado 3	-

Finalmente, cabe señalar que durante el proceso de pruebas de las distintas metodologías para calcular las medidas de precisión de la mediana se desestimó el uso de *BRR* y *Jackknife*, como

⁷Para más detalles sobre esta fórmula, revisar la literatura asociada y el código de fuente de la función *svrepdesign*.

⁸El cálculo del error estándar implementado en el paquete *Survey* de R implica un reescalamiento de los valores para cada una de las iteraciones. Para más detalles sobre este procedimiento, ver código de fuente de la función *svyquantile*.

⁹Si no se distribuyera el peso del conglomerado no seleccionado, la expansión de los datos no sería completa, lo cual haría imposible llegar a valores cercanos a los stocks poblacionales.

métodos de remuestreo. En el primer caso, el motivo corresponde a que el algoritmo requiere que en cada estrato el número de conglomerados sea par, lo cual no se cumple en el caso de la VIII EPF. Esto quiere decir que para poder utilizar un remuestreo tipo *BRR*, se requeriría fusionar estratos, de manera tal de no dejar ninguno con números impares. Dado que esto implicaba un cambio importante en el diseño muestral, se decidió no utilizar esta estrategia. En el caso de *Jackknife*, Lumley (2010) señala que es un método válido para calcular el error estándar de la media, sin embargo, no se recomienda para percentiles. Pese a esta recomendación, se implementaron algunas pruebas. Los resultados mostraron que mientras *bootstrap* y *subbootstrap* generan intervalos de confianza similares, en el caso de *Jackknife* se obtienen resultados bastante diferentes a los otros remuestreo, por lo que se desestimó su uso tempranamente.

3.2 Métodos analíticos

Los métodos analíticos intentan estimar la varianza de un percentil mediante estrategias que no requieren réplicas, pues se basan únicamente en los datos de la muestra efectivamente seleccionada. En el marco de este trabajo, se implementan cálculos a partir de las propuestas de Woodruff (1952), de Korn & Graubard (1998) y de Binder (1991).

En términos generales, el método de Woodruff consiste en encontrar los percentiles de la distribución entre los cuales se mueve la mediana (o cualquier otro percentil). Para ello, en primer lugar, se calcula el error estándar de la proporción de datos que está bajo la mediana, utilizando el diseño complejo:

$$se = \sqrt{\frac{1}{n-1} * \sum_{i=1}^n (x_i - x_{\frac{n}{2}})^2} \quad (3)$$

En segundo lugar, se utiliza dicho error estándar para construir el límite inferior y superior, mediante el siguiente procedimiento:

$$\begin{cases} inferior = 0.5 - (Z_{\frac{\alpha}{2}}) * se \\ superior = 0.5 + (Z_{1-\frac{\alpha}{2}}) * se \end{cases} \quad (4)$$

Donde se es el error estándar de la proporción de registros bajo la mediana, $Z_{\frac{\alpha}{2}}$ es el valor crítico asociado a un nivel de significancia α .

Los valores asociados a *inferior* y *superior* de la ecuación indican los percentiles entre los que se mueve la mediana. El paso final consiste en buscar los valores que están en los percentiles obtenidos mediante la operación anterior.

El método de Korn y Graubard funciona igual que el anterior, pero el intervalo de confianza para la proporción de valores bajo la mediana es calculado usando una función de distribución acumulada binomial.

Finalmente, el método de Binder propone un cálculo de los intervalos de confianza basado en un test invertido. Así, para un estimador $\hat{U}(\theta)$ se plantea el siguiente test:

$$\begin{cases} H_0 : U(\theta) = 0 \\ H_1 : U(\theta) = 1 \end{cases} \quad (5)$$

El intervalo de confianza se define como todos los valores de θ para los cuales H_0 no es rechazada. La región de no rechazo para H_0 se define como:

$$\frac{\hat{U}(\theta)^2}{mse(\hat{U}(\theta))} \leq Z_{1-\frac{\alpha}{2}}^2 \quad (6)$$

Donde $Z_{1-\frac{\alpha}{2}}$ representa un percentil de una distribución normal estándar. Lo anterior quiere decir que se asume que $\hat{U}(\theta)$ se aproxima a una normal, con media $U(\theta)$ y varianza $MSE(\hat{U}(\theta))$ y que $mse(\hat{U}(\theta))$ es un estimador consistente de $MSE(\hat{U}(\theta))$.

Según Binder (1991, *pág. 36*), para construir el intervalo de confianza de un parámetro de interés, éste último se define como la solución de $U(\theta_0) = 0$ y el estimador de $U(\theta)$ es $\hat{U}(\theta)$, con un error cuadrático medio como $mse(\hat{U}(\theta))$. Por lo tanto, vale la pena mencionar que si en la ecuación 6 se evalúa $mse(\hat{U}(\theta))$ asumiendo que $\theta = \hat{\theta}$, se obtienen los intervalos de confianza propuestos por Woodruff.

3.3 Resumen de los métodos

Considerando lo anterior, los métodos testeados en el presente trabajo son los siguientes:

1. Métodos basados en réplicas:
 - *Bootstrap*: Calcula un intervalo de confianza simétrico en base al error estándar de la distribución de medianas.
 - *Bootstrap* (Burn): Calcula un intervalo de confianza asimétrico en base al percentil 0,025 y 97,5 de la distribución de medianas.
 - *Subbootstrap*: Calcula un intervalo de confianza simétrico en base al error estándar de la distribución de medianas. Genera muestras de tamaño $n - 1$.
2. Métodos analíticos
 - Woodruff: Calcula el error estándar de los registros bajo la mediana.
 - Korn y Graubard: Similar a Woodruff, pero utiliza una función de distribución acumulada binomial.
 - Binder: Utiliza test invertido.

4 Resultados de las pruebas

En el apartado anterior se revisó el marco conceptual de los distintos métodos utilizados para calcular medidas de precisión para la mediana. Respecto a esto es importante señalar que existe una serie de indicadores que permiten medir la precisión de un estimador. En el marco de este trabajo se evalúan únicamente el coeficiente de variación y el intervalo de confianza.

A continuación, se presentan los resultados de los ejercicios realizados con los datos de la VIII EPF para evaluar las ventajas y desventajas de los distintos métodos¹⁰. El objetivo de esta evaluación es optar por uno de estos para publicar la estimación de la mediana en un tabulado complementario a los ya existentes.

Los resultados se presentan de la siguiente forma: en este apartado, se abordan las principales estimaciones de ingreso disponible y gasto de la VIII EPF, según áreas de estimación (nivel agregado), es decir, se realizan pruebas con los 6 métodos definidos previamente según Total de capitales regionales, Gran Santiago y Resto de capitales regionales. Por otro lado, en el siguiente apartado, se realizará el mismo análisis, pero con una desagregación mayor de los datos según características de los informantes de la encuesta. La idea es evaluar cómo se comportan los intervalos de confianza y los coeficientes de variación, cuando se realizan estimaciones con distintos niveles de desagregación.

4.1 Resultados a nivel agregado

4.1.1 Ingreso disponible mensual del hogar

En términos generales, los cálculos realizados tienen la característica de que la mediana es la misma en todos los casos, no obstante, los intervalos varían según el método utilizado. La principal diferencia reside en que los métodos analíticos registran intervalos que son asimétricos¹¹, mientras que los métodos basados en réplicas presentan intervalos simétricos. En consideración de lo anterior es que para simplificar la comparación de los métodos se utilizará la **amplitud** del intervalo de confianza, construida como la diferencia entre el límite superior y el inferior. Debido a la ya mencionada asimetría, no es tan sencillo utilizar una medida más común como el error relativo, pues la interpretación se torna un poco menos intuitiva¹².

Es común asumir que un intervalo más pequeño es preferible a uno amplio. Siguiendo este razonamiento, la decisión debiese inclinarse hacia el método que genere el intervalo más estrecho. Sin embargo, existe un *trade-off* en esta elección. Efectivamente un intervalo estrecho está asociado una estimación más precisa, pero también aumenta la probabilidad de error (en particular de cometer el error tipo I). Siguiendo un criterio conservador, podría argumentarse que un intervalo

¹⁰Todos los cálculos fueron realizados con la versión 3.5.1 de *R*. Versiones más actuales generan leves diferencias en el cálculo de los intervalos de confianza y el coeficiente de variación.

¹¹La asimetría no es tan evidente en los datos más agregados de la encuesta. En el siguiente apartado, cuando se revisen cifras más desagregadas, la asimetría se hace más patente.

¹²El hecho de que el intervalo de confianza sea asimétrico genera un error relativo con una interpretación poco usual, ya que hacia un lado es posible obtener un error, por ejemplo, de 5% hacia un lado y de 7% hacia el otro.

amplio es preferible a uno estrecho, pues una varianza mayor implica una menor probabilidad de rechazar una hipótesis cuando esta es verdadera. Ahora bien, un intervalo excesivamente amplio tampoco es deseable, ya que, si bien en este caso la probabilidad de error es baja, la información entregada por la estimación podría ser tan imprecisa, que la vuelve poco útil. En consecuencia, la elección del método no es algo trivial y requiere el análisis en conjunto de más de un factor.

Una decisión razonable sería la siguiente: seguir un criterio conservador en el sentido de entregar una amplitud que se pueda asumir como la máxima según el cálculo de la varianza.

Los resultados del ingreso disponible dan cuenta de coeficientes de variación e intervalos de confianza bastante similares entre los métodos. El cuadro 3 resume los coeficientes de variación para el ingreso disponible mensual del hogar según área de estimación.

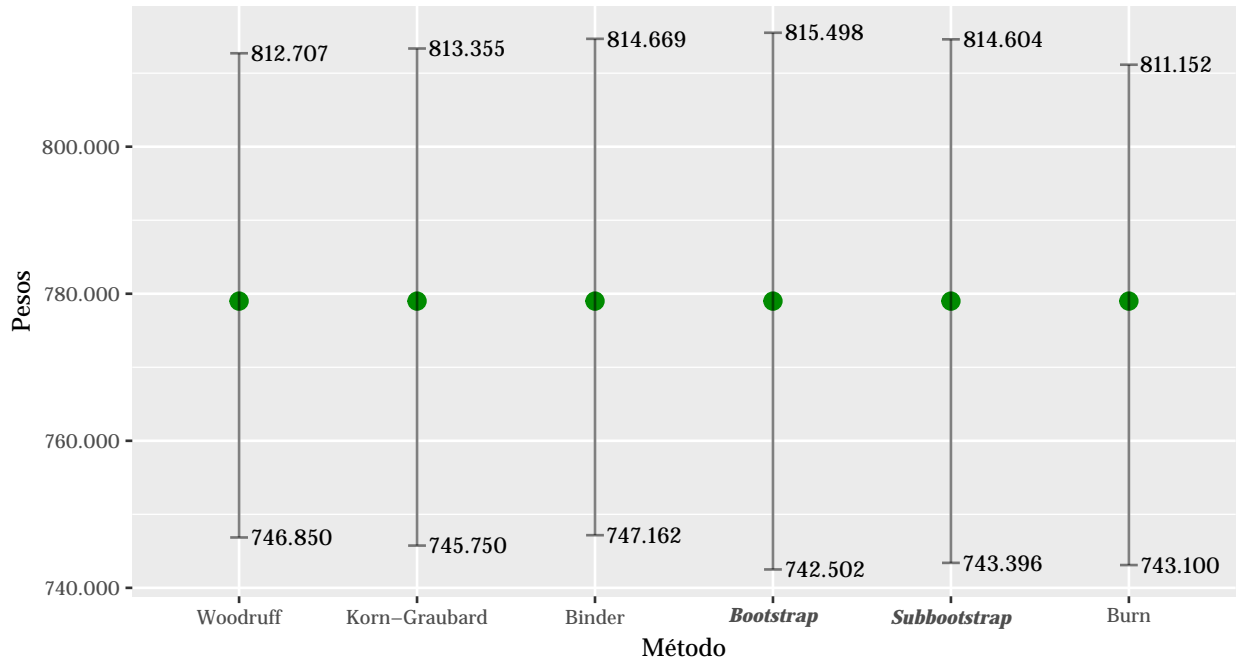
Cuadro 3: Coeficientes de variación del Ingreso disponible mediano mensual del hogar, según área de estimación

Estimación	Mediana	Woodruff	Korn-Graubard	Binder	<i>Bootstrap</i>	<i>Subbootstrap</i>	Burn
Total de capitales regionales	779.000	2,16	2,21	2,21	2,39	2,33	2,30
Gran Santiago	857.360	3,48	3,59	3,69	3,67	3,67	2,97
Resto de capitales regionales	690.800	3,13	3,19	3,13	3,20	3,02	3,56

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

Las cifras indican que no existe un patrón claro respecto al comportamiento de las medidas de precisión. En primera instancia podría señalarse que, en términos generales los métodos analíticos tienen un menor coeficiente de variación, por ende, una menor amplitud del intervalo de confianza (figura 5 y cuadro 4). Sin embargo, si se revisa el ingreso disponible del Resto de las capitales regionales, el método que registra un menor coeficiente es el de *subbootstrap* (método replicado), lo que refleja que no existe una dominancia absoluta de un método por sobre otro.

Figura 5: Intervalos de confianza para la mediana del Ingreso disponible mensual del hogar a nivel Total de capitales regionales



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

Cuadro 4: Amplitud del intervalo de confianza del Ingreso disponible mediano mensual del hogar, según área de estimación

Estimación	Mediana	Woodruff	Korn-Graubard	Binder	Bootstrap	Subbootstrap	Burn
Total de capitales regionales	779.000	65.857	67.605	67.507	72.996	71.208	68.052
Gran Santiago	857.360	116.953	120.641	123.999	123.441	123.479	98.557
Resto de capitales regionales	690.800	84.748	86.365	84.776	86.688	81.890	96.025

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

4.1.2 Gasto mensual del hogar

En relación con el gasto mensual de los hogares según área de estimación, en el cuadro 5 se observa que el método de Burn (réplicas) es el que presenta un menor coeficiente de variación, a excepción del área de Resto de capitales regionales, donde el método de Woodruff es el que presenta el valor más bajo. La ausencia de una tendencia plantea la interrogante de si existe una relación entre el número de observaciones y las medidas de precisión de los métodos analizados. La existencia de dicha relación constituye un antecedente relevante a tomar en consideración al momento de seleccionar un método. Este punto será abordado en el apartado 4.2 de resultados desagregados.

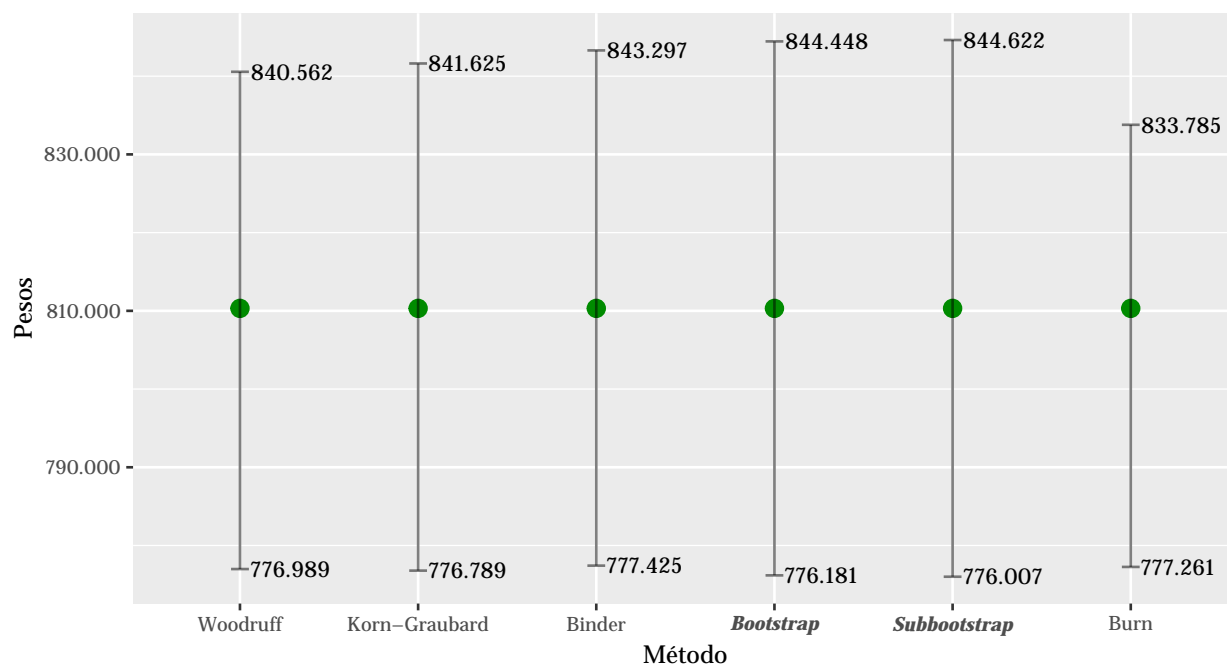
Cuadro 5: Coeficientes de variación del Gasto mediano mensual del hogar, según área de estimación

Estimación	Mediana	Woodruff	Korn-Graubard	Binder	Bootstrap	Subbootstrap	Burn
Total de capitales regionales	810.315	2,00	2,04	2,07	2,15	2,16	1,84
Gran Santiago	873.755	3,65	3,75	3,69	4,05	3,91	3,04
Resto de capitales regionales	735.881	3,13	3,25	3,14	3,24	3,17	3,54

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

La amplitud de los intervalos de confianza del gasto mensual a nivel Total de capitales regionales ilustra que el método de Burn es el que presenta una menor amplitud en relación con el resto. A su vez, es posible identificar la asimetría de este método el cual presenta un error superior de 23.470 y un error inferior de 33.054. No obstante, si se revisa el cuadro 6, las cifras por área de estimación, y consistente con lo planteado en los coeficientes de variación, el método de Woodruff presenta una menor amplitud en el Resto de capitales regionales. A pesar de esto, las diferencias son mínimas entre los métodos, lo cual permite tomar la decisión respecto al método a utilizar considerando resultados robustos y menos sensibles al método de cálculo de las medidas de precisión.

Figura 6: Intervalos de confianza para la mediana del Gasto mensual del hogar a nivel Total de capitales regionales



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

Cuadro 6: Amplitud del intervalo de confianza del Gasto mediano mensual del hogar, según área de estimación

Estimación	Mediana	Woodruff	Korn-Graubard	Binder	Bootstrap	Subbootstrap	Burn
Total de capitales regionales	810.315	63.573	64.836	65.871	68.267	68.615	56.524
Gran Santiago	873.755	125.140	128.373	126.500	138.860	133.794	102.693
Resto de capitales regionales	735.881	90.428	93.770	90.487	93.544	91.525	101.586

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

• Gasto mensual del hogar según división

Un resultado adicional corresponde al análisis del gasto mensual a nivel de división. En general, se observa que el método de *bootstrap* es el que presenta un mayor coeficiente de variación y, por tanto, también una mayor amplitud del intervalo. No obstante, en algunos casos es superado por el método de Binder (analítico). Dentro de los métodos replicados también se registran los valores más bajos del coeficiente de variación con el método de Burn, lo cual va en línea con la ausencia de dominancia de un enfoque por sobre otro.

Cuadro 7: Coeficientes de variación del Gasto mediano mensual del hogar, según división de la CCIF

Estimación	Mediana	Woodruff	Korn-Graubard	Binder	Bootstrap	Subbootstrap	Burn
Alimentos y bebidas no alcohólicas	173.780,00	1,79	1,82	1,76	1,74	1,74	1,34
Bebidas alcohólicas, tabaco y estupefacientes	5.595,32	5,47	5,65	5,57	6,12	5,76	5,08
Prendas de vestir y calzado	22.871,52	3,26	3,30	3,31	3,66	3,45	2,35
Vivienda y gastos en servicios básicos	101.116,92	3,03	3,06	3,05	3,15	3,23	2,32
Muebles y artículos para el hogar	23.977,90	5,12	5,14	5,05	5,35	5,33	5,67
Salud	24.982,00	4,87	5,03	4,85	5,10	5,17	3,60
Transporte	75.562,50	3,02	3,08	2,94	2,99	3,18	2,03
Comunicaciones	49.014,39	2,36	2,41	2,29	2,65	2,51	2,27
Recreación y cultura	32.528,99	3,46	3,48	3,39	3,43	3,25	2,63
Educación	416,67	28,06	28,06	NA	26,36	25,21	23,33
Restaurantes y hoteles	29.000,00	5,54	5,59	5,48	5,82	5,74	4,15
Bienes y servicios diversos	44.167,26	4,94	5,01	4,96	5,41	5,03	4,34

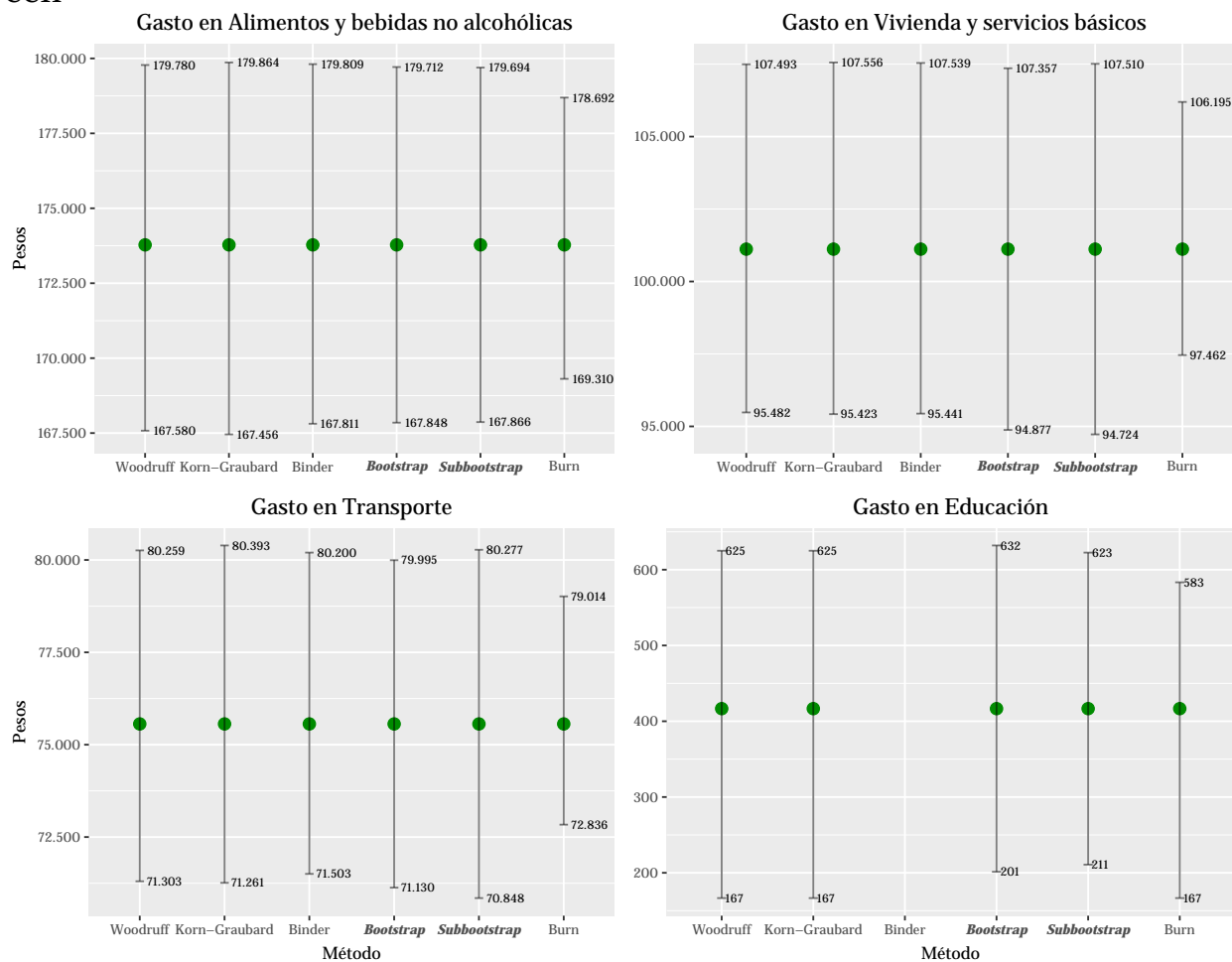
Fuente: Instituto Nacional de Estadísticas (INE) - VIII Encuesta de Presupuestos Familiares (EPF)

En la figura 7, a modo de resumen, se presentan los intervalos de confianza de cuatro de las doce divisiones de gasto. La elección de las divisiones se basa en que estas son las que tienen mayor participación en la estructura de gasto de los hogares¹³. En ellas se refleja de forma más nítida el comportamiento de los intervalos y la asimetría de algunos de los métodos propuestos. Cabe señalar que el método de Binder presenta una desventaja en los resultados presentados en este apartado. Y es que, por las características específicas de los datos de educación, el algoritmo no logra calcular

¹³Al observar la figura 7 se debe tener precaución con los ejes, debido a que los 4 gráficos ilustran distintos montos, por ende, diferentes valores en el eje *y*. Esto permite apreciar de forma más clara el comportamiento de los intervalos de confianza de las divisiones graficadas.

correctamente el límite inferior y superior del intervalo de confianza¹⁴. Debido a lo anterior, es que en la figura de los Gastos en Educación (figura 7) la barra del intervalo presenta un valor “NA”. Esta situación ocurre tanto en el cálculo del intervalo de confianza como al coeficiente de variación presentado en el cuadro 7 y el cuadro 8.

Figura 7: Intervalos de confianza para la mediana del Gasto mensual del hogar, según división de la CCIF



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

¹⁴Este método busca un punto donde el estadístico del test de score cruza un valor crítico con un 5% de confianza, sin embargo, no hay garantías de que tal punto exista. Al estudiar el código de fuente de la función *svyquantile* de R (específicamente el método *svyquantile.survey.design*) fue posible observar la existencia del cálculo de una raíz unitaria. Dicho cálculo requiere que el límite superior e inferior de la función para la cual se quiere calcular la raíz unitaria estén correctamente delimitados. Este paso presenta una falla de cálculo del intervalo de confianza cuando las distribuciones contienen una gran cantidad de valores cero.

Cuadro 8: Amplitud del intervalo de confianza del Gasto mediano mensual del hogar, según división de la CCIF

Estimación	Mediana	Woodruff	Korn-Graubard	Binder	Bootstrap	Subbootstrap	Burn
Alimentos y bebidas no alcohólicas	173.780	12.200	12.408	11.997	11.865	11.828	9.383
Bebidas alcohólicas, tabaco y estupefacientes	5.595	1.201	1.240	1.221	1.342	1.264	1.000
Prendas de vestir y calzado	22.872	2.926	2.955	2.963	3.277	3.097	2.094
Vivienda y gastos en servicios básicos	101.117	12.012	12.134	12.098	12.480	12.786	8.733
Muebles y artículos para el hogar	23.978	4.811	4.831	4.749	5.026	5.009	5.010
Salud	24.982	4.765	4.928	4.751	4.992	5.059	3.442
Transporte	75.562	8.957	9.132	8.697	8.865	9.429	6.178
Comunicaciones	49.014	4.533	4.630	4.395	5.085	4.814	4.038
Recreación y cultura	32.529	4.415	4.440	4.317	4.369	4.148	3.288
Educación	417	458	458	NA	431	412	417
Restaurantes y hoteles	29.000	6.302	6.358	6.235	6.613	6.530	4.620
Bienes y servicios diversos	44.167	8.549	8.671	8.586	9.375	8.716	7.353

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

A partir de los resultados agregados es posible hacer algunas observaciones generales. En primer lugar, los métodos generan resultados relativamente similares. En segundo lugar, y no obstante lo anterior, el método propuesto por Burn genera intervalos de confianza más estrechos que el resto, lo cual constituye una primera evidencia de que posiblemente no sea el más idóneo, ya que, siguiendo el argumento mencionado anteriormente, puede ser deseable construir intervalos de tamaño más conservador. En tercer lugar, el método de Binder no está disponible para todas las estimaciones revisadas a nivel agregado (división Educación), lo cual levanta una primera alerta respecto a la factibilidad de adoptarlo como método de cálculo de precisión debido a que existe una alta probabilidad que no sea factible para niveles de estimación desagregados.

4.2 Resultados desagregados

Tal como se mostró en el apartado 4.1, no existen grandes diferencias entre los distintos métodos cuando se realizan estimaciones con bajo nivel de desagregación (Total de capitales regionales y zonas). Ahora bien, esto no necesariamente se cumple para dominios de estimación más pequeños. A continuación, se muestran resultados desagregados de estimaciones de ingreso a nivel de personas. Específicamente, se revisan datos de ocupación (CISE y CIUO)¹⁵, edad, y escolaridad. En cada uno de estos casos se desagregan los resultados por sexo.

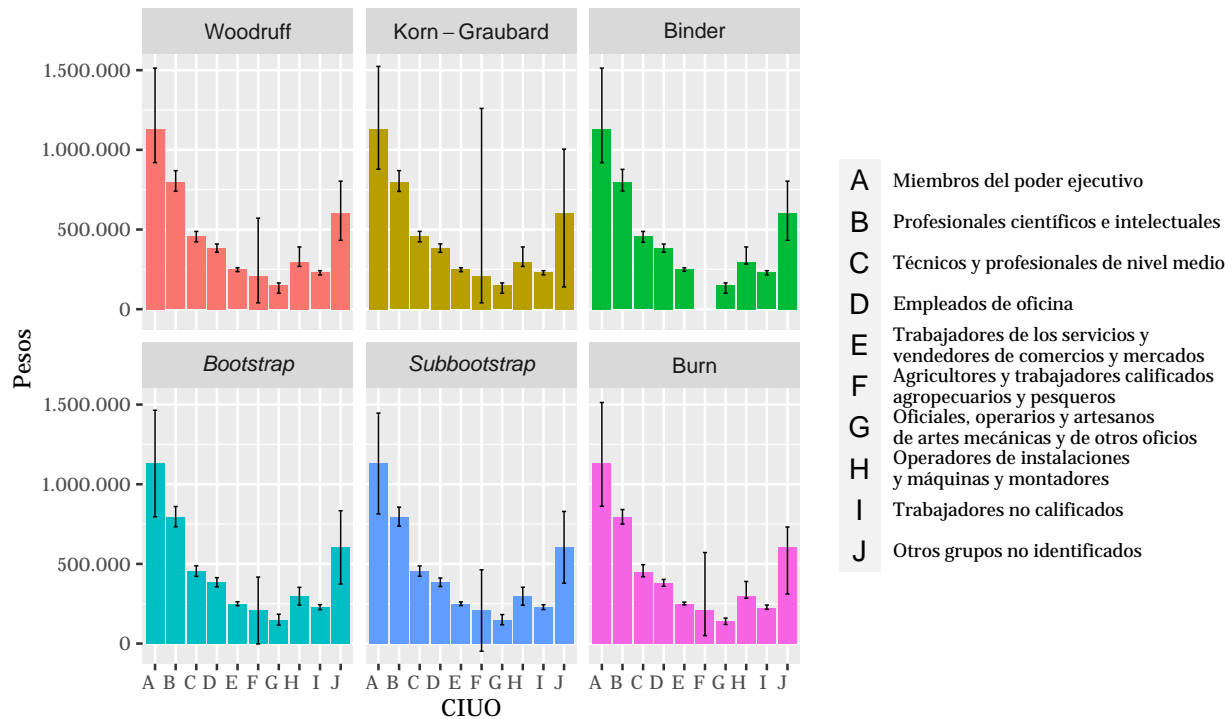
El motivo por el cual se utilizan datos de ingreso de las personas y no de gasto de los hogares dice relación con la necesidad de someter a los métodos a mayores niveles de exigencia, lo cual no es tan sencillo con datos de gasto, pues independiente de que un hogar realice o no determinado gasto, no es excluido del cálculo¹⁶. En definitiva, estas pruebas permiten llevar a cabo estimaciones considerando varias características simultáneamente, de modo de trabajar con subpoblaciones pequeñas. Con lo anterior se busca cumplir con el objetivo del documento asociado a tomar una decisión respecto a qué método utilizar para la construcción de intervalos de confianza para la mediana.

La figura 8 muestra medianas de ingreso para las distintas categorías ocupacionales de las mujeres. Lo primero que vale la pena mencionar es que, como ya se ha mostrado a nivel de Total de capitales regionales, el cálculo no está disponible para todas las categorías de esta variable. Así, la precisión para la categoría Agricultores y trabajadores calificados agropecuarios y pesqueros (F) en el método de Binder no es posible de ser calculada.

¹⁵Se utiliza el Clasificador Internacional de la Situación en el Empleo (CISE) y Clasificador Internacional Uniforme de Ocupaciones (CIUO-08).

¹⁶Tal como se presenta en la metodología de la encuesta, existen diversas formas de calcular el gasto promedio de los hogares. Para efectos del cálculo promedio de gasto presentado en los tabulados oficiales, en los hogares que no declaran gasto en determinados productos, se asume que su valor es cero (no realizaron gasto), por ende, entrarán en el cálculo. En cambio, a nivel de persona, no se presenta esta situación, pues los filtros en este caso excluyen las observaciones que no tienen la característica requerida (sexo, edad, situación en el empleo, etc.). Para efectos analíticos, a la hora de calcular el gasto promedio de los hogares que efectivamente realizan el gasto en determinados productos, al igual que el ingreso, es posible excluir del cálculo a aquellos hogares que no realizan el gasto. Para una explicación con mayores elementos de formalización, se sugiere revisar el apartado 6.3.2. de la metodología de la VIII EPF.

Figura 8: Estimación de la mediana e intervalos de confianza, según CIUO para mujeres a nivel de Total de capitales regionales



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

Un segundo aspecto para mencionar es que la amplitud de los intervalos de confianza varía de manera importante entre los distintos métodos. El caso que mejor ilustra esto es la categoría F, que en el método Korn-Graubard genera una amplitud de 1.220.000, distanciándose de manera importante de los demás métodos. Considerando la menor y la mayor diferencia, se observa que la amplitud del método Korn-Graubard es aproximadamente 2,4 veces más grande que la generada por el método de *subbootstrap* y 2,9 más que en el método de *bootstrap*. Estas diferencias se deben a que dicho método es muy sensible a la pérdida de grados de libertad¹⁷ y, de hecho, es el único que considera en su cálculo de manera explícita los grados de libertad disponibles.

Con el objeto de observar lo anterior, los cuadros 9 y 10 entregan información de precisión (amplitud del intervalo y coeficiente de variación) y grados de libertad en cada una de las estimaciones. Se advierte que para las categorías F y J existen 1 y 3 grados de libertad y, justamente, en estas categorías el método Korn-Graubard genera un intervalo de confianza notoriamente más amplio y un coeficiente de variación más elevado que los demás métodos.

¹⁷Los grados de libertad se calculan restando los conglomerados con información menos los estratos con información.

Cuadro 9: Grados de libertad y amplitud del intervalo de confianza para cada una de las categorías de CIUO (mujeres)

	Categoría	GL	Tamaño muestral	Woodruff	Korn-Graubard	Binder	Bootstrap	Subbootstrap	Burn
1	A	113	396	592.950	644.873	592.950	668.485	633.105	651.304
2	B	230	1851	129.512	130.687	135.642	126.877	119.031	91.104
3	C	243	1363	65.395	65.548	68.068	65.639	64.295	75.790
4	D	217	802	51.702	53.080	51.371	57.585	53.215	42.414
5	E	247	2937	24.266	24.550	21.580	24.202	23.520	17.016
6	F	1	21	531.258	1.220.000	NA	419.337	511.294	521.258
7	G	188	535	64.800	65.000	65.000	68.488	63.998	40.000
8	H	60	149	121.992	121.992	107.392	111.577	113.230	104.667
9	I	227	2161	26.746	26.746	26.746	31.806	30.087	26.746
10	J	3	26	370.899	865.036	370.899	459.991	450.292	420.204

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

Cuadro 10: Grados de libertad y coeficiente de variación para cada una de las categorías de CIUO (mujeres)

	Categoría	GL	Tamaño muestral	Woodruff	Korn-Graubard	Binder	Bootstrap	Subbootstrap	Burn
1	A	113	396	13,4	14,6	13,4	15,1	14,3	14,2
2	B	230	1851	4,1	4,2	4,3	4,1	3,8	3,0
3	C	243	1363	3,7	3,7	3,8	3,7	3,6	4,6
4	D	217	802	3,4	3,5	3,4	3,8	3,5	3,0
5	E	247	2937	2,5	2,5	2,2	2,5	2,4	1,5
6	F	1	21	65,3	150,0	NA	51,6	62,9	53,7
7	G	188	535	11,0	11,1	11,1	11,6	10,9	7,7
8	H	60	149	10,5	10,5	9,2	9,6	9,7	9,3
9	I	227	2161	3,0	3,0	3,0	3,5	3,4	4,9
10	J	3	26	15,7	36,5	15,7	19,4	19,0	18,3

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

La figura 9 muestra datos similares a los anteriores, pero ahora desagregando por la situación en el empleo. Nuevamente se presentan resultados para el caso de las mujeres. Adicionalmente, es importante mencionar que la categoría “Personal no remunerado” fue excluida, ya que por definición corresponde a personas sin ingreso, de modo que su inclusión no aporta información relevante para el análisis.

Por su parte, la categoría “Patrón o empleador” presenta un comportamiento particular, puesto que su mediana es cero y los límites inferior y superior también lo son, lo cual se debe a que la gran mayoría de los registros en la muestra es cero. Esto implica que los métodos que no utilizan remuestreo, por la forma en la que están contruidos, siempre arrojarán como resultado una amplitud de intervalo igual a cero¹⁸. Por su parte, los métodos que utilizan réplicas, teóricamente, tienen la probabilidad de generar medianas distintas de cero, pero debido a la gran cantidad de valores cero y dado que estos se encuentran distribuidos de manera más o menos homogénea entre todos los conglomerados¹⁹, es muy poco probable que la mediana calculada sea distinta de cero.

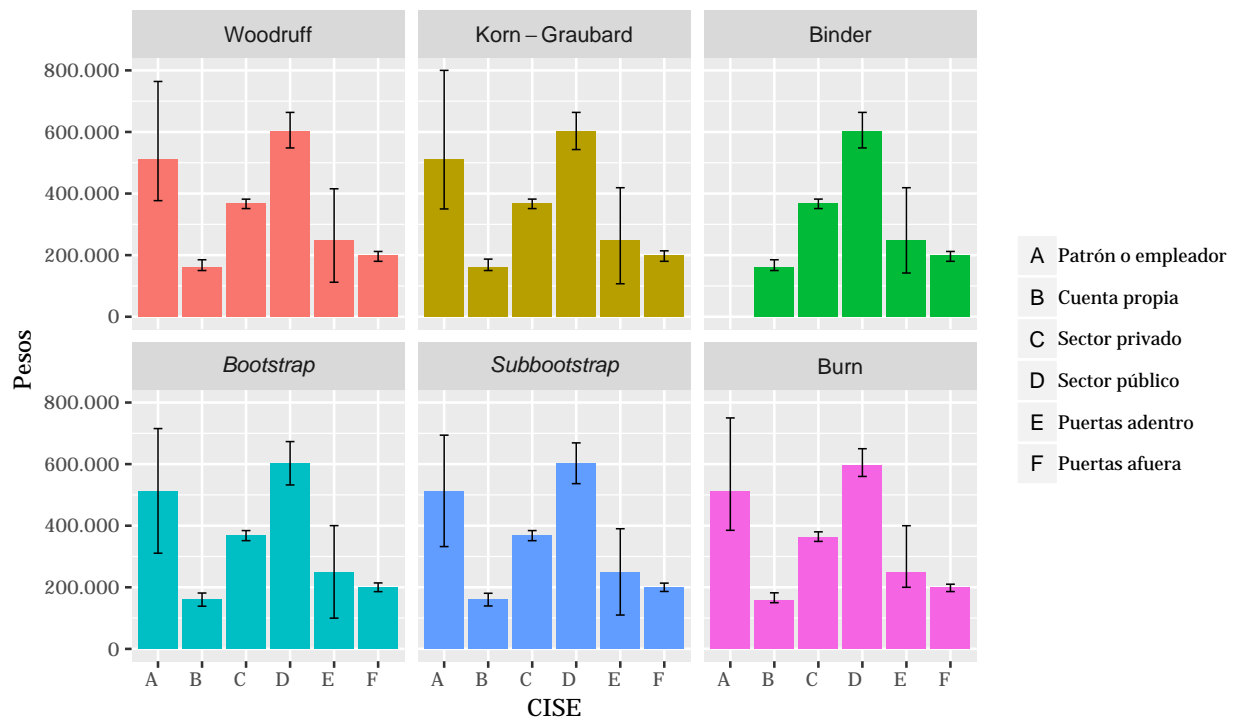
¹⁸Para más detalles, ver apartado 3.1 sobre descripción de los métodos de remuestreo.

¹⁹Para más detalles sobre el procedimiento de remuestreo, ver apartado 3.1.

Esta situación puede ser contraintuitiva, ya que el valor calculado es un estimador de la mediana, por ende, lo esperable es que tenga varianza, sin embargo, en casos en los que el mismo valor se repite muchas veces, los métodos utilizados arrojan como resultado que no existe error de estimación.

En lo que respecta a la comparación entre los distintos métodos, la inspección visual de la figura 9 da cuenta que en general no existen grandes diferencias. La única categoría que muestra diferencias un poco más visibles es la de “Servicio puertas adentro” (E). En este caso se observa que el método de Burn arroja un intervalo bastante más asimétrico que los demás. Por su parte, Korn-Graubard presenta una amplitud un poco mayor que el resto. En tanto, el método de Binder, al igual que en resultados anteriores, no genera resultados para la categoría “Patrón o empleador”.

Figura 9: Estimación de la mediana e intervalos de confianza, según situación en el empleo (CISE) para mujeres a nivel Total de capitales regionales



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

Los cuadros 11 y 12 corroboran de manera más exacta lo que muestra la figura 9, es decir, que en general la amplitud del intervalo y el coeficiente de variación son similares para los distintos métodos. Ahora bien, respecto a la ya mencionada categoría de Servicio puertas adentro (E) se observa la misma situación mostrada anteriormente, es decir, la disminución en los grados de libertad genera un intervalo más amplio y un coeficiente de variación más alto para el método de Korn-Graubard.

Cuadro 11: Grados de libertad y amplitud del intervalo de confianza según situación en el empleo (mujeres)

	Categoría	GL	Tamaño muestral	Woodruff	Korn-Graubard	Binder	Bootstrap	Subbootstrap	Burn
1	A	56	137	387.000	0	NA	404.694	361.730	365.000
2	B	248	2534	35.000	37.076	35.000	42.780	41.285	32.000
3	C	248	5308	30.549	30.841	30.749	32.675	32.780	31.122
4	D	240	1502	115.435	120.671	115.435	140.801	132.351	90.020
5	E	22	62	303.446	311.924	277.000	300.412	280.277	200.000
6	F	176	677	32.000	34.000	32.000	28.310	27.098	23.947

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

Cuadro 12: Grados de libertad y coeficiente de variación según situación en el empleo (mujeres)

	Categoría	GL	Tamaño muestral	Woodruff	Korn-Graubard	Binder	Bootstrap	Subbootstrap	Burn
1	A	56	137	19,2	22,4	NA	20,1	18,0	17,0
2	B	248	2534	5,6	5,9	5,6	6,8	6,6	6,7
3	C	248	5308	2,1	2,1	2,1	2,3	2,3	2,7
4	D	240	1502	4,9	5,1	4,9	6,0	5,6	4,7
5	E	22	62	31,0	31,8	28,3	30,7	28,6	27,2
6	F	176	677	4,1	4,3	4,1	3,6	3,5	2,5

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

4.2.1 Relación entre la precisión y características de la muestra

Los resultados desagregados del ingreso disponible de este apartado tienen la intención de ilustrar la existencia de algún tipo de patrón entre las medidas de precisión y las características de la muestra²⁰ de los seis métodos analizados. El objetivo de este análisis adicional es complementar los resultados de los apartados previos, y adicionalmente, identificar si un método es más idóneo que los demás, de modo de dilucidar con cuál de ellos se publicarán los tabulados oficiales de la mediana de ingresos y gastos de la encuesta.

Los análisis adicionales consisten en contrastar:

- la amplitud del intervalo de confianza con el número de observaciones
- el coeficiente de variación con el número de observaciones
- el coeficiente de variación con los grados de libertad²¹

En primer lugar, la figura 10 ilustra cómo disminuye la amplitud del intervalo de confianza cuando aumenta el número de observaciones, lo cual es esperable, ya que un mayor tamaño muestral debe generar estimaciones más precisas. Se observa una pendiente negativa más pronunciada en las variables de CIUO y educación, que en CISE y edad, debido a que estos últimos registran una mayor dispersión entre los métodos.

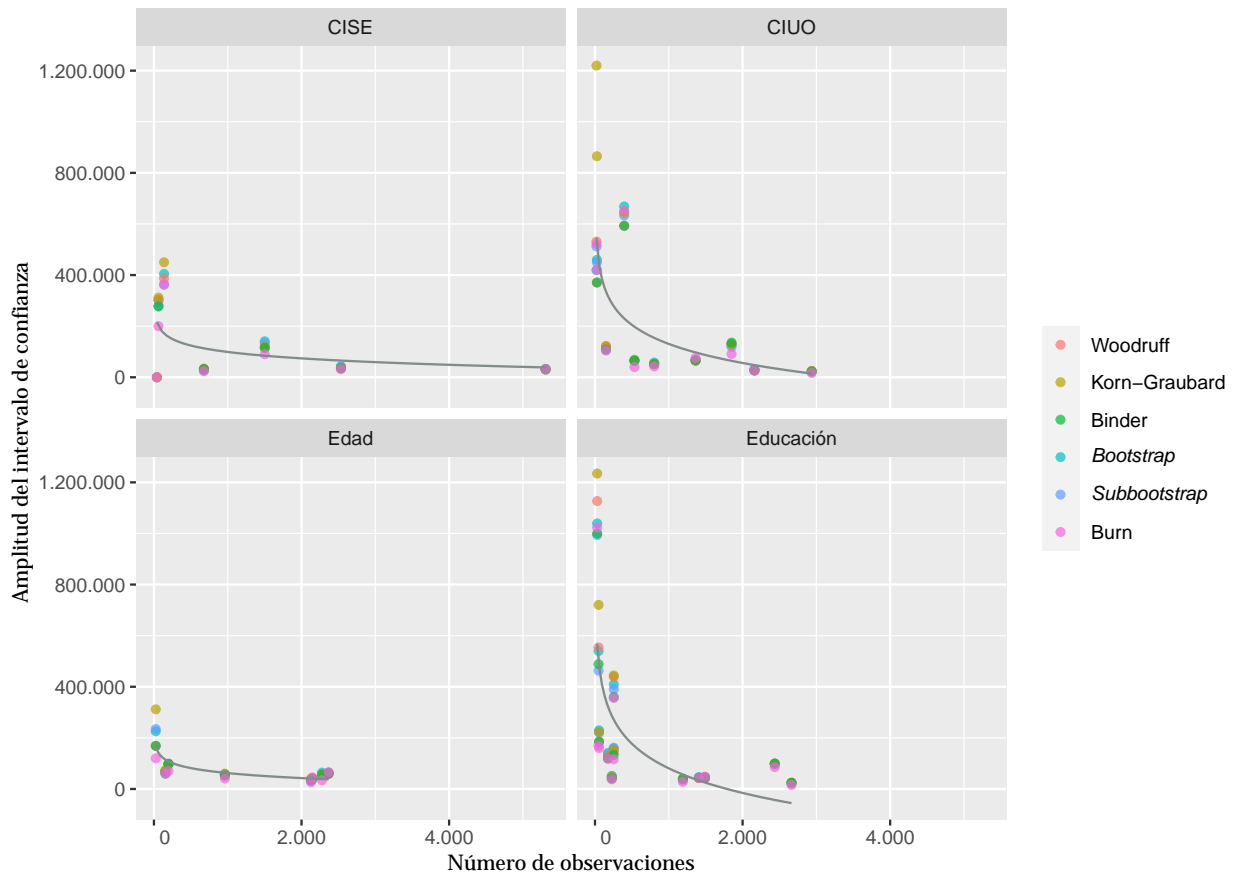
²⁰Tales como el número de observaciones y los grados de libertad de las distintas subpoblaciones estudiadas.

²¹También se contrastó la amplitud del intervalo de confianza con los grados de libertad, no obstante, da cuenta de los mismos resultados que al comparar el coeficiente de variación con los grados de libertad.

Si bien existen diferencias entre los distintos métodos, estas tienden a desaparecer conforme crece el tamaño muestral²². Así, cuando el número de observaciones es pequeño, se advierten diferencias, que van desapareciendo a medida que aumenta el tamaño muestral, lo cual se aprecia a partir de la superposición de puntos en la figura 10.

El comportamiento descrito refuerza lo observado en el apartado de resultados a nivel agregado, esto es, cuando el tamaño muestral es suficientemente grande, la elección de un método u otro se vuelve menos crítica.

Figura 10: Amplitud del intervalo de confianza según el número de observaciones (Mujeres)



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

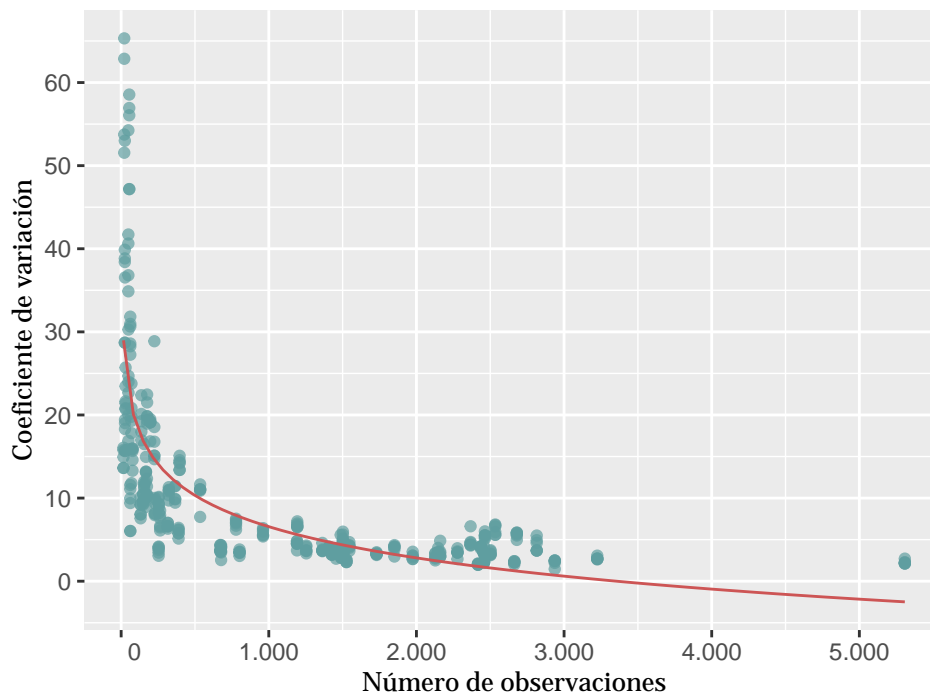
En segundo lugar, la figura 11 da cuenta de una relación negativa entre el coeficiente de variación y número de observaciones, es decir, a medida que aumenta el tamaño muestral, el coeficiente de variación se reduce²³. El resultado de esta relación establece que la dispersión de las subpoblaciones estudiadas disminuye al incrementar el número de observaciones de la muestra, esto significa que

²²En la figura 10 los puntos de colores representan el método utilizado según características del informante (edad, educación, CISE y CIUO) independiente de las categorías de cada una de ellas.

²³Con fines ilustrativos, la figura 11 se construyó truncando el número de observaciones a 6.000 casos y el coeficiente de variación a menos de 75%. Esta adaptación no modifica el análisis de los datos, su propósito es ilustrar de forma más clara la relación entre estas dos variables y la superposición de los datos.

un menor tamaño muestral registra una mayor dispersión en las cifras. Este comportamiento es un resultado independiente de las variables y métodos analizados. Asimismo, cabe señalar que el coeficiente de variación disminuye rápidamente al aumentar el número de observaciones, es decir, la precisión crece rápidamente cuando se realizan estimaciones de la mediana con mayores tamaños de muestra. Esto se encuentra en completa concordancia con los resultados de amplitud mostrados en la figura 10.

Figura 11: Coeficiente de variación según el número de observaciones de los seis métodos (CISE, CIUO, Edad, Educación)

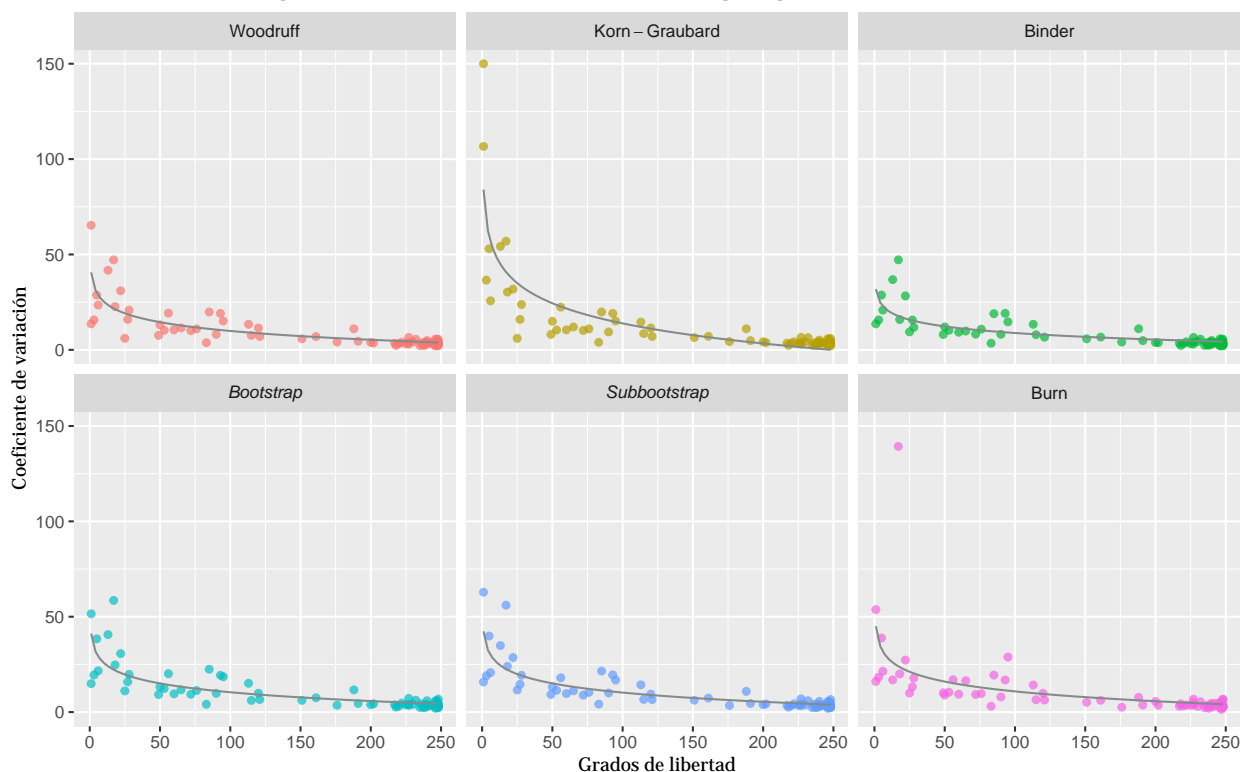


Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF)

El último ejercicio corresponde al contraste del coeficiente de variación con los grados de libertad de las estimaciones realizadas. Por lo tanto, la figura 12 identifica esta relación según los seis métodos estudiados. En términos generales, se registra una mayor dispersión de las cifras cuando los grados de libertad son bajos, independiente del método. Junto con lo anterior, existe una tendencia de los coeficientes de variación a converger a un menor valor cuando hay una gran cantidad de grados de libertad.

Sin embargo, el análisis de los resultados desagregados para dominios de estimación más pequeños debería incidir en pérdida de grados de libertad, tal como se indicó en el apartado 4.2. Este hecho se ve reflejado solo en el método de Korn-Graubard el cual tiene una pendiente más pronunciada que el resto de los métodos. La principal razón es que el cálculo de dicho método considera como variable los grados de libertad, por lo tanto, cada vez que se realiza una estimación en una subpoblación más pequeña el método de Korn-Graubard considera esta modificación, provocando una tendencia diferente en relación con los otros métodos.

Figura 12: Coeficiente de variación según grados de libertad



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

A partir de los resultados obtenidos, es posible establecer algunas consideraciones relevantes. En primer lugar, con el método de Binder no es posible calcular la precisión para todos los niveles de estimación, por lo que se toma la decisión de descartar su uso. En segundo lugar, el método de Korn-Graubard es sensible a la cantidad de grados de libertad, lo que en principio es algo positivo, sin embargo, en ciertos casos este método genera intervalos de confianza excesivamente amplios y asimétricos, lo cual hace desestimar su uso. En tercer lugar, un mayor número de observaciones tiende a reducir tanto la amplitud del intervalo de confianza como el coeficiente de variación. Es importante escoger un método que permita mantener un equilibrio entre la sobreestimación de la varianza y la disminución de la probabilidad de cometer el error tipo I.

Considerando esto último, los métodos de Woodruff y Burn son descartados debido a que generan intervalos asimétricos y un poco más estrechos que los de *bootstrap* y *subbootstrap*. De hecho, los resultados dan cuenta de que los métodos de *bootstrap* y *subbootstrap* presentan coeficientes de variación e intervalos de confianza más altos que los demás métodos. De este modo mantienen el criterio normativo de generar un intervalo un poco más amplio. Adicionalmente, es importante señalar que en el marco de este trabajo se valora positivamente el hecho de que los remuestreos *bootstrap* y *subbootstrap* generen intervalos de confianza simétricos, puesto que comunicacionalmente hace más sencilla la tarea de publicación de nuevos resultados. A raíz de ello, en lo que resta del documento, se busca definir entre *bootstrap* y *subbootstrap* como método de publicación de resultados, considerando cuál de los dos es el más idóneo para los datos de la EPF.

4.3 *Bootstrap versus subbootstrap*

Decidir entre utilizar un muestreo tipo *bootstrap* o *subbootstrap* no es sencillo. Desde un punto de vista conceptual, ambos son muy similares. La única diferencia, como se ha señalado anteriormente, radica en el tamaño de la muestra seleccionada. Mientras en *bootstrap* esta es de tamaño n , en *subbootstrap* es de tamaño $n - 1$. Por otro lado, desde un punto de vista empírico, los resultados revisados hasta el momento dan cuenta de que no existen diferencias importantes entre ambos muestreos en lo que respecta a la generación de medidas de precisión para la mediana.

Dado que no se cuenta con estadísticas de los “verdaderos” valores de las estadísticas respecto a las que se hace inferencia, se torna complejo decidir entre un tipo de remuestreo y otro. Con el objeto de abordar este problema, en el presente apartado se lleva a cabo una comparación del intervalo de confianza teórico y el generado por cada uno de los métodos. Debido a que la distribución de la media es conocida, se puede estimar su precisión por medio del diseño complejo. De este modo, es posible constatar cuál de los dos remuestreos (*bootstrap* o *subbootstrap*) se acerca más al intervalo de confianza teórico²⁴.

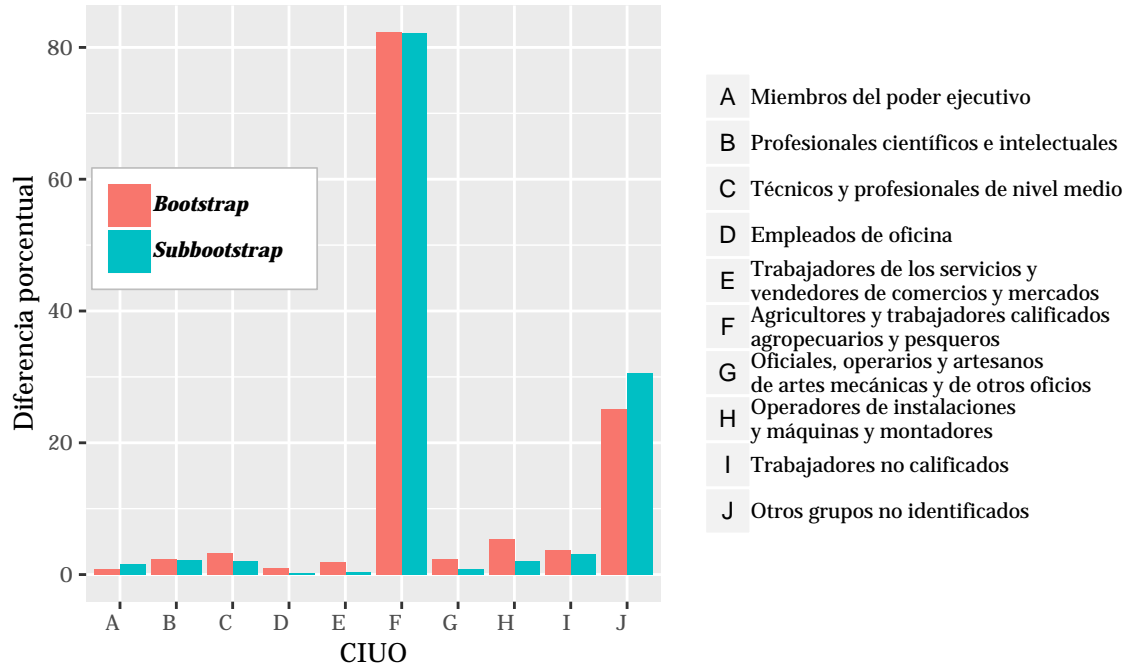
Las figuras 13 a la 17 muestran la diferencia porcentual que existe entre el intervalo generado para la media por cada método. Para ello, sencillamente, se calcula la diferencia entre la amplitud del intervalo que se calcula con el diseño de muestras complejas (dc), menos la amplitud generada por cada uno de los remuestreos. A dicha diferencia se le aplica el valor absoluto y ese valor se divide por el intervalo efectivo (dc), del siguiente modo:

$$dif = \left| \frac{amplitud_{dc} - amplitud_{remuestreo}}{amplitud_{dc}} \right| \quad (7)$$

En los gráficos se advierte que en la mayoría de las categorías ambos métodos arrojan resultados muy similares, sin embargo, en ciertos casos se observan diferencias importantes. Un ejemplo de ello es el caso de asalariados del sector público en la figura 14, donde se aprecia que *bootstrap* se aleja más del intervalo teórico que *subbootstrap*. En el sentido inverso, es posible encontrar situaciones en las que *bootstrap* se acerca más que *subbootstrap* al intervalo del diseño complejo. Tal es el caso del promedio de ingresos de las personas con doctorado en la figura 15.

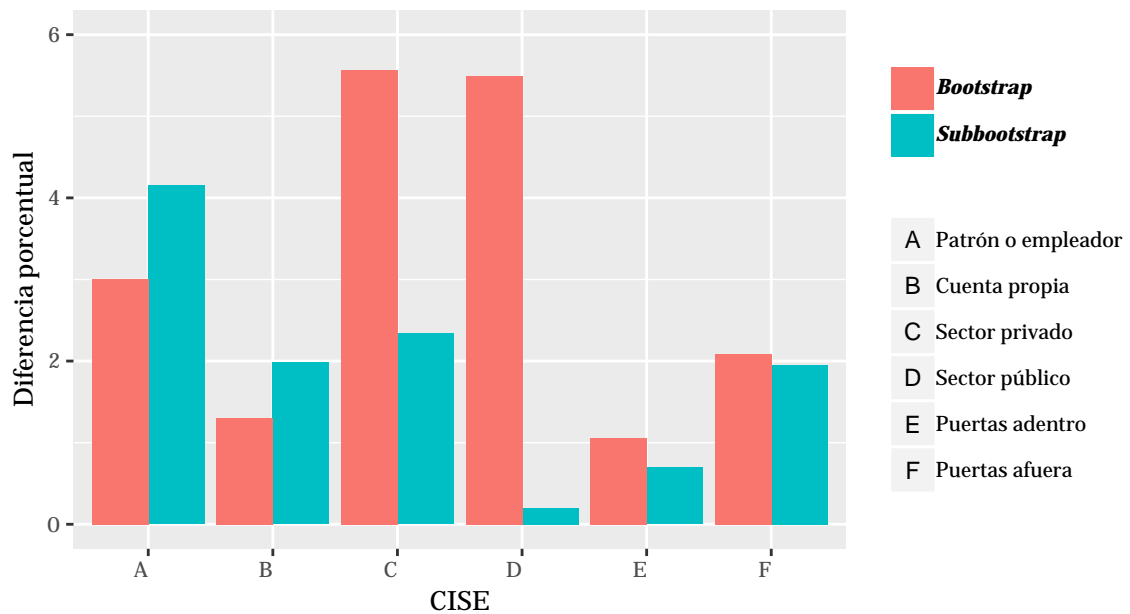
²⁴Para efectos de este ejercicio, se denomina intervalo teórico a aquel obtenido a partir del diseño complejo de la encuesta.

Figura 13: Diferencia porcentual respecto al intervalo de confianza del diseño complejo - Clasificación internacional uniforme de ocupaciones (CIUO-o8)



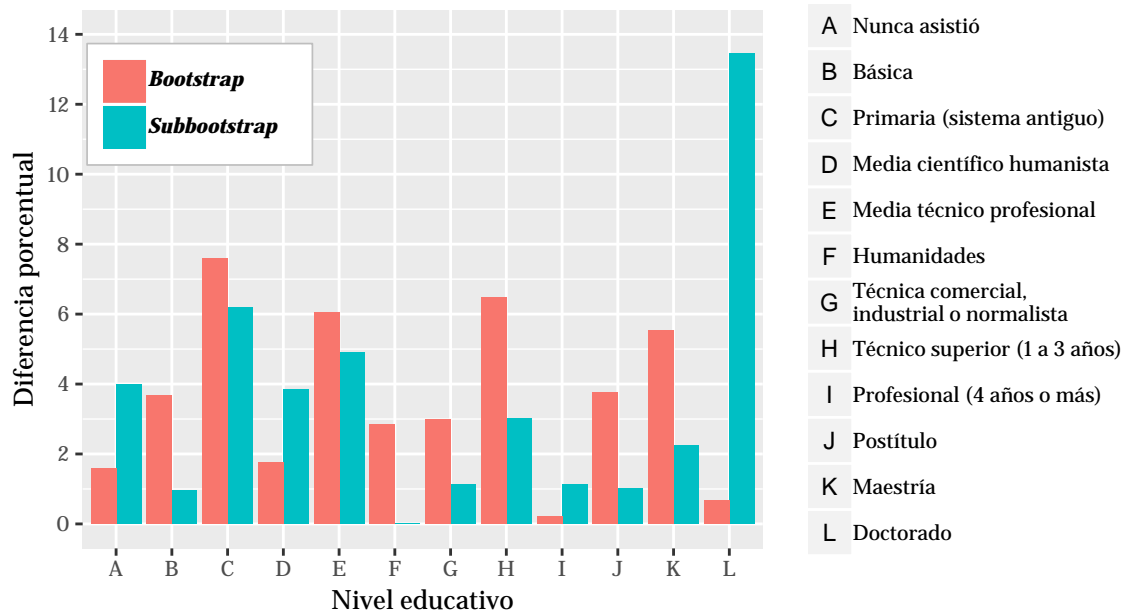
Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF)

Figura 14: Diferencia porcentual respecto al intervalo de confianza del diseño complejo - Clasificación internacional de la situación en el empleo (CISE)



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

Figura 15: Diferencia porcentual respecto al intervalo de confianza del diseño complejo - Nivel educativo

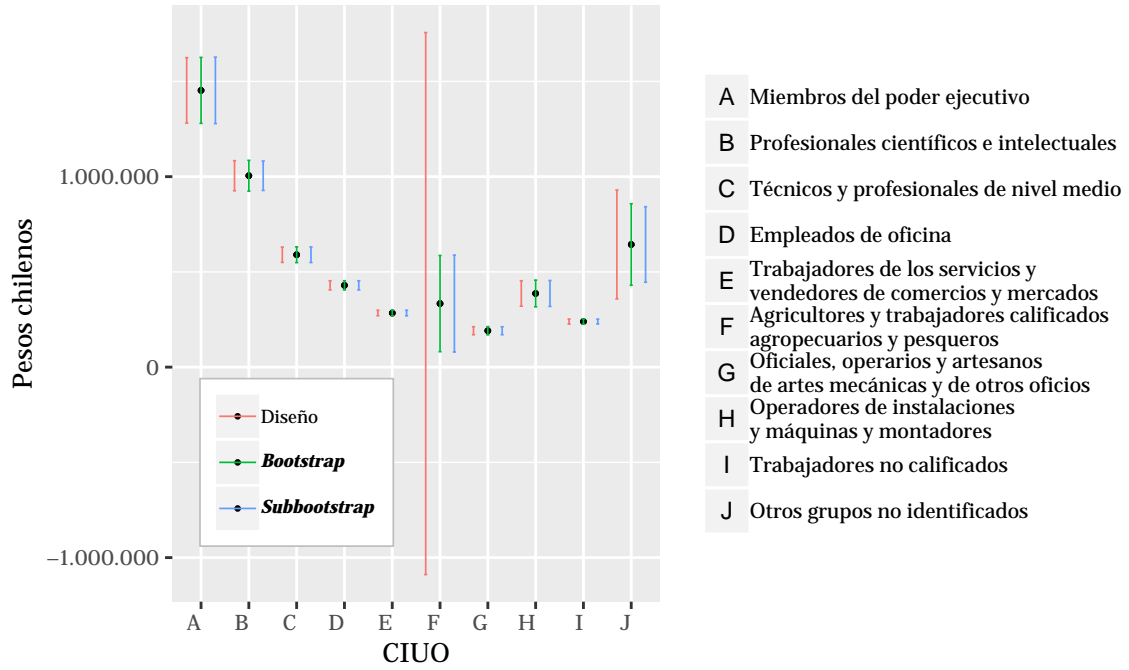


Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF)

Un fenómeno que vale la pena mencionar es que para ciertas categorías ambos tipos de remuestreos se alejan de la estimación proveniente del diseño complejo (efectivo). Esto ocurre cada vez que una estimación se lleva a cabo con pocos grados de libertad. En estos casos, las técnicas de remuestreo tienden a sobreestimar la variabilidad con la que se construyen las medidas de precisión del estimador. Un ejemplo de ello es la categoría Agricultores y trabajadores calificados agropecuarios y pesqueros (F) de la figura 16, donde se observa que ambos métodos se alejan notablemente del intervalo teórico. Esto sucede debido a que el cálculo del intervalo de confianza derivado del diseño complejo se lleva a cabo solo con un grado de libertad, a partir de lo cual se genera una alta imprecisión. Por su parte, los métodos de remuestreo construyen el intervalo de confianza con infinitos grados de libertad (ver apartado 3 sobre métodos), generando una amplitud mucho menor.

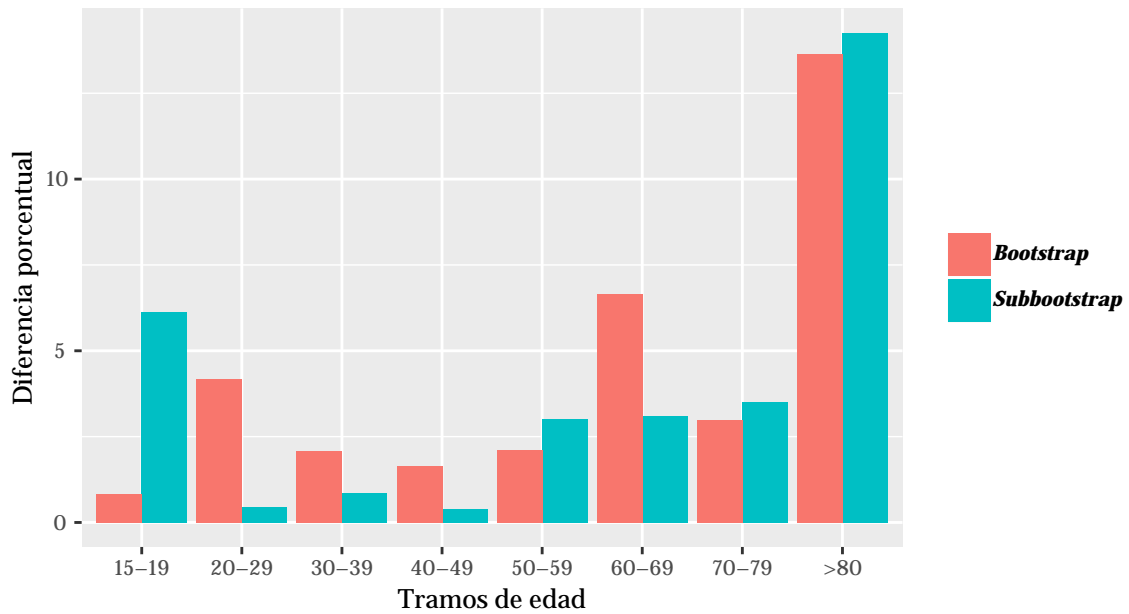
Respecto a este último punto, probablemente no sea adecuado llevar a cabo estimaciones con tan pocos grados de libertad, pues la precisión obtenida es muy baja, lo cual hace poco útil la información del estimador puntual.

Figura 16: Intervalos de confianza del diseño complejo, *bootstrap* y *subbootstrap* - Clasificación internacional uniforme de ocupaciones (CIUO-o8)



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

Figura 17: Diferencia porcentual respecto al intervalo de confianza del diseño complejo - Tramos de edad



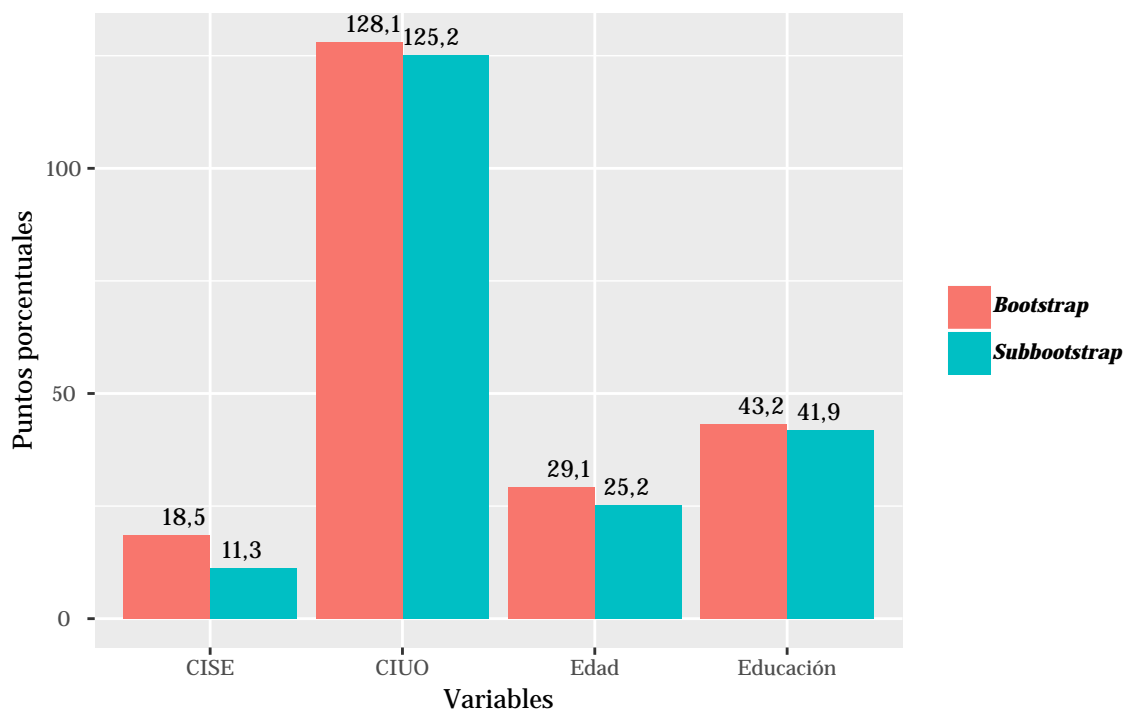
Fuente: Instituto Nacional de Estadísticas (INE) – VIII Encuesta de Presupuestos Familiares (EPF)

Las figuras 13 a la 17 permiten identificar claramente diferencias a nivel de categoría, pero no facilitan una comparación agregada. Es por ello que a continuación se presenta información agregada de cada una de las variables estudiadas. Esto se lleva a cabo simplemente sumando las diferencias de cada una de las categorías, a partir de lo cual se genera un indicador global de la distancia que tiene cada variable (CIUO, CISE, nivel educativo y edad) respecto al diseño complejo.

Los datos muestran que, para las cuatro variables (figura 18), el remuestreo *subbootstrap* genera intervalos de confianza más cercanos al intervalo construido a partir del diseño complejo (*dc*) que el remuestreo mediante *bootstrap*. Cabe señalar que se hicieron más ejercicios de este tipo con una serie de otras variables y se observó que, en general, el muestreo mediante *subbootstrap* tendía a generar intervalos de confianza más cercanos al intervalo del diseño complejo que el método de *bootstrap*.

Considerando estos resultados, el remuestreo mediante *subbootstrap* se priorizó como el más adecuado para generar los indicadores de precisión de la mediana. Cabe señalar que este método se utiliza actualmente para calcular coeficientes de variación para las medianas de ingreso de la Encuesta Suplementaria de Ingreso (ESI), de modo que es una estrategia que ya ha sido estudiada y validada por el Departamento de Estadísticas del Trabajo y por el Departamento de Metodologías e Innovación Estadística de la Institución.

Figura 18: Diferencia porcentual respecto al intervalo de confianza del diseño complejo



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

5 Conclusiones

El presente trabajo ha tenido el objetivo de revisar algunos de los métodos existentes para el cálculo de la precisión de la mediana (u otros percentiles). Esto con el objeto de identificar cuál es la estrategia que mejor se adecúa a las necesidades de la encuesta, en miras a la publicación de tabulados de ingresos y gastos. En ese sentido, este trabajo busca servir de apoyo a dicha publicación y al mismo tiempo poner a disposición de la ciudadanía la metodología utilizada.

A lo largo del documento se revisaron dos estrategias para el cálculo de la precisión de la mediana: 1) métodos que utilizan réplicas y 2) métodos analíticos. En primer lugar, se realizaron comparaciones de estimaciones a nivel agregado (Total de capitales regionales y por zonas), para luego mostrar resultados a niveles más desagregados, con subpoblaciones pequeñas. El objetivo de ello es identificar la sensibilidad de los métodos respecto a la disminución del tamaño muestral y a la pérdida de grados de libertad.

A nivel agregado, los datos dan cuenta de que no existen diferencias entre los distintos métodos, es decir, todos generan coeficientes de variación e intervalos de confianza relativamente similares. Esto es positivo, puesto que en lo que respecta al objetivo del diseño muestral, la decisión de qué método utilizar se vuelve menos crítica, pues los resultados no varían de manera importante entre un método y otro. De hecho, la asimetría de los métodos analíticos es bastante baja, lo cual permite que sean buenos candidatos a nivel agregado.

Sin perjuicio de lo anterior, a nivel de Total de capitales regionales se encontraron diferencias que apuntan a que los remuestreos *bootstrap* y *subbootstrap* generan intervalos un poco más amplios y coeficientes de variación más elevados que el resto de los métodos, sin embargo, vale la pena mencionar que esta situación no se cumple en la totalidad de las estimaciones. Los resultados a nivel de Total de capitales regionales y por zona aportan evidencia de que los remuestreos *bootstrap* y *subbootstrap* son preferibles a los demás métodos, en el sentido de que, sin la necesidad de perder demasiada precisión, generan mayor seguridad al considerar que el intervalo construido contiene la mediana de la población.

Otro hallazgo obtenido a nivel nacional es que el método de Binder no permite hacer el cálculo de medidas de precisión para la división de gasto en educación debido a que el algoritmo no es adecuado para las características de los datos. A raíz de ello se descarta su utilización, pues el método seleccionado debe, al menos, tener la capacidad de generar cálculos a nivel de Total de capitales regionales y por divisiones de gasto.

Respecto a las estimaciones con mayor nivel de desagregación se observan diferencias en las medidas de precisión generadas por los distintos métodos. En el caso de los métodos analíticos y de la propuesta de Burn, la asimetría de los intervalos de confianza a nivel desagregado se hace más evidente que a nivel agregado, lo cual es una característica poco deseable en las estimaciones de dichos métodos para ciertos casos.

Un segundo aspecto importante para considerar a la hora de decidir el método a utilizar es que el método de Korn-Graubard es sensible a los grados de libertad, generando una amplitud de intervalo y un coeficiente de variación más elevados que el resto de los métodos cuando existen pocos grados de libertad. Al respecto cabe señalar que existe una alta probabilidad de que los demás métodos estén

sobreestimando la precisión cuando la estimación se lleva a cabo con pocos grados de libertad.

Un último hallazgo relevante dice relación con el hecho de que, si bien a nivel desagregado se observan diferencias entre los métodos, los resultados tienden a converger rápidamente conforme aumenta el tamaño muestral. Esto es positivo, ya que, tal como se mencionó, la decisión sobre el método de cálculo de la mediana se vuelve menos crítica. Ligado con lo anterior, también se observa que a medida que aumenta el tamaño muestral, tanto los coeficientes de variación como la amplitud del intervalo tienden a disminuir, conforme a lo esperado, pues a medida que crece el número de observaciones las estimaciones de la mediana tienen una mayor precisión.

A partir de la información recabada se decide optar por una estrategia de remuestreo tipo *subbootstrap*. Los motivos para preferir este método por sobre los demás son los siguientes:

- 1) Genera un intervalo de confianza simétrico, lo cual hace más sencillo comunicar a los usuarios los resultados. Actualmente, el Instituto Nacional de Estadísticas no publica intervalos de confianza asimétricos, de modo que hacerlo en una publicación de resultados oficiales podría ser comunicacionalmente confuso.
- 2) Utilizar estrategias de remuestreo para calcular la varianza de un estimador no requiere hacer supuestos sobre la distribución del estadístico en cuestión, lo cual es algo deseable desde el punto de vista del trabajo con los datos. El único requerimiento es tener suficiente capacidad computacional para llevar a cabo la cantidad necesaria de remuestreos.
- 3) Actualmente, el Departamento de Estadísticas del Trabajo utiliza un remuestreo tipo *subbootstrap* para generar coeficientes de variación de medianas de ingreso en la Encuesta Suplementaria de Ingresos (ESI). En ese sentido, es una metodología que ya ha sido estudiada, validada y aplicada por la Institución.

Es importante mencionar algunos desafíos y aristas no consideradas en el marco de este trabajo, que debiesen ser abordadas en un futuro.

- La función del paquete *survey* de *R* para calcular la precisión de la mediana mediante remuestreo no es sensible a la pérdida de grados de libertad. Tal como se mencionó en el apartado 3 sobre los métodos, la forma de calcular el intervalo de confianza consiste en multiplicar el error estándar por un valor crítico de 1,96, lo cual implica asumir infinitos grados de libertad. Esto genera que incluso para estimaciones de subpoblaciones pequeñas con muy pocos grados de libertad, el intervalo de confianza obtenido siga siendo razonable. Dicho de otro modo, el método de cálculo no es sensible a la pérdida de grados de libertad, lo cual puede conducir a una sobreestimación de la precisión. En este sentido, no debiesen considerarse seriamente estimaciones hechas con muy pocos grados de libertad, ya que la imprecisión en estos casos es tan alta que vuelve poco útil la información obtenida. Pese a ello, es deseable que el método utilizado dé cuenta correctamente de la pérdida de precisión que se genera producto de la disminución de grados de libertad. Dado que no es posible conocer el “verdadero” intervalo de confianza de la mediana, identificar qué tan bien o mal calculado está aquel es algo sumamente complejo. Afortunadamente, el método de Korn–Graubard considera en su cálculo los grados de libertad disponibles. Los resultados muestran que cuando existen pocos grados de libertad, dicho método genera intervalos notoriamente más amplios que los métodos replicados, lo cual da una idea de que la

precisión en el caso de *bootstrap* puede estar sobreestimada. En este sentido, un desafío futuro es desarrollar una estrategia de remuestreo que contemple la cantidad de grados de libertad disponibles.

- Dado que los métodos probados generan resultados diferentes en dominios pequeños de estimación, es posible desarrollar una estrategia que combine de alguna manera los resultados de cada uno de ellos en un único valor.
- En el marco de este trabajo, los remuestreos realizados utilizaron los factores de expansión finales, es decir, no se calibraron los pesos para cada muestra obtenida según un único *stock* poblacional. Esta decisión tiene la ventaja de que no requiere información auxiliar para llevar a cabo el cálculo de precisión de la mediana, lo cual hace que sea relativamente sencillo para los usuarios implementar cualquiera de los métodos aquí revisados, entregándoles la libertad de decidir qué metodología usar. Ahora bien, la desventaja de trabajar con los factores de expansión finales es que esto no permite llegar a los mismos *stocks* poblacionales en cada una de las iteraciones del remuestreo, introduciendo una suerte de inconsistencia. En ese sentido, un desafío futuro es implementar un remuestreo que funcione idénticamente al utilizado en este trabajo, pero con la diferencia de que incorpore una calibración específica para cada una de las iteraciones. Esto permitiría tener un mayor nivel de consistencia, pero al mismo tiempo haría un poco más complejo para los usuarios tomar la decisión respecto a qué metodología usar porque no podrían replicar los cálculos de este documento.

Finalmente, cabe mencionar que este documento no pretende agotar el tema sobre cálculo de precisión para percentiles. La experiencia nacional al respecto es limitada, por lo que este tema requiere mayor desarrollo y análisis empírico. La generación de nuevas pruebas podría contribuir a tener una evaluación más clara del desempeño que tiene la metodología aquí utilizada en esta y en otras encuestas de hogares. En ese sentido, este documento busca incentivar la discusión sobre el tema, de modo que surjan más trabajos relacionados, que permitan el desarrollo de nuevas propuestas para el cálculo de medidas de precisión de percentiles, en el contexto de encuestas de hogares.

Referencias

- Binder, D. (1991). Use of Estimating Functions For Interval Estimation from Complex Surveys. *Proceedings of the ASA Survey Research Methods Section*, 34-42.
- Burn, D. (2006). The use of resampling for estimating confidence intervals for single site and pooled frequency analysis. *Hydrological Sciences Journal*, 48(1), 25-38.
- Carpenter. (1999). Test inversion bootstrap confidence intervals, 61(N° 1), 159-172.
- Efron. (1979). Bootstrap Methods: Another Look at the Jackknife, 7(N° 1), 1-26.
- INE. (2018). *Metodología VIII EPF*.
- Korn, E., & Graubard, B. (1998). Confidence Intervals for Proportions With Small Expected Number of Positive Counts Estimated From Survey Data. *Survey Methodology Statistics Canada*, 24(2), 193-201.
- Lumley, T. (2010). *Complex Surveys. A guide to Analysis Using R*. John Wiley & Sons, Inc.
- Woodruff, R. (1952). Confidence Intervals for Medians and Other Position Measures. *Journal of the American Statistical Association*, 47(260), 635-646.