

## **Ponderaciones ENGHo-2004/05 e IPC-GBA base abril 2008=100**

Como es sabido, una de las principales fuentes para el cálculo de las ponderaciones de un Índice de Precios al Consumidor (IPC) es la Encuesta Nacional de Gastos de los Hogares (ENGHo). Dicha encuesta tiene como fin proporcionar información sobre las condiciones de vida de la población en general y de grupos de hogares en particular, desde el punto de vista de su participación en la distribución y en el consumo de los bienes y servicios que produce la sociedad.

Al tiempo, uno de los mayores problemas relativos a estas encuestas es la subdeclaración de los ingresos y de los gastos (en particular, los gastos de consumo). La ENGHo-2004/2005 de la República Argentina ha presentado, como otras realizadas en período anteriores, las limitaciones señaladas.

Con el objeto de asegurar una mayor correspondencia entre los resultados obtenidos en la ENGHo-2004/2005 y la real estructura de ingresos y gastos de los hogares –corregida por subdeclaración y subestimación- y las ponderaciones a ser aplicada en el IPC-GBA base abril 2008=100, se desarrolló una metodología consistente en los siguientes pasos:

- En primer lugar, se partió de la estructura de gastos de consumo obtenida de la ENGHo-2004/05. La clasificación de bienes y servicios para construir la estructura se elaboró según los criterios de la Clasificación de Consumo Individual por Finalidades (COICOP<sup>1</sup>, por sus siglas en inglés).
- En segundo lugar, se corrigió la estructura observada en la ENGHo-2004/05 utilizando el procedimiento descrito en el trabajo “Un método para corregir la subdeclaración de ingresos y gastos de consumo a partir de la Encuesta Nacional de Gastos de los Hogares 2004-2005”<sup>2</sup>, que se puede encontrar

---

<sup>1</sup> Naciones Unidas (2001): Clasificaciones de gastos por finalidades, Serie M N° 84, División de Estadística, Departamento asuntos económicos y sociales, Naciones Unidas, Nueva York. <http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5>

<sup>2</sup> “Un método para corregir la subdeclaración de ingresos y gastos de consumo a partir de la Encuesta Nacional de Gastos de los Hogares 2004-2005”, JEL:E01,E21, Presentado en la Reunión Anual de la Asociación de Economía Política, Mendoza, 2009

como Anexo 1. El método aplicado corrigió la subdeclaración del ingreso detectado en la ENGHo-2004/2005 a partir de fuentes externas, teniendo en cuenta la categoría ocupacional y la región geográfica de pertenencia de los hogares, y se asignó un vector de consumo al monto corregido de ingreso. Para este fin, fue necesario utilizar la Clasificación Central de Productos (CPC<sup>3</sup>). Mediante métodos econométricos, se corrigieron los sesgos en la declaración del ingreso que surgieron de la comparación entre los datos de la Encuesta y la información de salarios registrados en el Sistema Nacional de Jubilaciones y Pensiones (SIJP, actual SIPA) para el año 2004. Así como la subdeclaración del ingreso suele estar asociada a los controles tributarios y esto impacta de manera más fuerte en los percentiles de ingreso más altos, la subdeclaración de gastos está comúnmente vinculada al olvido en el registro de gastos, fundamentalmente en los períodos de referencia más amplios, así como también a la omisión deliberada de artículos sobre los cuales pueda pesar algún tipo de sanción legal o social. Asimismo, puede existir la voluntad de mantener cierta proporción entre el nivel declarado de gastos y el monto de ingresos subdeclarados en la encuesta. Entonces, al gasto que declara el hogar en la Encuesta se le sumará el gasto corregido por subdeclaración, teniendo en cuenta la estructura del consumo del ventil de ingreso al que pertenece luego de la corrección.

- En tercer lugar, dado que el IPC utiliza la COICOP para clasificar los bienes y servicios que se encuentran en su canasta, fue necesario recodificar y reponderar los artículos. De este apareo surgió una nueva estructura de gastos de consumo. Como ejemplos se pueden citar los siguientes casos: el agrupado clasificado en CPC como *“Aparatos ortopédicos; tablillas y otros aparatos y artículos para fracturas; partes artificiales del cuerpo; audífonos y otros aparatos que se llevan o implantan en el cuerpo para compensar un defecto o una incapacidad”* no tiene correspondencia en COICOP por lo cual se desestimó su gasto; el agrupado CPC *“Grifos, llaves, válvulas y accesorios análogos para tuberías, calderas, tanques, cubas y recipientes análogos”* se apareó con su correspondiente *“Juego de grifería”* en COICOP.

---

<sup>3</sup> Naciones Unidas (2002): Clasificación Central de Productos (CPC), Serie M N° 77 ver. 1.1, División de Estadística, Departamento asuntos económicos y sociales, Naciones Unidas, Nueva York.

- Finalmente, ya que el cálculo del IPC utiliza una fórmula de Laspeyres, (lo que implica que el año de referencia de las ponderaciones debe ser el mismo que el año que se utiliza como base para las comparaciones de precios) fue necesario actualizar las ponderaciones corregidas del aglomerado Gran Buenos Aires mediante la evolución de sus índices de precios utilizando el IPC-GBA base 1999=100. La actualización de las ponderaciones se realizó utilizando la siguiente fórmula:

$$\omega_i = \frac{I_i^{o,b} \omega_i^{ENGHo-2004/05}}{\sum_i I_i^{o,b} \omega_i^{ENGHo-2004/5}}$$

donde  $I$  es un relativo entre índice de precios base 1999=100,  $b$  se refiere al período de referencia de la ENGHo-2004/05,  $o$  se refiere al mes abril 2008 y  $\omega^{ENGHo-2004/05}$  es la ponderación ya corregida de la ENGHo.

Los resultados se presentan en la siguiente tabla:

Capítulo	ENGHo 2004/05 (COICOP)	ENGHo-2004/05 corregida por subdeclaración (CPC)	ENGHo 2004/05 apareada con IPC (COICOP)	Ponderación IPC base abril 2008=100 *
1- Alimentos y bebidas	31,5	31,4	36,9	37,9
2- Indumentaria	7,1	7,2	6,9	7,3
3- Vivienda y servicios básicos	10,5	10,1	11,9	12,1
4- Equipamiento y mantenimiento del hogar	7,1	7,5	4,9	4,9
5- Atención medica y gastos para la salud	8,2	6,1	5,8	5,6
6- Transportes y telecomunicaciones	16,6	17,3	18,3	16,6
7- Esparcimiento	8,9	6,8	5,3	5,1
8- Educación	3,9	4,4	3,7	4,3
9- Otros bienes y servicios	6,3	9,2	6,2	6,3

\* Actualización de las ponderaciones corregidas de la ENGHo-2004/2005 del aglomerado Gran Buenos Aires mediante la evolución de sus índices de precios utilizando el IPC-GBA base 1999=100.

## **Anexo I**

**REUNIÓN ANUAL DE LA  
ASOCIACIÓN ARGENTINA DE ECONOMÍA POLÍTICA  
MENDOZA 2009**

**UN MÉTODO PARA CORREGIR LA SUBDECLARACIÓN DE INGRESOS Y GASTO DE CONSUMO A  
PARTIR DE LA ENCUESTA NACIONAL DE GASTO DE LOS HOGARES 2004-2005**

## **RESUMEN**

Se expone un método para corregir la subestimación de consumos declarados en la Encuesta Nacional de Gasto de los Hogares 2004-2005 desarrollando el utilizado para la versión anterior de la encuesta.

## **ABSTRACT**

***JEL: E01, E21***

## Introducción

Las encuestas a hogares suelen tener problemas en la captación de ingresos y gastos. Es sabido entre los especialistas que los resultados de estas variables están generalmente sujetos a una subestimación producto de errores muestrales y no muestrales. Entre los primeros se pueden mencionar los problemas de captación de hogares, lo cual depende básicamente del diseño de la muestra; un ejemplo de los segundos es el caso de hogares que ocultan información. Por tales motivos se vuelve necesario estimar la subdeclaración en la información para producir estadísticas más cercanas a la realidad.

Existen, en principio, dos tipos de metodologías para la corrección de datos en encuestas a hogares: los métodos estadísticos y los métodos contables. Los primeros procuran estimar los datos faltantes a partir del reconocimiento de patrones inmanentes a los datos de las encuestas, o bien en relación con datos de registro, mientras que los segundos comparan la información a nivel agregado con terceras fuentes, verificando su consistencia, como se da con las Cuentas Nacionales. En nuestro medio, un ejemplo del primer tipo es el trabajo de Sosa y Alaimo (2000); del segundo, el de Camelo (1998). Ambas estrategias no constituyen alternativas, sino que se complementan. Si la información aportada por las encuestas a hogares adolece de fallas producto de errores no muestrales, éstos no pueden ser corregidos en su totalidad apelando a los datos recolectados, pues la información faltante puede presentar un sesgo sistemático. Por lo tanto, para corregirla será preciso contar con datos externos. Del otro lado, sin embargo, corregir la información de ingresos proveniente de encuestas apelando exclusivamente a las Cuentas Nacionales tampoco ofrece un resultado óptimo, dado que dichas encuestas son a su vez un insumo básico para el cálculo de las Cuentas Nacionales, de manera que el eventual sesgo en el origen podría estar trasladándose al final.<sup>4</sup>

Parece conveniente, por tanto, estimar en primer lugar la subdeclaración con la información proveniente de las encuestas u otras bases de datos (como las que surgen de los registros tributarios) mediante métodos econométricos, para en un segundo paso conciliarla con el resto de las Cuentas Nacionales. La propuesta presentada en este trabajo se limita a corregir los ingresos y gastos utilizando el primer método mencionado, dejando el segundo tipo de corrección para una instancia ulterior. Se expone un método de corrección haciendo uso solamente de los datos de la Encuesta Nacional de Gasto de los Hogares 2004-2005 (ENGH) y de la información de salarios registrados en el Sistema Nacional de Jubilaciones y Pensiones (SIJP) para el año 2004. No se incorporan, por tanto, correcciones que surgen de la comparación con otras fuentes, como las bases de bienes registrables o aquellas que es menester realizar al comparar el vector de consumo con la matriz de oferta de bienes en la economía. Por lo demás, el objetivo es proponer un método y no ofrecer estimaciones, dado que a la fecha la ENGH no ha sido publicada aún.

## Fuentes de la subdeclaración

Para realizar una estimación del ingreso o del consumo de los hogares es preciso reconocer las fuentes de esta subestimación, de manera de desarrollar una estrategia que pueda estimar, en segundo lugar, su magnitud.

La forma como este sesgo se manifiesta puede provenir de tres situaciones distintas: a) hogares que no contestan la encuesta; b) hogares que no contestan la pregunta relativa al ingreso; y c) hogares o individuos que declaran un ingreso inferior

---

<sup>4</sup> Camelo (1998) aclara esta salvedad en sus estimaciones.

al que realmente ganan. El segundo caso es resuelto por la oficina responsable de la elaboración de la ENGH mediante un proceso de imputación de ingresos. El presente trabajo se concentrará, por tanto, en las dos fuentes de error restantes.

Las familias suelen asociar la información sobre sus ingresos con los controles tributarios, de manera que creen contar con motivos para declarar menos ingresos que los que efectivamente ganan. Esta situación afecta a los hogares de ingresos más altos y cuyos ingresos no son de fuente asalariada registrada (Minardi 2002). En contraste, los asalariados registrados no tienen motivos para ocultar información, pues informan su salario de bolsillo ya tiene descontados los impuestos.

Por su parte, en el caso del consumo, los errores no muestrales se deben a otros motivos: a) olvido o error en el registro de los consumos realizados; b) omisión deliberada, producto de que pese algún tipo de sanción legal o social sobre el consumo del artículo en cuestión; c) voluntad de mantener cierta proporción entre el nivel declarado de gastos y el monto de ingresos subdeclarados en la encuesta.

### **Estimación de la subcaptación de ingresos en la ENGH 2004-2005**

Como fue mencionado, existen tres fuentes posibles de subdeclaración de ingresos. De ellas, la ausencia de respuesta de la pregunta de ingresos por parte de hogares encuestados es estimada por la Dirección de Estudios de Ingresos y Gastos de los Hogares del INDEC (encargada de la elaboración de la ENGH), mediante un procedimiento de imputación tipo *hot-deck*, de manera que no será objeto de reestimación en este trabajo. Restan, por tanto, dos fuentes de sesgo: a) la subdeclaración propiamente dicha, y b) el rechazo a responder la encuesta en su totalidad.

Partiendo, por tanto, de una base de datos que contaba ya con una imputación de ingresos en aquellos hogares que no habían contestado la pregunta respectiva, fue corregido, en primer lugar, el sesgo en la distribución del ingreso producido por el hecho de que algunos hogares rechacen responder la encuesta; en segundo lugar, se realizó una corrección por subdeclaración.

#### *Corrección de la subestimación de ingresos por falta de respuesta a la encuesta*

Existen razones para considerar que la falta de respuesta total a la encuesta es más importante en los hogares de ingresos más elevados, de manera que ello estaría generando un sesgo hacia abajo en los valores de ingresos medios estimados. Por lo tanto, una comparación de la información procedente de la ENGH con información de registro debería reflejar dicho sesgo en los percentiles superiores de la distribución. Por tal motivo se compararon los salarios medios por percentil de los asalariados registrados de la ENGH<sup>5</sup> con los correspondientes al SIJP (exceptuando el sector agropecuario, los trabajadores estatales y el servicio doméstico, mal representados en esta última fuente), por región. De allí surgen relaciones que fueron aplicadas a las categorías ocupacionales restantes, siguiendo el supuesto de que la subdeclaración de los asalariados registrados es la mínima. Los resultados muestran valores superiores a 1 en los percentiles superiores (desde p85 en adelante, según la región). En los percentiles cuyo salario medio era mayor en la ENGH que en el SIJP fue mantenido el dato original de la ENGH, por considerarse la declaración del individuo como valor mínimo. El resultado de esta corrección en la distribución del ingreso puede verse en el cuadro expuesto a continuación.

Si se tiene en cuenta que en cada hogar puede haber ingresos de fuentes diversas, la corrección final al ingreso total del hogar dependerá de la participación de

---

<sup>5</sup> Se incluyó la primera ocupación asalariada registrada más los casos en que la segunda ocupación era de estas características, mientras que la primera fuera no asalariada o no registrada.

las fuentes corregidas en el total y del percentil al que pertenecen el/los ingresos individual/es registrados.<sup>6</sup>

**Impacto de la corrección del sesgo por no encuesta en el ingreso total, por veintiles de ingreso per capita y región**

Veintil de ingresos	GBA	Región Pampeana	NOA	NEA	Cuyo	Patagonia	Total país
1	0%	0%	0%	0%	0%	0%	0%
2	0%	0%	0%	0%	0%	0%	0%
3	0%	0%	0%	0%	0%	0%	0%
4	0%	0%	0%	0%	0%	0%	0%
5	0%	0%	0%	0%	0%	0%	0%
6	0%	0%	0%	0%	0%	0%	0%
7	0%	0%	0%	0%	0%	0%	0%
8	0%	0%	0%	0%	0%	0%	0%
9	0%	1%	0%	0%	0%	0%	0%
10	0%	0%	0%	0%	0%	0%	0%
11	0%	1%	0%	0%	0%	0%	0%
12	1%	1%	0%	0%	0%	0%	1%
13	1%	2%	0%	0%	0%	0%	1%
14	2%	2%	0%	0%	0%	0%	1%
15	2%	2%	0%	0%	1%	0%	1%
16	2%	3%	0%	0%	1%	0%	2%
17	5%	4%	0%	0%	1%	0%	4%
18	7%	7%	1%	0%	1%	1%	5%
19	15%	13%	2%	1%	5%	2%	11%
20	33%	29%	5%	16%	18%	6%	26%
Total	9%	8%	1%	4%	4%	1%	7%

El segundo paso, por tanto, fue estimar la subdeclaración diferencial de las ocupaciones no registradas (básicamente, patrones y trabajadores autónomos), que se suma a la corrección realizada en el paso anterior. El método se expone en el apartado siguiente.

*Modelo de corrección de ingresos por subdeclaración*

Tomando la corrección anterior como básica para todos los ingresos laborales, se estimó en segundo lugar un modelo similar al realizado para calcular la Cuenta de Generación del Ingreso base 1993. Dicho modelo parte de comparar el gasto en alimentos y los ingresos declarados de dos grupos, a saber: asalariados registrados y un grupo definido como subdeclarante, bajo el supuesto de que el rubro “gasto en alimentos” no es objeto de subdeclaración por ninguno de ellos. De tal manera, para hogares de similares características y nivel de gasto en dicho rubro, la elasticidad gasto-ingreso debería ser la misma. Sin embargo, esto no ocurre así. De tal manera, será posible calcular para el grupo subdeclarante el factor que disminuye el valor real de la elasticidad y obtener el porcentaje de subdeclaración del grupo (Sosa Escudero y Alaimo 2000). Por lo tanto, el esquema supone que la subdeclaración de los hogares en los que predomina el ingreso de fuente asalariada registrada o jubilación es mínima, y calcula la diferencia de las restantes categorías con dicho grupo.

Formalmente:

<sup>6</sup> Las otras fuentes de ingreso son ocupaciones anteriores, jubilaciones, rentas y transferencias.

$$E(G^{al}) = \alpha + \beta \bar{Y} + \delta Z \quad (1)$$

$$E(G_{AR}^{al_i}) = E(G_S^{al_i}) \quad (2)$$

donde  $G^{al}$  es el gasto en alimentos,  $Y$  es el ingreso del hogar y  $Z$  un vector de características del hogar. Por su parte, los subíndices  $AR$  y  $S$  refieren a los asalariados registrados y al grupo subdeclarante, respectivamente.

Ahora:

$$Y^* = Y - S \quad (3)$$

donde  $Y^*$  es el ingreso reportado y  $S$  la tasa de subdeclaración que es nuestra incógnita.

De tal forma, el consumo autónomo ( $\alpha$ ) de un modelo que no discrimina si hay grupos subdeclarantes incluye una componente que corresponde a la subdeclaración. En otros términos:

$$\alpha_0 + \beta_0 \bar{Y} = (\alpha_s + \beta_s S) + \beta_0 \bar{Y}^* = \alpha^* + \beta_0 \bar{Y}^* \quad (4)$$

De (1), (2) y (3) se obtiene

$$\alpha_s = \alpha_0 + (Y^* + S)(\beta_0 - \beta_s) \quad (5)$$

Entonces, dado que de (4) sabemos que  $\alpha^* = \alpha_s + \beta_s S$ , reemplazando  $\alpha_s$  en (5) y despejando  $S$  obtenemos:

$$S = \frac{(\alpha_0 - \alpha^*) + (\beta_s - \beta_0)Y^*}{\beta_0} \quad (6)$$

con lo que se alcanza una expresión de  $S$  en términos de magnitudes observables (Sosa y Alaimo 2000).

Como se podrá advertir, el problema que surge es que la categoría ocupacional es un rasgo de los individuos, pero la unidad de gasto es el hogar. Por lo tanto, para asociar ambas variables se asignó al hogar la categoría ocupacional cuyos ingresos fueran predominantes en el total.<sup>7</sup>

La estructura formal del modelo aplicado a las distintas regiones del país es la siguiente:

$$G^{al} = \alpha_0 + \beta_0 Y + \gamma_i X_i + \phi_i Y X_i + \delta Z$$

donde:

$$\gamma = \alpha_0 - \alpha^*$$

$$\phi = \beta_s - \beta_0$$

$G^{al}$  = logaritmo del gasto en alimentos

$Y$  = logaritmo del ingreso del hogar

$i$  = subíndice que indica las categorías ocupacionales

<sup>7</sup> La metodología utilizada por Felcman, Kidyba y Ruffo (2003) tiene algunas diferencias: distingue sólo entre GBA y total de aglomerados del interior; calcula la subdeclaración para un solo grupo considerado subdeclarante, conformado por los patrones y los autónomos que realizaran tareas técnicas o profesionales; finalmente, se consideró autónomo al hogar cuyos ingresos provenientes de dicha fuente superaran el 25% del total. Además, el modelo se calculó eliminando los hogares cuya única fuente de ingresos fuera la jubilación y aquellos con jefes de hogar mayor a 65 años (Felcman, Kidyba y Ruffo, 2003).

X= matriz de variables binarias referidas a las categorías ocupacionales (*i*)  
 Z= matriz de variables que capturan características del hogar (cantidad de adultos equivalentes, nivel de instrucción del jefe).<sup>8</sup>

Además de la especificación aquí expuesta también fue probado un modelo que discriminaba la subdeclaración por quintiles de ingreso, pero los coeficientes de las variables desagregados por quintiles no resultaron estadísticamente significativos. De ello se sigue que la tasa de subdeclaración no depende del nivel de ingresos, sino principalmente de su fuente, de manera que la subdeclaración por quintiles de ingreso varía en función de la proporción que representa cada fuente en el total del quintil.<sup>9</sup> Asimismo, fueron probadas otras variables, como edad o sexo del jefe, pero la pérdida de grados de libertad afectaba al comportamiento del modelo global (medida por el estadístico F) sin mejoras significativas en la bondad del ajuste.

Desde el punto de vista de la técnica econométrica, los modelos fueron calculados en tres etapas: en primer lugar, se corrió la regresión robusta para obtener nuevos ponderadores, de manera de reducir el efecto de los casos extremos; luego, se corrió la regresión por mínimos cuadrados con dichos ponderadores. Finalmente, como los residuos presentaban heterocedasticidad, se hizo lo propio con una regresión ponderada por los coeficientes de la segunda regresión, obteniendo de esta manera resultados eficientes.

Los modelos y sus resultados son expuestos en el anexo. Los valores de los coeficientes de subdeclaración obtenidos son los siguientes:

#### Coefficientes de subdeclaración por categoría ocupacional y región

	Total País	GBA	Interior	Pampeana	NOA	NEA	NEA+NOA +Cuyo	Cuyo	Patagonia
P	43,2 (*)	33,6 (*)	48,7 (*)	51,8 (*)	12,2	80,0 (*)	36,7 (*)	34,1 (*)	9,3
CP	18,8 (*)	17,4 (*)	19,1	15,7	14,4	28,1	18,8	14,1	36,8
ANR	2,5 (*)	2,0 (*)	-0,1 (*)	-6,7 (*)	-0,6 (*)	7,1 (*)	4,3 (*)	8,4 (*)	7,7
Casos	28.578	6.661	21.942	7.496	5.562	2.845	11.227	2.818	3.233

(\*) indica un valor-P < 0,05 para ambos coeficientes ( $\alpha$  y  $\beta$ )

Los modelos para el NOA y la Patagonia no arrojaron resultados válidos. Por tal motivo, para dichos casos se tomaron los coeficientes obtenidos para agregados superiores: a NOA le fueron asignadas las tasas correspondientes al agregado NEA, NOA y Cuyo, mientras que a Patagonia, las que surgen para el total del interior.

Obtenidos los coeficientes de subdeclaración se los aplicó a la totalidad de los ingresos del hogar, en función de la asociación realizada entre hogar y categoría ocupacional predominante, según fue explicado más arriba. El resultado de la corrección por veintiles de ingreso per capita es el siguiente:

#### Corrección del ingreso total, por veintiles de ingreso per capita y región

Veintil de ingresos	GBA	Región Pampeana	NOA	NEA	Cuyo	Patagonia	Total país
1	6%	7%	7%	19%	10%	4%	7%
2	5%	6%	8%	14%	9%	5%	6%

<sup>8</sup> El nivel de instrucción del jefe fue categorizado en tres niveles: hasta secundaria incompleta, secundaria completa (*jnivinsSeC* en los modelos del anexo) y estudios superiores o universitarios completos (*jnivinsSuC* en los modelos del anexo).

<sup>9</sup> Este hecho aquí corroborado (o, más precisamente, no refutable con base en la evidencia empírica) es asumido como supuesto por Camelo (1998).

3	7%	7%	9%	22%	7%	5%	7%
4	4%	3%	6%	14%	8%	6%	5%
5	6%	6%	7%	28%	6%	4%	7%
6	3%	5%	5%	14%	7%	4%	5%
7	6%	6%	6%	12%	6%	1%	6%
8	4%	9%	5%	15%	5%	5%	6%
9	6%	6%	7%	12%	5%	3%	6%
10	6%	7%	5%	11%	5%	3%	6%
11	7%	10%	6%	14%	7%	5%	8%
12	6%	8%	6%	24%	8%	7%	8%
13	6%	10%	7%	27%	5%	7%	9%
14	8%	11%	5%	10%	4%	5%	8%
15	11%	11%	5%	17%	8%	7%	11%
16	9%	11%	7%	24%	11%	6%	10%
17	15%	22%	7%	25%	9%	11%	17%
18	17%	23%	8%	35%	9%	22%	19%
19	28%	45%	10%	34%	17%	12%	32%
20	59%	98%	28%	194%	46%	39%	77%
Total	20%	31%	11%	59%	16%	14%	25%

La corrección del NEA llama la atención por su magnitud. Al respecto, es menester reconocer en primer lugar que los coeficientes son significativos para el caso de los hogares identificados como patrones y asalariados no registrados. Una mirada a los datos de base indica que en el NEA es menor la proporción de hogares patrones, mayor la de cuentapropistas y menor la de asalariados registrados. Además, la relación del ingreso medio de los hogares patrones y los asalariados registrados es algo superior que en el resto de las regiones, mientras que la relación de consumos declarados es similar. Ello da como resultado que los modelos tengan un coeficiente de consumo autónomo general menor y un coeficiente de consumo autónomo por categoría más alto, lo que arroja una mayor tasa de subdeclaración.<sup>10</sup>

Un resumen del impacto de la corrección puede apreciarse si se calculan los coeficientes de Gini para la distribución del ingreso original y para la corregida:

	GBA	Región Pampeana	NOA	NEA	Cuyo	Patagonia	Total país
Gini original	0,4462	0,4432	0,4193	0,4590	0,4228	0,4552	0,4551
Gini corregido	0,5137	0,5491	0,4461	0,6188	0,4713	0,4997	0,5364

Como se sigue del método empleado, la corrección no incluye la subdeclaración de ingresos proveniente de fuentes diversas al ingreso laboral, como los ingresos de la propiedad (rentas, dividendos, etc.) y otros ingresos de no mercado (transferencias).

### *Corrección del consumo*

La corrección por subdeclaración del consumo se estimó asignando un vector de consumo al ingreso diferencial que surge de la estimación de la subdeclaración de ingresos. Debe tenerse en cuenta, sin embargo, que la ausencia de respuesta a la encuesta por parte de muchos hogares puede introducir sesgos u omisiones en el vector de consumo total, debido a que el proceso de reponderación a que son sometidos los hogares que sí contestaron la encuesta implica extender sus consumos

<sup>10</sup> También debe mencionarse que la provincia de Formosa tiene pocos datos (180), muchos de los cuales han sido imputados. Con todo, los resultados del modelo excluyendo esta provincia reducen el porcentaje de subdeclaración en patrones de 80% al 72%, manteniendo prácticamente inalterada la corrección de cuentapropistas.

a aquellos hogares que no lo hicieron. De tal forma, si existe algún sesgo socioeconómico en la tasa de no respuesta, los consumos asociados a los hogares no respondientes estarán mal representados.

Al aplicar al total del gasto la tasa de subdeclaración de los ingresos surge un nuevo ordenamiento de los hogares. Al gasto sumado a cada hogar se le asignó la estructura del vector de consumo típico del veintil de ingresos al cual el hogar pasó a pertenecer luego de la corrección de ingresos. De tal forma, el consumo corregido de cada hogar estará conformado por: a) la declaración original que surge de la encuesta más b) el gasto adicionado producto de la corrección por subdeclaración, repartido con la estructura del consumo del veintil de ingresos al que pasó a pertenecer tras la corrección de ingresos, según informa la ENGH. Los hogares que originalmente estaban en el veintil 20, como es obvio, sólo vieron modificado el monto de sus consumos tras la corrección.

#### Corrección del gasto total de consumo, por veintiles de ingreso per capita y región

Veintil de ingresos	GBA	Región Pampeana	NOA	NEA	Cuyo	Patagonia	Total país
1	3%	4%	4%	13%	6%	3%	4%
2	4%	5%	6%	10%	7%	5%	5%
3	5%	5%	7%	16%	5%	5%	6%
4	4%	3%	5%	14%	7%	6%	4%
5	5%	5%	6%	26%	5%	4%	6%
6	3%	5%	5%	13%	6%	5%	5%
7	5%	6%	6%	12%	6%	2%	6%
8	4%	8%	5%	15%	4%	6%	6%
9	7%	7%	7%	11%	5%	3%	7%
10	6%	7%	6%	12%	5%	3%	6%
11	7%	11%	6%	16%	7%	7%	9%
12	6%	9%	7%	23%	8%	8%	8%
13	6%	10%	7%	32%	6%	9%	9%
14	8%	12%	5%	11%	5%	7%	9%
15	12%	14%	6%	19%	9%	8%	12%
16	10%	13%	8%	29%	11%	8%	12%
17	16%	27%	7%	26%	9%	15%	19%
18	20%	32%	10%	43%	12%	38%	25%
19	35%	60%	13%	45%	21%	20%	41%
20	89%	197%	42%	385%	86%	86%	134%
Total	22%	37%	12%	69%	18%	19%	29%

Como se observa, la corrección del gasto de consumo es más pronunciada que la observada en los ingresos, cuando se la mira según la estructura veintílica de ingreso per capita.

#### Corrección por autoconsumo

A continuación de la corrección anterior se estimó la porción de bienes consumidos cuyo origen fuera la producción propia, contraprestación por el trabajo realizado o retiro del propio negocio. A tal efecto fue seleccionado un vector de

artículos susceptibles de cumplir con algunos de dichos rasgos.<sup>11</sup> El monto consignado en la base como autoconsumo o retiro del propio negocio por parte de cada hogar fue distribuido proporcionalmente entre dichos bienes de consumo. Dicho monto es finalmente restado del vector de consumo final de los hogares.

### Corrección contable del consumo

Posteriormente se aplicó una corrección ulterior, a partir de las declaraciones juradas (DDJJ) del impuesto a las ganancias Personas Físicas para el año 2005 (IGPF). En la discriminación de los ingresos declarados por tramos para dicho impuesto existen niveles de ingreso que no aparecen en la distribución de la ENGH corregida. En este sentido, una categorización del ingreso total del hogar según tramos en esta fuente no reconoce casos más allá del millón de pesos anuales. En cambio, las DDJJ del IGPF sí registran ingresos que superan tal magnitud. Se supuso que dicha información corresponde a ingresos no captados por la encuesta, por lo que fueron incluidos para el cálculo del vector de consumo final de los hogares.

Para estimar la propensión a consumir de dichos tramos de ingreso se calculó una regresión log-lineal entre las propensiones medias a consumir por tramo y el ingreso total del hogar mediano de dichos tramos.<sup>12</sup>

Tramo de ingreso anual (miles)	Propensión a consumir
0-10	1.20
10-20	0.97
20-30	0.88
30-40	0.81
40-50	0.76
50-60	0.70
60-80	0.67
80-100	0.56
100-150	0.51
150-200	0.43
200-300	0.40
300-500	0.13
500-1000	0.08

Fuente: Elaboración propia a partir de datos de la ENGH '04-'05

```
.reg lny lnx
```

Source	SS	df	MS			
Model	6.52398501	1	6.52398501	Number of obs =	11	
Residual	.212913814	9	.02365709	F( 1, 9) =	275.77	
				Prob > F =	0.0000	
				R-squared =	0.9684	
				Adj R-squared =	0.9649	
Total	6.73689883	10	.673689883	Root MSE =	.15381	

  

lny	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnx	-.4477297	.0269613	-16.61	0.000	-.5087203	-.3867391
_cons	1.318603	.128517	10.26	0.000	1.027878	1.609329

<sup>11</sup> Básicamente, productos de huerta, carnes, leche, queso y algunos servicios asociados al mantenimiento del hogar (pintura, carpintería, etc.) o de artículos del hogar (artefactos eléctricos, computadoras). En cuanto a los bienes retirados del propio negocio fueron incluidos los artículos de almacén, de granja, carnicería, pescadería, transporte (taxi, remis) o contables.

<sup>12</sup> Para forzar una convergencia de la propensión marginal en torno a un valor del 5% se incorporó un tramo ficticio de 10 millones de pesos anuales. Por otra parte, fueron eliminados los datos correspondientes a los tramos de ingreso 300-500 y 500-1000, por contar con tan sólo 6 y 2 casos, respectivamente.

Del modelo surgen las siguientes propensiones medias a consumir para los tramos de ingreso a los que se les debe asignar consumo:

Tramo de ingresos	Propensión media a consumir estimada
1000-1500	0.17
1500-2000	0.14
2000-3000	0.12
3000-5000	0.10
5000 y más	0.08

Se atribuyó a dicho gasto de consumo resultante una distribución según el vector de consumo del veintil 20 de la ENGH. Al respecto, debe considerarse que en la corrección por *no encuesta* se comparaban las distribuciones de los salarios registrados de ENGH y del SIJP. Dado que la cantidad de asalariados registrados por el SIJP es superior a la estimación que ofrece la ENGH, resta una corrección de cantidades que incide en la masa salarial total destinada a consumo y que es captada por vía de las DDJJ.

Fueron incluidos los ingresos declarados en concepto de las categorías tercera y cuarta. La tercera categoría corresponde a los ingresos que tienen como fuente beneficios de las empresas. Como se están considerando las DDJJ de personas físicas, la categoría incluye sólo a los patrones de empresas y trabajadores autónomos. Como es sabido, la cuarta categoría está reservada a los salarios que superan el mínimo no imponible.

Esta corrección suma poco más de unos 3 mil millones de pesos (1,4%) a la estimación anterior del consumo final de los hogares.

## Conclusiones

En el presente trabajo se ha expresado una metodología para corregir la subdeclaración de ingresos y gastos de los hogares en la ENGH '04-'05. La ventaja del método utilizado es que parte de unos supuestos mínimos inevitables en función de la información con que se cuenta y no hace uso de supuestos *ad hoc*.<sup>13</sup> Las correcciones al consumo se realizan asignando un vector de consumo al monto corregido del ingreso. Éste fue, por su parte, objeto de una doble corrección: en primer lugar, se comparó la distribución de ingresos asalariados registrados de la ENGH con la que surge de la AFIP, de donde se obtuvieron correcciones para los percentiles de ingreso superiores; en segundo lugar, se estimó mediante un modelo econométrico la tasa de subdeclaración según la categoría ocupacional predominante en los ingresos del hogar.

Además de haberse discriminado el grupo subdeclarante por categoría ocupacional, la estimación de las tasas de subdeclaración fue realizada por regiones geográficas.

Por otra parte, se agregó gasto de consumo para los tramos de ingreso que no son captados por la encuesta, pero que sí aparecen en las declaraciones juradas de impuesto a las ganancias personas físicas.

Finalmente, no fue asignada al consumo proporción alguna de los ingresos por rentas, ni en la corrección realizada con base en la encuesta, ni incorporando las

<sup>13</sup> En la metodología del cálculo de gasto de consumo para la Matriz Insumo-Producto 1997 se suponía, por ejemplo, que cada veintil de ingresos incrementaba la subdeclaración en un 5% respecto del precedente. Tal supuesto no sólo no se verifica en los resultados de nuestro modelo, sino que implicaría una importantísima subestimación de los resultados.

declaraciones de las categorías primera y segunda (renta de la tierra y de inversiones financieras) del impuesto a las ganancias.

### **Bibliografía**

- AFIP (2006), *Anuario Estadísticas Tributarias 2006*, disponible en el sitio oficial de la AFIP [<http://www.afip.gov.ar>].
- Camelo (1998), "Subdeclaración de ingresos medios en las encuestas de hogares, según quintiles de hogares y fuente del ingreso", *2° Taller MECOVI*, Buenos Aires
- Sosa Escudero, W. y V. Alaimo (2000), "La economía oculta en la Argentina: evidencia basada en Encuestas de Gasto", en AAVV, *La economía oculta en la Argentina*, FIEL, 57-69.
- Felcman, D., S. Kidyba y H. Ruffo (2003), "Medición del ingreso laboral: Ajustes de los datos de la Encuesta Permanente de Hogares para el análisis de la distribución del ingreso", *Anales de la Asociación Argentina de Economía Política*.
- Manzano, G. y R. Maurizio (2004), "Informe final. Vector de consumo de la MIP-Córdoba", Documento inédito.
- Minardi, G. (2002), "Informe final proyecto MECOVI: Calidad en encuestas a hogares", mimeo.

## Anexo

Todas las variables (incluyendo el término constante) están ponderadas por los coeficientes obtenidos en una regresión robusta.

	Total país		Total interior		GBA		Pampeana	
N° de observaciones	28558		21927		6623		7487	
F	(10, 28545)	2556.38	(10, 21920)	1735.15	(10, 6612)	688.07	(10, 7476)	573.02
Prob > F	0.000		0.000		0.000		0.000	
In_gsalim	Coef.	P>t	Coef.	P>t	Coef.	P>t	Coef.	P>t
In_ingtot	0.503	0.000	0.472	0.000	0.517	0.000	0.452	0.000
sum_adeq	0.111	0.000	0.115	0.000	0.123	0.000	0.134	0.000
hog_P	1.372	0.000	1.519	0.000	0.842	0.030	1.662	0.000
hog_CP	0.275	0.000	0.068	0.351	0.658	0.000	0.119	0.366
hog_ANR	0.277	0.000	0.295	0.000	0.366	0.009	0.319	0.014
In_ingP	-0.168	0.000	-0.190	0.000	-0.094	0.056	-0.209	0.000
In_ingCP	-0.026	0.006	0.003	0.772	-0.080	0.000	-0.007	0.716
In_ingANR	-0.039	0.000	-0.044	0.000	-0.050	0.016	-0.051	0.010
jnivinsSeC	0.097	0.000	0.083	0.000	0.119	0.000	0.126	0.000
jnivinsSuC	0.120	0.000	0.092	0.000	0.147	0.000	0.120	0.000
c	1.955	0.000	2.122	0.000	1.913	0.000	2.245	0.000
R2 aj. modelo sin ponderar:	0.4722		0.4416		0.5092		0.4331	

	NOA		NEA		CUYO		PATAGONIA		NEA + NOA + Cuyo	
N° de observaciones	5558		2845		2817		3233		11220	
F	(10, 5547)	548.3	(10, 2834)	213.01	(10, 2806)	260.18	(10, 3222)	216.32	(10, 11211)	1000.83
Prob > F	0.000		0.000		0.000		0.000		0.000	
In_gsalim	Coef.	P>t	Coef.	P>t	Coef.	P>t	Coef.	P>t	Coef.	P>t
In_ingtot	0.514	0.000	0.570	0.000	0.550	0.000	0.370	0.000	0.537	0.000
sum_adeq	0.106	0.000	0.080	0.000	0.107	0.000	0.133	0.000	0.100	0.000
hog_P	0.319	0.316	3.079	0.000	2.039	0.000	-0.469	0.516	1.538	0.000
hog_CP	0.218	0.121	0.307	0.127	0.351	0.064	0.010	0.969	0.239	0.014
hog_ANR	0.591	0.000	0.830	0.000	0.711	0.000	-1.078	0.000	0.654	0.000
In_ingP	-0.039	0.372	-0.404	0.000	-0.275	0.000	0.070	0.439	-0.203	0.000
In_ingCP	-0.022	0.318	-0.023	0.477	-0.041	0.163	0.018	0.631	-0.021	0.169
In_ingANR	-0.090	0.000	-0.122	0.000	-0.099	0.001	0.154	0.000	-0.096	0.000
jnivinsSeC	0.016	0.319	0.006	0.823	0.033	0.145	0.052	0.064	0.021	0.085
jnivinsSuC	0.037	0.316	-0.040	0.487	0.051	0.270	0.163	0.004	0.025	0.346
c	1.913	0.000	1.525	0.000	1.596	0.000	2.623	0.000	1.730	0.000
R2 aj. modelo sin ponderar:	0.4961		0.4271		0.4793		0.3998		0.4712	