

COMPARACIÓN DE LA VARIACIÓN INDUSTRIAL DE LAS CUENTAS NACIONALES CON UN INDICADOR DE MERCADO

Francisco Ruiz A.
Felipe Stanger V.*

I. INTRODUCCIÓN

En el cálculo de la variación mensual del Valor Agregado Industrial de Cuentas Nacionales (CCNN) se utiliza principalmente información de las tasas de variación de las ramas de actividad del Índice de Producción y Venta Física (IPVF) del INE. La información pública de este índice tiene un mayor nivel de agregación. El INE hace públicas las tasas de variación de las ramas industriales del IPVF a tres dígitos CIU. En general, se utilizan estos índices a tres dígitos para 68% del valor agregado de CCNN, a cuatro dígitos CIU para 28,8% y fuentes distintas al IPVF para el 3,2% restante.

El objetivo de esta nota es comparar las tasas de crecimiento del sector Industria del Imacec con las que comúnmente estiman los analistas económicos utilizando la información a su alcance, esto es, con un nivel mayor de agregación y con ponderaciones diferentes a las disponibles en CCNN.

Se analizan tres métodos alternativos para aproximarse a la evolución de la actividad industrial presente en el Imacec. El análisis permite ver la bondad de ajuste de cada una de las alternativas, con el objeto de determinar la magnitud de los errores y eventuales sesgos que cometerían los analistas por no contar con toda la información necesaria para hacer los cálculos coyunturales de CCNN.

El ejercicio se realizó para el período enero 1997 - diciembre 2003, y los resultados indican una alta correlación entre las estimaciones, con diferencias absolutas mensuales que en el extremo nunca superan un punto porcentual. Este resultado demuestra que los analistas están en condiciones de aproximarse satisfactoriamente a las estimaciones de CCNN con métodos de cálculo relativamente simples.

II. CÁLCULO DEL IMACEC INDUSTRIAL

El Banco Central estima la actividad industrial de acuerdo con la variación del valor agregado mensual de los diferentes subsectores, para lo que usa un índice de Laspeyres con base 1996. El valor agregado de la actividad industrial se determinó para la matriz insumo-producto (MIP) del año 1996, desagregado en 37 agrupaciones (líneas) industriales.¹ Por lo tanto, a partir de ese año estos valores se van proyectando a través de indicadores económicos que representan de la manera más fidedigna posible la realidad de cada subsector. Las principales características del índice calculado son las siguientes:

- Año base: 1996
- Para las subactividades de alimentos y bebidas, 14 líneas utilizan las tasas de variación del IPVF a cuatro dígitos CIU. Este nivel de desagregación, cuatro dígitos CIU, no es público para los analistas económicos, debido a una política de representatividad y secreto estadístico de parte del INE. Por otro lado, dos de estas líneas incorporan información adicional al índice (aceites y pesca).
- Las líneas que utilizan las tasas de variación del IPVF a tres dígitos CIU² son 23.

La construcción del índice agregado (índice de cuántum de Laspeyres) es la siguiente:

$$I_t = \sum_i I_{i,t} \cdot \alpha_i \quad ; \quad i = 1, \dots, 37$$

donde

I_t : índice agregado del mes "t"

$I_{i,t}$: índice línea "i" del mes "t"

α_i : ponderación de la línea "i" en la matriz de insumo-producto de 1996

* Departamento de Cuentas Nacionales, Gerencia de Información e Investigación Estadística, División de Estudios, Banco Central de Chile.

¹ La matriz insumo-producto se presenta en su versión "ampliada", que incluye 73 actividades desagregadas, de las cuales 37 corresponden a líneas industriales. Para referencia ver:

<http://www.bcentral.cl/esp/estudiospublicaciones/publicaciones/cuentasnacionales/xls/matriz/tablas73x73.xls>

² Tres de estas 23 líneas resultan de reponderaciones del IPVF a tres dígitos CIU.

Este índice de Laspeyres se valora a precios promedio del año 1996, y bajo el supuesto de productividad constante. Para el cálculo de la actividad industrial implícita del Imacec, se extrapola usando indicadores de producción. Sin embargo, una vez que se tienen las cuentas anuales, provisionales o revisadas, se corrigen estos niveles mensuales. Por este motivo, la forma de actualizar el crecimiento en doce meses de cada línea y del agregado, para el análisis de la coyuntura, es la siguiente:

donde

$$\left(\frac{I_t}{I_{t-12}} - 1\right) = \sum_i \left(\frac{I_{i,t}}{I_{i,t-12}} - 1\right) \cdot \left(\frac{\alpha_i \cdot I_{i,t-12}}{I_{t-12}}\right) \Leftrightarrow$$

$$\left(\frac{I_t}{I_{t-12}} - 1\right) = \sum_i \left(\frac{I_{i,t}}{I_{i,t-12}} - 1\right) \cdot \alpha_{i,t-12} \quad ; \quad i = 1, 2, \dots, 37$$

$\alpha_{i,t-12}$: nueva ponderación base, una vez actualizada la evolución de las líneas en el mismo mes del año anterior con respecto al total de ese mes.

En síntesis, se actualizan las ponderaciones de cada mes por el valor agregado del mismo mes del año anterior sobre el valor agregado total del respectivo mes.

III. ESTIMACIÓN DE ANALISTAS

El propósito de los siguientes ejercicios es revisar distintas formas de construir una aproximación a las variaciones de la actividad industrial entregada por CCNN, con la información disponible en el mercado (con un menor nivel de agregación). Las estimaciones se realizan para el período enero 1997 – diciembre 2003.

La primera opción consiste en utilizar directamente la tasa del IPVF como *proxy* de la estimación del Imacec. Una segunda aproximación consiste en construir una tasa agregada a partir de los índices a tres dígitos CIU del INE, ahora ponderados por la participación fija obtenida en la MIP del año 1996. La tercera opción —y la más similar a la que emplea actualmente el Banco Central para la elaboración de dicha cifra— consiste en agregar los índices a tres dígitos CIU del INE usando el método de Laspeyres con las ponderaciones de la MIP del año 1996, vale decir, permitiendo una actualización de ponderaciones implícita propia de este tipo de

índices. Tanto la alternativa segunda como la tercera utilizan información a tres dígitos CIU, pues este es el nivel de desagregación del que se dispone públicamente (28 líneas CIU a tres dígitos). Sin embargo, es importante destacar que las variaciones de algunas de estas 28 líneas no coinciden con las variaciones de CCNN, a este nivel de desagregación, por lo expuesto en el punto II.

1. Resultados con el Método 1 (uso directo de la tasa del IPVF)

Al comparar la evolución del índice agregado del INE (variaciones) con la del Imacec, se observa una diferencia promedio entre ambas estimaciones que fluctúa en torno a 1.2% con un coeficiente de variación de 0.77. El coeficiente de correlación entre ambas series es de 0.95, lo que refleja un movimiento de tasas muy similar. Mirando los percentiles, se extrae que 90% de las diferencias se encuentra en el rango entre -2.5% y 2.2% (cuadro A1). Los resultados se presentan en el gráfico 1.

Una de las ventajas de mirar este indicador radica en que la información del INE tiene una metodología clara y de libre acceso para los analistas, además de ser de carácter oficial. Sin embargo, las ponderaciones de ambos índices son diferentes, por lo que se procede, en una segunda aproximación, a reponderar los índices a tres dígitos CIU del INE con la matriz de 1996.

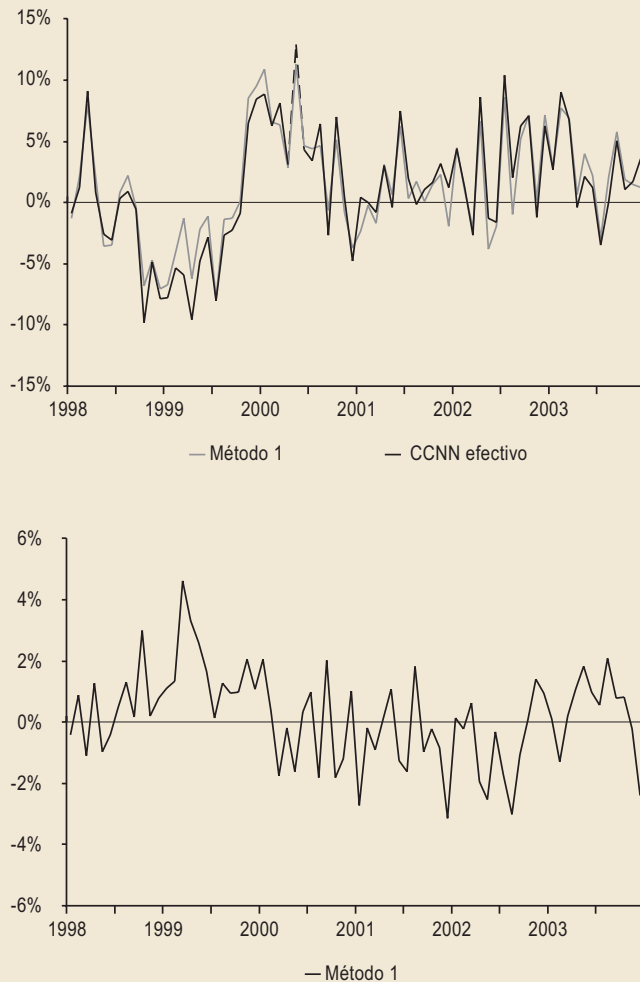
2. Resultados con el Método 2 (reponderando los índices con la MIP 96)

En este método se utilizan los valores agregados de la MIP del año 1996, y luego se extrapolan a través de los índices (a tres dígitos CIU) del INE, manteniendo las ponderaciones de dichos valores agregados en el año base. La forma de estimar la variación del índice es la siguiente:

$$\left(\frac{I_t}{I_{t-12}} - 1\right) = \sum_i \left(\frac{I_{i,t}}{I_{i,t-12}} - 1\right) \cdot \alpha_i \quad ; \quad i = 1, \dots, 28$$

En este caso, la diferencia promedio (en valores absolutos) está en torno a 1.1% con un coeficiente de variación de 0.99. La correlación entre ambas series, para el período antes especificado, es de 0.96.

Tasas de Variación en 12 Meses y Diferencias Respecto de la Tasa Efectiva



En cuanto a los percentiles, el rango para 90% de las diferencias es de -2.7% a 2.1% , dejando ver un leve sesgo negativo a las diferencias, ya que tiende a subestimar la cifra efectiva (cuadro A1). Los resultados se presentan en el gráfico 2.

Al comparar este método con el anterior, se observa que no hay mejoras sustanciales en algunos estadísticos (y un leve deterioro en otros). Ello puede obedecer a que el procedimiento para construir una aproximación de la variación del índice de producción industrial presume que la estructura productiva no ha sufrido modificaciones, por lo que no incorpora las actualizaciones implícitas en un índice de Laspeyres, lo que es poco realista. Dada

esta deficiencia, se elabora un tercer indicador con ponderaciones actualizadas obtenidas a partir de la MIP de 1996.

3. Resultados con el Método 3 (ponderación actualizada)

Con este método también se utilizan los valores agregados de la matriz insumo-producto del año 1996 (ponderación) y se extrapolan a través de los índices (a tres dígitos CIU) del INE. La variante radica en que las ponderaciones resultantes se actualizan según la metodología implícita en un índice de Laspeyres. De este modo, el método 3 es equivalente a lo que se realiza en CCNN, salvo el nivel de desagregación con que se trabaja la información. En efecto, en el caso de CCNN se dispone de información para las 37 líneas entregadas en la MIP, mientras que la información disponible para los analistas externos es solamente de 28 líneas.

La forma en que los analistas externos estiman la variación del índice con la información disponible, a un nivel de 28 líneas, es la siguiente:

$$\left(\frac{I_t}{I_{t-12}} - 1 \right) = \sum_i \left(\frac{I_{i,t}}{I_{i,t-12}} - 1 \right) \cdot \left(\frac{\alpha_i \cdot I_{i,t-12}}{I_{t-12}} \right); i = 1, 2, \dots, 28$$

Los resultados para este caso muestran una mejora relativa, ya que la diferencia promedio (en valores absolutos) es 0.9% (menor que los casos anteriores) con un coeficiente de variación de 0.99 . La correlación entre ambas series, para el período antes especificado, es de 1.03 , la más alta entre las alternativas. Respecto de los percentiles, el 90% de las diferencias se encuentran acotadas entre -2.5% y 1.7% con un sesgo promedio aún inclinado hacia la subestimación aunque de menor amplitud que las

otras opciones (cuadro A1). Los resultados se presentan en el gráfico 3.

En este tercer caso, podemos decir que la precisión ha mejorado. Si bien no se observan diferencias sustanciales entre los resultados de las tres opciones, parecería ser que esta última es más precisa, ya que el promedio es el menor y la correlación es la más alta. Respecto de las diferencias, la amplitud de la serie al 90% es menor que las otras disyuntivas aunque, en promedio, en los extremos tiende a subestimar el valor efectivo.

Adicionalmente, a modo de ejercicio, se realizó una sofisticación a este método 3, consistente en suponer que a los analistas se les entregan los valores agregados actualizados a tres dígitos del año anterior, lo que les permitiría tener los ponderadores correctos a ese nivel. Los resultados muestran que dicha información no mejora significativamente los resultados ya obtenidos a través de los métodos expuestos (ver los resultados en el apéndice, bajo el título “Método 4”).

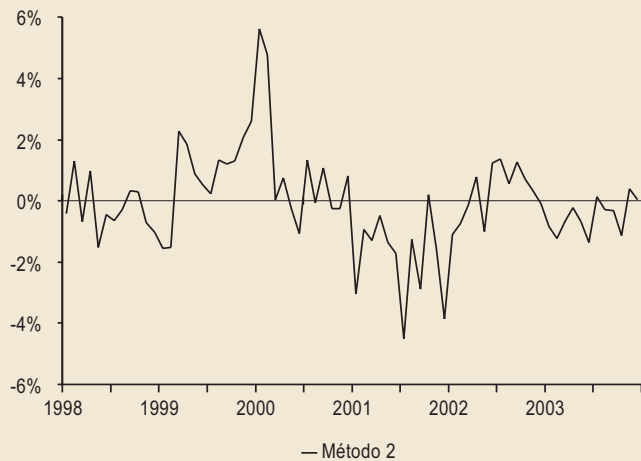
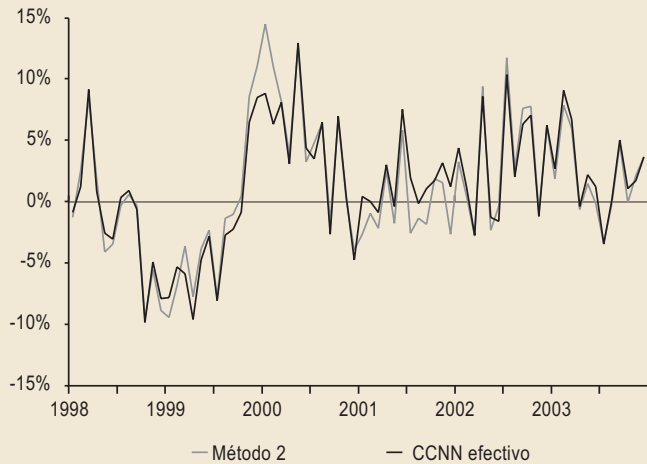
IV. PREVISIÓN DE LOS ERRORES

Adicionalmente, se revisó el comportamiento de los errores en los que se incurre al estimar la variación de la industria con la mejor metodología utilizable con los datos disponibles (método 3). La revisión consistió en determinar si los errores presentan algún patrón, que permita descartar o no la hipótesis de “ruido blanco”, con el propósito de evaluar la capacidad predictiva del modelo seleccionado.

Para realizar este ejercicio, se revisaron las diferencias entre los datos efectivos y los datos estimados para el período comprendido entre enero 1997 y agosto 2003. La serie seleccionada no presenta una raíz unitaria, y se procedió a identificar una representación Sarma.

GRÁFICO 2

Tasas de Variación en 12 Meses y Diferencias Respecto de la Tasa Efectiva



El modelo seleccionado fue:

$$(1 - \phi_1 L - \phi_3 L^3) \eta_t = \mu + (1 - \theta_{12} L^{12}) v_t$$

donde

η_t : son las diferencias entre las tasas estimadas y las tasas efectivas;

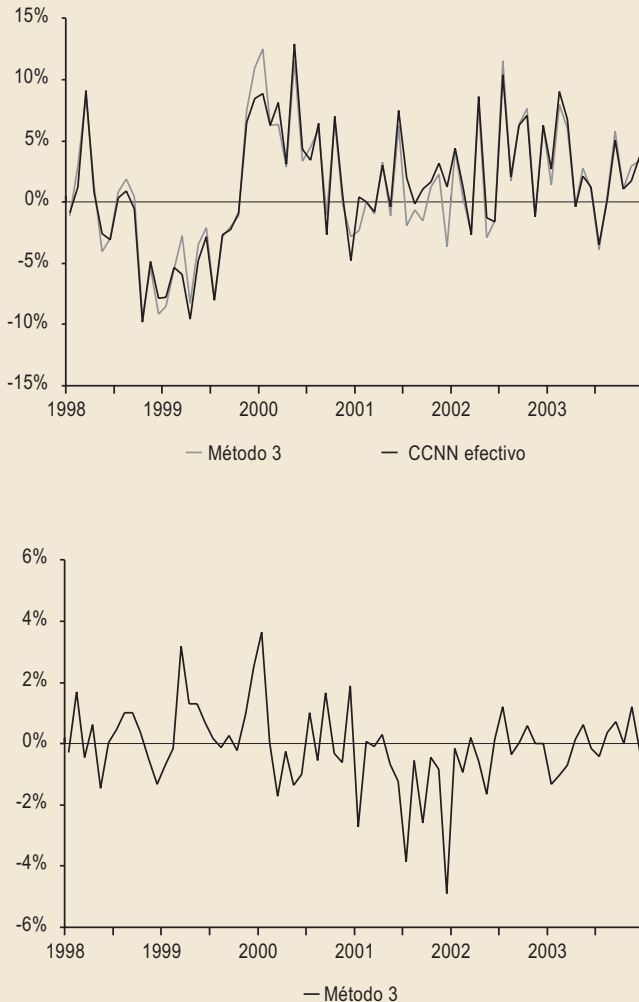
$v_t \sim \text{Ruido Blanco } (0, \sigma^2)$: son errores estimados a partir de la especificación anterior.

$$\eta_t = -0.001 + 0.341 * \eta_{t-1} + 0.216 * \eta_{t-5} + v_t - 0.718 * v_{t-12}$$

(0.0010) (0.0976) (0.0995) (0.1327)

(los valores entre paréntesis corresponden a las desviaciones estándar).

Tasas de Variación en 12 Meses y Diferencias Respecto de la Tasa Efectiva



Al comparar los valores estimados según el método 3, corregidos por la proyección de los errores, con las tasas efectivas de Cuentas Nacionales, se encuentra que no habría una mejora sustancial. Este resultado no sorprende, ya que la evaluación cubrió los últimos dos años, para los que la aproximación del método 3 a las variaciones efectivas es especialmente buena. En efecto, los errores de estimación se encuentran, en promedio, en torno a 0.5% durante ese período. Sin embargo, en los períodos previos, la diferencia promedio llega a sobrepasar el punto y medio porcentual. Es para esos períodos que la estimación de errores ayuda a tener una mejor aproximación del IPFV del INE al valor agregado del Imacec de industria, reduciendo

la diferencia promedio prácticamente a la mitad.

Finalmente, si bien cabe la duda respecto de para qué períodos la proyección del error permitirá hacer una mejor proyección de la variación de la actividad industrial, en el peor de los casos, esta estimación de las diferencias no deteriorará la aproximación del método 3.

V. CONCLUSIONES

Una vez revisados los resultados, se puede concluir que:

- Es posible obtener una aproximación razonable de la evolución de la actividad industrial con la información disponible. Las alternativas de cálculo descritas se correlacionan con la variación del índice efectivo, en promedio, en 0.96; las diferencias entre una tasa y otra se mueven en torno al 1% con una desviación estándar de 1% y, en promedio para el 90% de la muestra, no son superiores a 2.6%. Esto último resulta útil a la hora de analizar las variaciones, ya que si bien un error de 2.5% puede acarrear cambios de signo, en el escenario probable no alterará la tendencia.
- Aunque las diferencias no son exageradas, pueden incluso tener signo distinto cuando son muy bajas. En términos agregados, las tasas arrojan diferencias debido a que difieren las ponderaciones, la apertura y el tratamiento que se da a actividades como aceites y pesca.
- Las diferencias remanentes entre las variaciones de la actividad industrial utilizada en el cálculo del Imacec y las que se pueden obtener con alguno de los métodos alternativos responden principalmente a i) el tratamiento que se da a la línea de alimentos y bebidas, la que se introduce desagregada a cuatro dígitos y con participaciones de la MIP 1996, en vez

de las ponderaciones usadas por el INE y ii) información adicional exógena al IPVF³ en las líneas de aceites y pesca. Aunque no es posible

entregar esta información detallada a los analistas, ello no les impide realizar estimaciones bien comportadas.

³ Si bien para los últimos dos años la incidencia promedio de considerar información exógena al índice del INE ha sido menos de 0.2 punto porcentual, la mayor de estas estuvo en torno a 1 punto en ese período. Ver gráfico A2 en el apéndice.

APÉNDICE

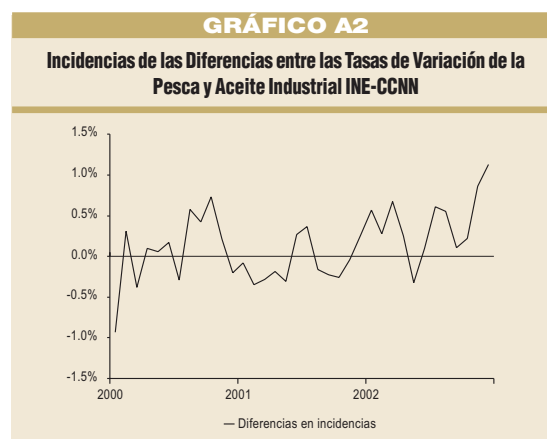
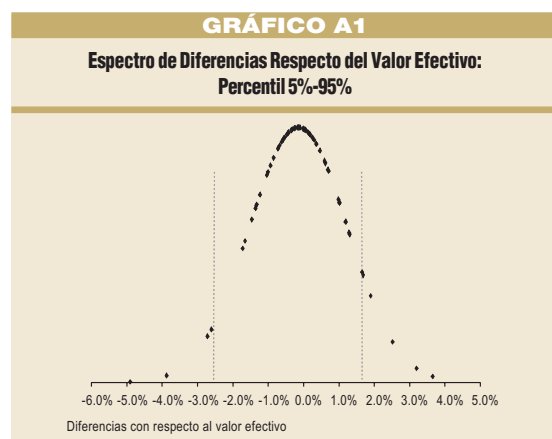
ANÁLISIS ESTADÍSTICO DE RESULTADOS

Con el propósito de analizar la bondad de ajuste de cada uno de estos tres métodos, se construyen las diferencias de las tasas de variación resultantes de estos con respecto al valor efectivo. Las principales estadísticas de cada uno de los escenarios antes mencionados se presentan en el cuadro A1.

El cuadro muestra que la mayor diferencia (en valor absoluto) entre el valor efectivo y el valor estimado corresponde al método 2 (5,3%). Las diferencias mínimas son prácticamente nulas para los tres métodos. En cuanto a los promedios, se desprende que en términos generales no son muy diferentes de

1% en cada método. La mediana nos muestra que la mitad de las diferencias absolutas no supera el 1%. Por su parte, las desviaciones estándar fluctúan en torno a 1%.

Por otro lado, para analizar las correlaciones que existen entre ambos métodos, se utilizan las diferencias efectivas (al igual que para el cálculo de los percentiles). Las correlaciones son prácticamente las mismas para los tres métodos, aunque para el método 3 arroja el mejor resultado, con un ajuste de 0,97. De los percentiles se puede extraer que el método 1 está centrado, en cambio los métodos 2 y 3 muestran un sesgo negativo (el método, en promedio, subestima la tasa efectiva). El gráfico A1 (distribución de las diferencias de los valores obtenidos a través del método 3) ilustra este sesgo.



CUADRO A1

Diferencias con Respecto a la Tasa Efectiva

	Método 1 INE	Método 2 INE CCNN (fija)	Método 3 INE CCNN (Lasp.)	Método 4 INE CCNN (Lasp. - Act.)
Max*	4.6%	5.6%	4.9%	4.8%
Min*	0.1%	0.0%	0.0%	0.0%
Promedio*	1.2%	1.1%	0.9%	1.0%
Mediana*	1.0%	0.8%	0.6%	0.7%
D.S.*	0.9%	1.1%	0.9%	1.0%
Correlación	0.95	0.96	0.96	0.97
Percentil 5	-2.5%	-2.7%	-2.5%	-2.6%
Percentil 10	-1.8%	-1.5%	-1.4%	-1.5%
Percentil 25	-0.9%	-0.9%	-0.6%	-1.0%
Percentil 75	1.1%	0.7%	0.6%	0.2%
Percentil 90	2.0%	1.3%	1.3%	1.0%
Percentil 95	2.2%	2.1%	1.7%	1.3%

Fuente: elaboración propia basada en información pública.

* diferencias en valor absoluto