

## Introducción

Las encuestas a hogares constituyen un instrumento fundamental para la obtención de información que permita dar cuenta de las condiciones de vida de una población. En particular, el ingreso de los hogares es la variable frecuentemente utilizada para la construcción de indicadores que reflejen el bienestar general.

Sin embargo, existe un amplio consenso internacional respecto a los problemas derivados de la captación del ingreso en las encuestas a hogares, que tienden a subestimar el dato real. Los mismos se relacionan con problemas tanto relativos al diseño y aspectos técnicos-estadísticos del relevamiento, como con problemas de otra índole (rechazo a la encuesta, intencionalidad del informante, falta de respuesta, etc.) (Grosskoff, 1998).

En particular, tres problemas distorsionan la medición de los ingresos. El primero se refiere a la *No Respuesta* de Ingresos, que se presenta cuando los individuos no declaran los ingresos percibidos a pesar de estar ocupados y contar con diferentes fuentes de ingresos. Se ha observado una alta y variable proporción de ingresos no declarados, la cual se advierte más elevada para el caso de los autónomos.

El segundo problema es el de la *Subdeclaración* de Ingresos, donde a diferencia del caso anterior, se lleva a cabo la declaración de ingresos, sin embargo el mismo es declarado por un nivel inferior al realmente percibido. Este es un problema generalizado, aunque puede afectar en mayor medida a los trabajadores autónomos.

Por último, encontramos el problema de la *No Respuesta total (No Encuesta)*, el cual se da cuando los hogares que forman parte de la muestra no contestan la totalidad de las preguntas de la encuesta. Es decir que la muestra prevista no se completa, lo cual altera su aleatoriedad. Hay evidencia de que este problema alcanza a más del 15% de la muestra, aunque este porcentaje difiere en cada onda.

Los problemas de no encuesta y no respuesta pueden sesgar la estimación en el caso de que los individuos que no responden tengan, en promedio, un ingreso diferente al de los que sí lo hacen. Este problema se agrava al incluir la subdeclaración.

Algunos trabajos comparan los datos obtenidos por medio de las encuestas de hogares con otras fuentes de información, para una cuantificación aproximada de este sesgo. Los trabajos de Llach y Montoya (1999), Camelo (1993), Gasparini, Marchionni y Sosa Escudero (2000) y otros, dan cuenta de divergencias entre las Cuentas Nacionales y los datos de ingresos obtenidos en base a la Encuesta Permanente de Hogares (EPH). Para realizar un análisis de pobreza y distribución del ingreso, esta serie de trabajos aplicó diversos métodos de corrección a los datos de la encuesta, en parte basada en correcciones de ponderación, e imputación de ingresos de la encuesta, y en parte por comparación con los resultados de Cuentas Nacionales.

El presente trabajo busca ensayar un método de corrección de ingresos independiente de los resultados de Cuentas Nacionales, basado sólo en la comparación de microdatos. El objetivo es obtener periódicamente una estimación del ingreso de fuente laboral corregido de los posibles sesgos de las encuestas a hogares y que aporte a las propias estimaciones de Cuentas Nacionales.

El análisis se centra en el ingreso de la actividad principal de los ocupados. La fuente utilizada es la Encuesta Permanente de Hogares, un relevamiento periódico realizado en los meses de mayo y octubre de cada año, relevándose solo aglomerados urbanos, los cuales representan el 70% de la población urbana del país. Para estimar el sesgo de subdeclaración y no respuesta total se han utilizado la Encuesta de Gastos de los Hogares de 1996 y el Sistema Integrado de Jubilaciones y Pensiones (SIJP) desde 1994.

Las correcciones se realizaron para todas las ondas entre mayo de 1993 hasta mayo de 2002, para el total de los aglomerados relevados en la EPH.

En la sección 1 se explican los métodos ensayados para la corrección del ingreso asalariado y del ingreso mixto<sup>1</sup>, y se exponen los resultados de las correcciones y la forma en que mejora la disponibilidad de datos de ingreso.

En la sección 2 se analizan los cambios en la distribución del ingreso laboral, realizando comparaciones entre los ingresos originales (reportados a la EPH) y los corregidos, y se construyen distintos indicadores de desigualdad.

Luego de las conclusiones se presenta el anexo metodológico, donde se puede encontrar más detalle sobre los métodos empleados para la solución de los problemas de captación del ingreso.

Finalmente, se presenta un anexo con cuadros adicionales.

### Sección 1. Métodos de Corrección

Siendo la Encuesta Permanente de Hogares (EPH-INDEC) una importante y casi exclusiva fuente de información para estimar el ingreso laboral de todos los trabajadores, y dados los problemas identificados en la captación del ingreso, se han elaborado procedimientos con el objetivo de corregir el ingreso de la ocupación principal de los ocupados. En esta sección se intenta determinar la magnitud de estos problemas, y se explicitan brevemente los métodos utilizados para el ajuste correspondiente<sup>2</sup>.

#### Corrección por No Respuesta de Ingresos

La no declaración del ingreso no es una cuestión menor en las encuestas a hogares. En el Cuadro 1 se puede apreciar la proporción de no respuesta de ingresos (no respuesta parcial) en la EPH, desagregada entre GBA e Interior. En algunas ondas se observa que este problema alcanza a más del 10% de los ocupados, observándose una tendencia creciente en los últimos años. La proporción de no respuesta varía entre ondas y categorías, siendo más habitual entre los empleadores y cuenta propias (en el Cuadro 1 se observa cual es el porcentaje de no respuesta dentro de cada categoría).

**Cuadro 1. Proporción de No Declaración de ingresos en la EPH**

Onda	Total	GBA	Interior	Empleadores	Cuenta Propia	Asalariados Registrados	Asalariados No Registrados
May-93	10,6%	11,9%	8,6%	20,8%	13,5%	5,9%	6,8%
Oct-93	11,2%	12,9%	8,8%	22,6%	14,1%	6,6%	8,2%
May-94	8,5%	8,5%	8,6%	16,1%	11,0%	4,4%	6,3%
Oct-94	9,8%	10,8%	8,5%	23,3%	11,9%	5,6%	7,2%
May-95	9,4%	10,2%	8,3%	22,1%	12,0%	5,0%	7,3%
Oct-95	9,4%	10,6%	7,9%	22,3%	11,7%	4,9%	7,1%
May-96	9,9%	10,6%	9,0%	22,2%	11,9%	4,8%	8,6%
Oct-96	11,7%	12,2%	11,0%	24,2%	15,2%	6,8%	9,1%
May-97	7,7%	7,5%	7,9%	17,0%	8,6%	4,1%	5,3%
Oct-97	9,5%	9,7%	9,3%	24,9%	10,8%	5,9%	6,8%
May-98	8,8%	8,6%	9,1%	18,9%	11,2%	5,5%	6,1%
Oct-98	9,1%	9,4%	8,7%	23,7%	9,6%	5,8%	7,1%
May-99	9,5%	9,6%	9,2%	23,3%	11,4%	6,1%	7,0%
Oct-99	10,5%	11,0%	9,8%	25,1%	13,2%	6,5%	7,8%
May-00	10,1%	10,7%	9,2%	24,6%	13,0%	6,5%	7,5%
Oct-00	9,9%	10,5%	9,2%	27,6%	12,5%	6,4%	6,5%
May-01	10,4%	10,5%	10,2%	22,2%	13,7%	7,0%	7,0%
Oct-01	12,1%	12,8%	11,3%	24,9%	16,0%	8,9%	8,7%
May-02	14,3%	15,6%	12,8%	25,1%	22,4%	9,6%	9,3%

*Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.*

El problema de la no respuesta de ingresos puede afectar las estimaciones en dos sentidos. Por un lado, puede introducir un sesgo si se considera que la no declaración puede estar asociada a características particulares relacionadas con el ingreso. Por ejemplo, la probabilidad de no respuesta en empleadores (los ocupados con mayor ingreso laboral) es entre tres y cuatro veces la de los asalariados. Por otra parte, cuando la proporción de no respuesta varía en el tiempo y entre regiones, puede afectar las comparaciones temporales o regionales.

Para corregir este problema se imputa un ingreso a cada no respuesta parcial. De esta forma, se realiza un modelo econométrico que, en base a características del trabajador, del establecimiento y del puesto de trabajo, estima los determinantes del ingreso laboral de los ocupados que lo declaran. Luego se aplican estos coeficientes a quienes no declaran ingresos para obtener una aproximación de su ingreso real<sup>3</sup>.

Dado que no todas las observaciones presentan las características necesarias para realizar la imputación, persiste un remanente de no respuesta. La cantidad de observaciones corregidas se muestra en el Cuadro 2, como proporción sobre el total de ocupados de la EPH.

En cualquier caso, la proporción de no respuesta se reduce a más de la mitad. En el Cuadro 3 se observa que en todas las ondas la misma ha disminuido más del 50% a partir de la corrección realizada, con una notable mejora en las últimas ondas.

**Cuadro 2. Proporción de correcciones por no respuesta**

Onda	GBA	Interior
May-93	7,26%	4,56%
Oct-93	8,64%	4,86%
May-94	5,16%	4,76%
Oct-94	6,60%	5,76%
May-95	7,49%	4,74%
Oct-95	7,40%	4,49%
May-96	7,26%	5,71%
Oct-96	8,22%	7,21%
May-97	4,63%	4,71%
Oct-97	6,58%	5,88%
May-98	6,18%	5,78%
Oct-98	6,47%	5,47%
May-99	7,24%	5,93%
Oct-99	7,95%	6,37%
May-00	8,19%	6,26%
Oct-00	7,50%	5,45%
May-01	8,24%	6,69%
Oct-01	9,97%	7,95%
May-02	12,52%	8,63%

*Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.*

**Cuadro 3. Porcentaje de mejora en la no respuesta**

Onda	GBA	Interior
May-93	61%	53%
Oct-93	67%	56%
May-94	61%	55%
Oct-94	61%	68%
May-95	74%	57%
Oct-95	69%	57%
May-96	69%	64%
Oct-96	68%	65%
May-97	62%	59%
Oct-97	68%	64%
May-98	72%	63%
Oct-98	69%	63%
May-99	75%	64%
Oct-99	72%	65%
May-00	76%	68%
Oct-00	72%	59%
May-01	78%	66%
Oct-01	78%	71%
May-02	80%	68%

*Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.*

### Corrección por Subdeclaración

El presente método intenta medir y corregir la subdeclaración de ingresos de los no asalariados en las encuestas a hogares a partir del método de discrepancias del gasto. El mismo se basa en interpretar como subdeclaración a las diferencias en la relación de ingreso-gasto que no pueden ser explicadas por características familiares. En otras palabras, se supone que un hogar está subdeclarando ingresos cuando declara mayores gastos que los habituales para su ingreso, de acuerdo a sus características. La comparación se realiza por grupos, uno de los cuales presenta mayores incentivos y posibilidades de subdeclarar. En este grupo se ubica a los autónomos (empleadores y cuenta propistas), dadas las características de su actividad y la mayor variabilidad de sus ingresos. En el otro grupo se ubican los asalariados.<sup>4</sup>

Así, partiendo del supuesto de que ambos grupos declaran correctamente su nivel de consumo<sup>5</sup> y que los asalariados declaran correctamente su ingreso, el modelo compara las declaraciones de gastos e ingresos entre asalariados y autónomos. Si por cada nivel de ingreso declarado (dadas similares características familiares) los no asalariados declaran mayores gastos que los asalariados, se supone que los primeros están subdeclarando el ingreso percibido.

Las estimaciones se basaron en la Encuesta de Gasto de los Hogares, realizada por el INDEC en 1996. Los hogares fueron separados entre asalariados y autónomos (estos últimos son los que perciben más de un 25% de su ingreso por ingreso laboral de ocupaciones autónomas). En esta estimación, el grupo de autónomos se ha reducido a aquellos empleadores o cuenta propistas que realizan tareas profesionales o técnicas. Se estimaron ecuaciones econométricas, en las cuales el logaritmo del gasto (en alimentos) es explicado por el logaritmo del ingreso y otras características del hogar. Para casi todas las especificaciones probadas, se comprobó que, a igual nivel de ingreso, el hogar autónomo consume más que el asalariado y, por lo tanto, se advierte la existencia de subdeclaración. La misma es cuantificada en 32.9% para el GBA y 22.9% para el interior. Con estos coeficientes es posible corregir el ingreso de los trabajadores autónomos de la EPH.

En el Cuadro 4 se exponen la cantidad de observaciones ajustadas por este método en proporción al total de ocupados de la EPH. Hasta el año 1997, el porcentaje de correcciones para el interior del país es marcadamente mayor que para GBA, duplicándolo en las primeras ondas. Esta tendencia se equipara para los años posteriores.

**Cuadro 4. Proporción de correcciones por subdeclaración**

Onda	GBA	Interior
May-93	9,67%	19,95%
Oct-93	9,72%	18,72%
May-94	9,36%	13,19%
Oct-94	8,82%	12,94%
May-95	9,48%	12,19%
Oct-95	9,31%	11,45%
May-96	10,16%	12,17%
Oct-96	9,43%	10,74%
May-97	8,93%	9,55%
Oct-97	9,62%	8,06%
May-98	9,20%	7,82%
Oct-98	9,18%	7,67%
May-99	9,31%	7,66%
Oct-99	8,89%	8,02%
May-00	9,22%	7,72%
Oct-00	8,76%	8,05%
May-01	8,69%	7,55%
Oct-01	8,75%	7,88%
May-02	7,62%	7,27%

Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.

### Corrección por No Respuesta Total

El problema de la no respuesta total (hogares que no responden la encuesta) genera un sesgo si los hogares que no responden difieren del hogar promedio. En particular, las estimaciones de ingresos serán sesgadas si la no respuesta total está ligada al ingreso de las familias o a su patrimonio, siendo este el caso más frecuente.<sup>6</sup>

Dado que el total de la no encuesta abarca entre un 15% y 16% de la EPH (GBA en mayo y octubre de 2001), este problema puede introducir un sesgo significativo en la medición de ingresos.

Para estimar el ajuste se realiza una comparación de un sub-universo de la EPH (ocupados con ocupación principal privada asalariada, con descuento jubilatorio) con un universo específico (puestos declarados en el Sistema Integrado de Jubilaciones y Pensiones, excluyendo el sector público), suponiendo que la diferencia entre ambas fuentes se debe a la no respuesta total.

A fin de realizar una comparación apropiada, la misma se realiza dividiendo la muestra en seis regiones compatibles, eliminando del SIJP las áreas no cubiertas por la EPH. Asimismo, el análisis se realiza sobre la distribución de la variable ingreso de la ocupación principal de los asalariados formales de la EPH, enfrentándolo a un concepto comparable de la distribución del SIJP (la remuneración total neta habitual del puesto principal del CUIL, sin pagos extras ni indemnizaciones, más asignaciones familiares, menos el pago de ganancias).

La comparación realizada refleja una diferencia entre ambas distribuciones, principalmente en los últimos cincuentiles, donde los ingresos declarados al SIJP son significativamente mayores que los de la EPH. Esto daría cuenta de un truncamiento en la EPH en ingresos superiores.<sup>7</sup> Esta información se utiliza para realizar el ajuste de ingresos: se comparan las distribuciones por cincuentiles y se ajustan las declaraciones de los ocupados cuyos ingresos estén en el rango del cincuentil que presenta diferencias, de forma tal que ambas distribuciones se asemejen<sup>8</sup>. En este proceso también se modifican los ingresos de las otras categorías ocupacionales<sup>9</sup>.

En el Cuadro 5 se expone la proporción de observaciones ajustadas por este método en relación al total de ocupados de la EPH. Para el caso de GBA, se observa una tendencia decreciente en esta corrección, mientras que la proporción de la misma se mantiene relativamente constante para el Interior.

**Cuadro 5. Proporción de correcciones por no encuesta**

Onda	GBA	Interior
May-93	16,04%	7,01%
Oct-93	17,24%	7,52%
May-94	8,71%	2,22%
Oct-94	9,57%	3,23%
May-95	18,08%	6,02%
Oct-95	16,89%	5,68%
May-96	17,38%	9,56%
Oct-96	16,97%	8,96%
May-97	16,91%	8,03%
Oct-97	17,08%	7,98%
May-98	9,01%	7,17%
Oct-98	9,53%	7,33%
May-99	7,94%	6,24%
Oct-99	7,94%	5,99%
May-00	5,08%	7,21%
Oct-00	5,52%	7,24%
May-01	3,11%	7,35%
Oct-01	3,36%	6,91%
May-02	3,07%	7,44%

Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.

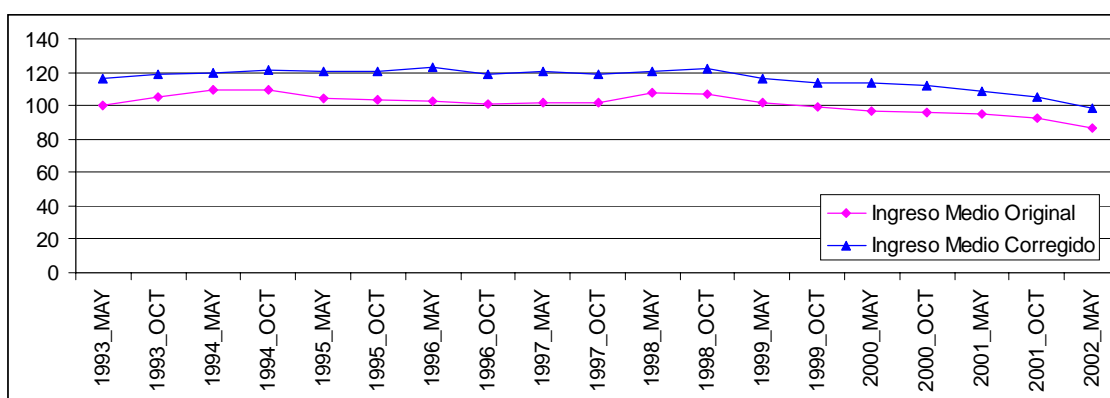
Las correcciones se realizan en el orden que fueron explicadas, y cada una corrige sobre los resultados obtenidos en la corrección anterior.

## Sección 2. Análisis de los resultados

Luego de realizar los ajustes pertinentes al ingreso reportado en la EPH, a partir de los problemas identificados en la captación del mismo, en esta sección se analizará el impacto de dichas correcciones sobre la distribución del ingreso laboral.

En primer lugar, cabe destacar que los ajustes generan un incremento del ingreso medio total, como puede observarse en el gráfico 1. La diferencia entre ambas series es relativamente estable (entre un 10% y un 21%), sobre todo a partir de 1996. Previamente se pueden identificar períodos en donde la evolución es diferente. En particular, la serie original muestra un incremento significativo en 1994, que no es replicado por la serie corregida.

**Gráfico 1. Ingreso Medio original y corregido. Serie mayo 1993-mayo 2002. (Ingreso Medio Original para mayo de 1993=100)**



Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.

En el Cuadro 6 se pueden observar en qué proporción fueron corregidos los promedios de ingresos laborales de cada decil, tomando como referencia tres ondas diferentes (Octubre de 1996, 1998 y 2000).

En las ondas elegidas, las tres correcciones son relevantes y positivas. Para el caso de la corrección por no declaración, se observa que el porcentaje de corrección entre los deciles es semejante, aunque es levemente menor para los deciles intermedios. Esto difiere en las dos correcciones restantes, donde se advierte un mayor ajuste en los últimos deciles. Así, la corrección por subdeclaración afecta principalmente a los dos últimos deciles y la no encuesta al último.

**Cuadro 6. Porcentaje de corrección para el promedio de ingresos por decil**

deciles	NO RESPUESTA			SUBDECLARACION			NO ENCUESTA			TOTAL		
	Oct-96	Oct-98	Oct-00	Oct-96	Oct-98	Oct-00	Oct-96	Oct-98	Oct-00	Oct-96	Oct-98	Oct-00
1	4%	5%	4%	3%	0%	1%	0%	0%	0%	6%	5%	5%
2	3%	4%	3%	2%	1%	1%	0%	0%	0%	4%	5%	3%
3	2%	2%	6%	1%	0%	0%	0%	0%	0%	2%	3%	7%
4	2%	3%	3%	1%	0%	1%	0%	0%	0%	3%	4%	4%
5	2%	2%	3%	1%	1%	1%	0%	0%	0%	3%	2%	3%
6	2%	2%	8%	1%	1%	1%	0%	0%	0%	3%	4%	9%
7	2%	2%	6%	1%	1%	3%	0%	0%	0%	4%	3%	10%
8	4%	2%	8%	2%	1%	2%	0%	0%	0%	5%	3%	10%
9	4%	4%	7%	7%	4%	6%	3%	1%	1%	15%	10%	14%
10	5%	3%	7%	16%	11%	11%	16%	15%	11%	41%	31%	32%
Ingreso Medio	3%	3%	6%	7%	5%	6%	6%	6%	4%	18%	14%	17%

Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.

El Cuadro A.1 del anexo muestra los ajustes por decil para cada categoría ocupacional, para octubre de 1998. Los ingresos de los empleadores son los que

presentan mayor corrección en los tres métodos, con una alta incidencia en la corrección al ingreso laboral total.

En la última columna del Cuadro 6 se observa la incidencia de las tres correcciones en los promedios de ingresos de cada decil. El hecho de que la corrección sea desigual para los distintos deciles tendrá un efecto en la distribución del ingreso entre los mismos.

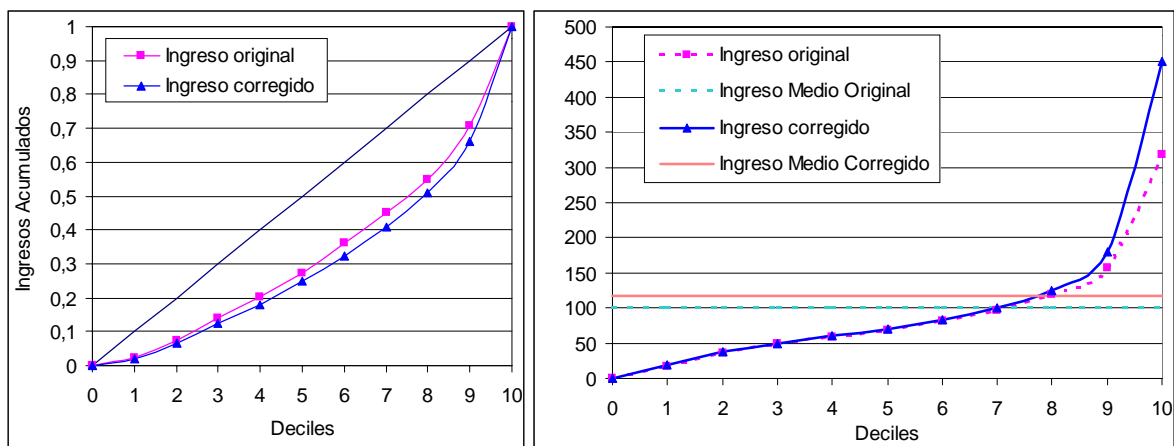
A continuación se analizan los cambios en la distribución del ingreso laboral introducidos por los ajustes, a partir de la construcción de indicadores de desigualdad.

### Curva de Lorenz e ingresos medios

Los gráficos 2, 3 y 4 realizados para las ondas de octubre de 1996, 1998 y 2000 respectivamente, permiten ver las diferencias entre la distribución de los ingresos reportados y los corregidos, a partir de la curva de Lorenz y los ingresos medios por decil.

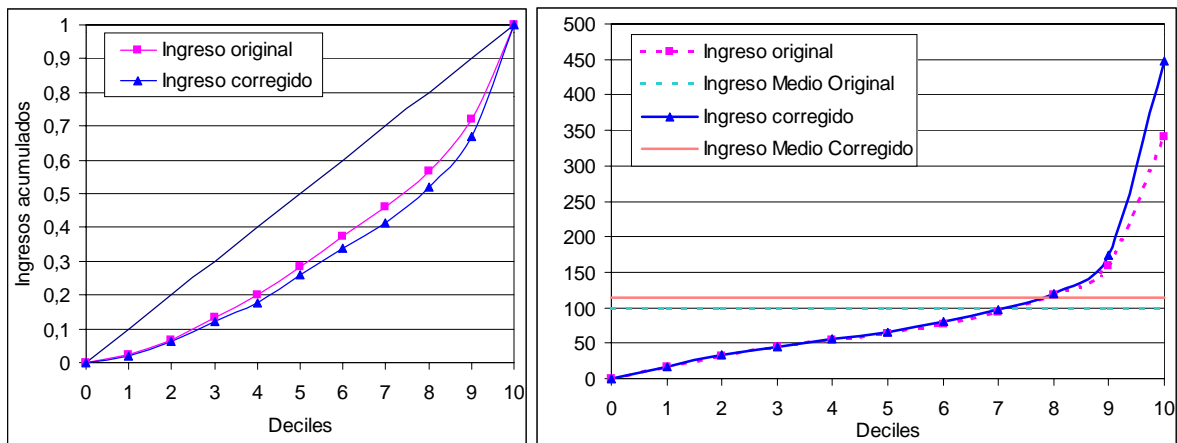
Luego de las correcciones, se advierte un aumento en la desigualdad, lo cual se puede observar en los gráficos de la curva de Lorenz. La desigualdad se acentúa con las correcciones por subdeclaración y no encuesta, dado que ambas afectan en mayor proporción a los deciles más altos, aumentando la distancia entre estos y los deciles más bajos. Esto se aprecia en los gráficos de ingreso medio, donde se verifica que el incremento del mismo luego de las correcciones se concentra en los últimos deciles.

**Gráfico 2. Curva de Lorenz e ingreso medio por decil para Octubre de 1996. (Ingreso Medio Original=100)**



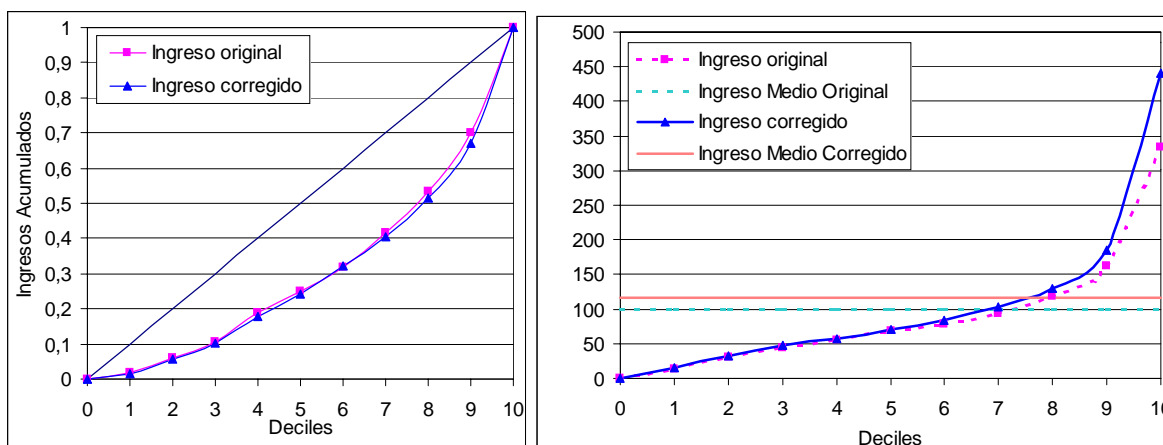
Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados

**Gráfico 3. Curva de Lorenz e ingreso medio por decil para Octubre de 1998. (Ingreso Medio Original=100)**



Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados

**Gráfico 4. Curva de Lorenz e ingreso medio por decil para Octubre de 2000. (Ingreso Medio Original=100)**

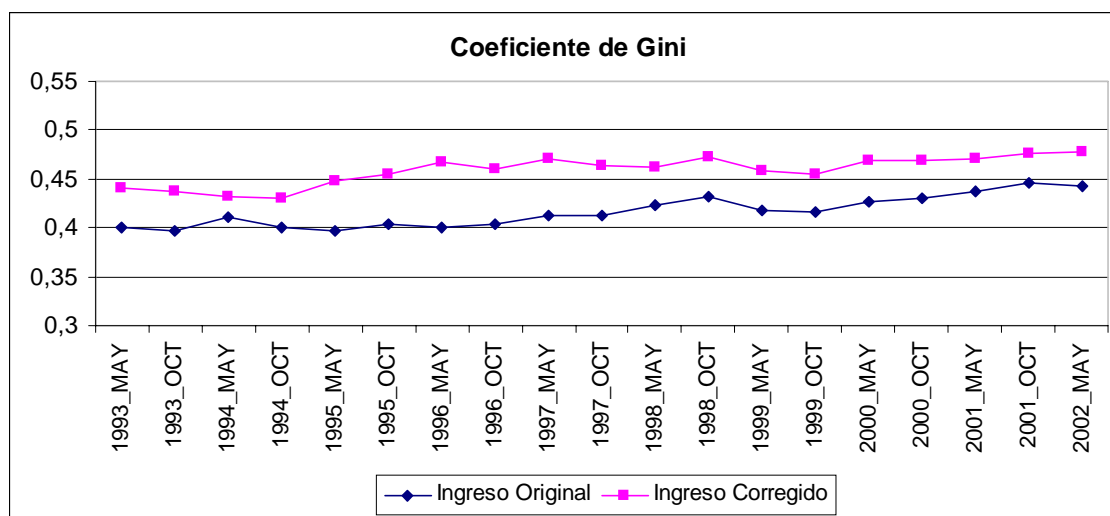


Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados

### **Coefficiente de Gini**

Para medir el impacto de las correcciones en la distribución del ingreso, se han construido los coeficientes de Gini para las ondas entre mayo de 1993 y mayo del 2002, para los ingresos reportados y los resultantes de las tres correcciones.

**Gráfico 5. Coeficiente de Gini. Serie mayo 1993-mayo 2002.**



Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.

Se verifica que el ajuste de ingresos provoca un aumento en la desigualdad. Dicha corrección incrementa, en promedio, un 10,25% al nivel del indicador. En particular, la corrección por no declaración incrementa un 0,87% el Gini, la corrección por subdeclaración un 4,46% adicional y la corrección por no encuesta un 4,62%. Una serie de los coeficientes de Gini, para los ingresos de cada una de las tres correcciones, se presenta en el cuadro A.2 del anexo.

El nivel de corrección es relativamente estable, sobre todo a partir de 1998. Ambas series muestran una tendencia creciente a lo largo del período bajo análisis. Sin embargo, la serie corregida muestra un incremento significativo de la desigualdad del ingreso laboral entre octubre de 1994 y mayo de 1996, un período con alto y persistente desempleo, incremento en la informalidad y en el diferencial de ingresos entre categorías. La serie original no muestra este aumento, sino otro entre octubre de 1996 y octubre de 1998.



### Indice de Atkinson

Otro indicador de desigualdad es el índice de Atkinson, cuyo parámetro de aversión a la desigualdad se asocia a la ponderación de la parte baja de la distribución: cuanto mayor sea el coeficiente, mayor importancia se le da a los ingresos de los primeros deciles (Cowell, 1995, y Medina, 2001).

Se construyeron los índices de Atkinson para toda la serie, comparando la distribución de ingresos reportados en la EPH con la que resulta luego de las correcciones. Se seleccionaron varios coeficientes de aversión a la desigualdad: 0,5; 2 y 3. En todos los casos el índice aumenta luego de las correcciones (ver cuadro A.2 del anexo).

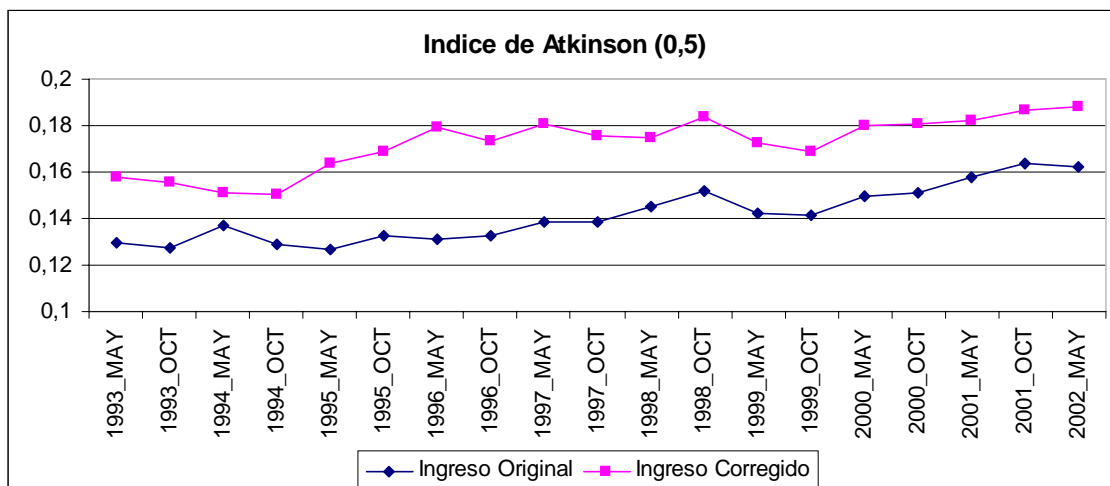
Dado que la corrección por no declaración incide en forma similar sobre el ingreso de los distintos deciles, su influencia en la desigualdad, y por lo tanto en el incremento del índice, no es significativo. Por el contrario, para coeficientes de aversión elevados, en algunos períodos esta corrección reduce la desigualdad, ya que, como surge del cuadro 6, el ajuste de los primeros deciles es mayor al de los deciles intermedios.

La corrección por subdeclaración tiene una mayor contribución sobre el incremento de la desigualdad, pero es la corrección por no encuesta la que más afecta al índice, debido a que la misma modifica únicamente los ingresos de los dos últimos deciles.

Se observa que al utilizar un coeficiente relativamente pequeño (0,5), con mayor sensibilidad a los cambios en la parte alta de la distribución, las correcciones provocan una reacción mayor en el índice de Atkinson que al utilizar un coeficiente más alto (3). Cuando el coeficiente de aversión a la desigualdad es igual a 0,5, el índice de Atkinson aumenta con las correcciones un 22,03%. En particular, en promedio para todo el período, la corrección por no declaración incrementa un 1,44% el índice, la corrección por subdeclaración un 9,27% adicional y la corrección por no encuesta un 10,05%. Cuando el coeficiente es igual a 2, el incremento del índice es solo del 9,96% (con incrementos por corrección de 0,51%; 4,52% y 4,65%, respectivamente), mientras que cuando se aplica un coeficiente igual a 3 el índice se incrementa en 5,99% (con incrementos de 0,07%; 2,85% y 2,97% por corrección)<sup>10</sup>.

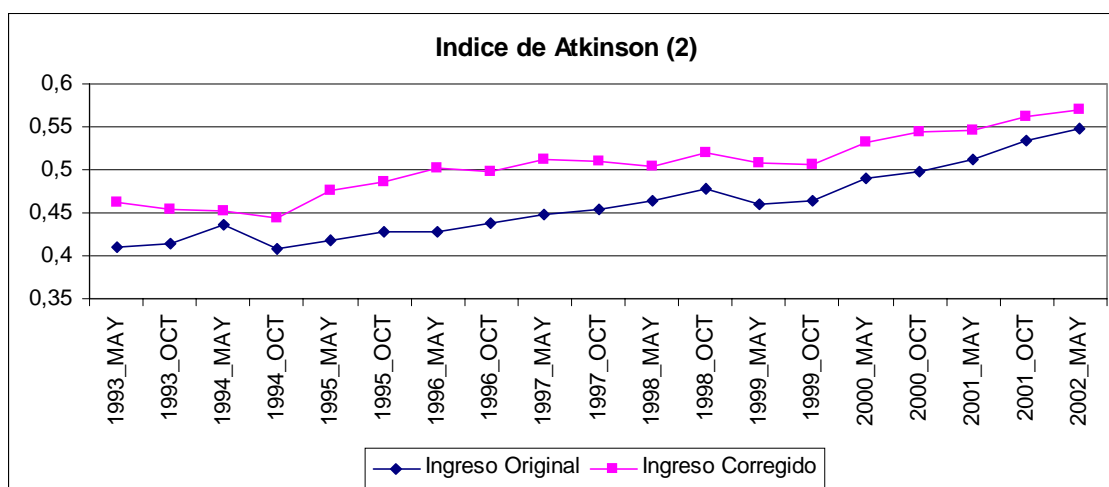
Al igual que para el caso del Gini, se observa una tendencia creciente para toda la serie, tanto en la original como en la corregida, siendo la misma más pronunciada para índices con coeficientes más elevados.

**Gráfico 6. Índice de Atkinson con coeficiente de aversión a la desigualdad = 0,5. Serie mayo 1993-mayo 2002.**



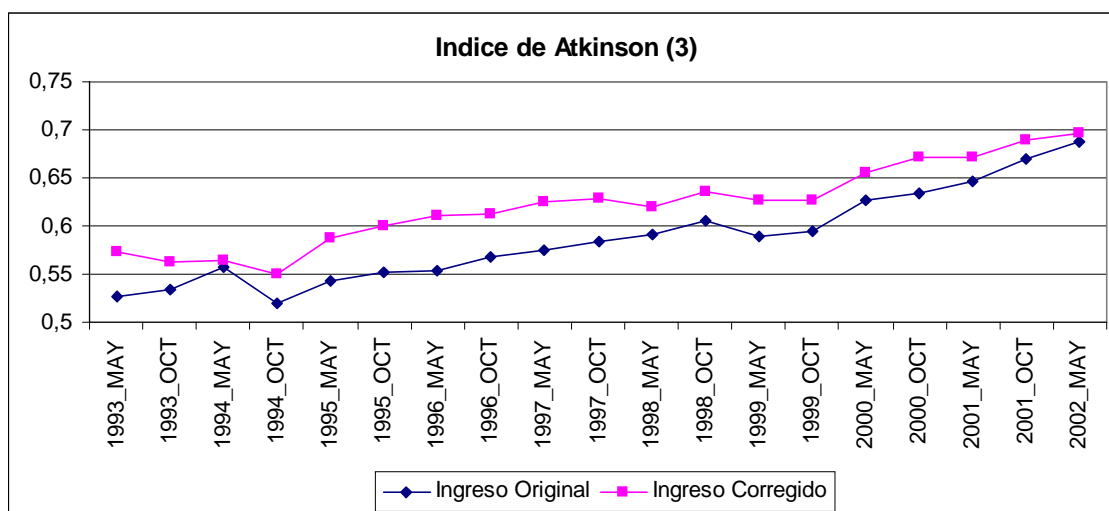
Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.

**Gráfico 7. Índice de Atkinson con coeficiente de aversión a la desigualdad = 2. Serie mayo 1993-mayo 2002.**



Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.

**Gráfico 7. Índice de Atkinson con coeficiente de aversión a la desigualdad = 3. Serie mayo 1993-mayo 2002.**



Fuente: Elaboración propia en base a EPH-INDEC. Total de aglomerados.

## Conclusiones

En el presente trabajo, se identificaron y corrigieron los principales problemas que presentan las encuestas a hogares en la captación del ingreso laboral: la no respuesta de ingresos, la subdeclaración y la no respuesta total. Dada la importancia de dichas encuestas como fuentes de información, y teniendo en cuenta que los problemas identificados introducen sesgos significativos en las estimaciones de ingresos, los mismos deberían ser corregidos para realizar estimaciones consistentes con otras fuentes de información, en particular, con las de Cuentas Nacionales.<sup>11</sup>

Como fuente se utilizó la EPH, tomando el ingreso de la ocupación principal. El trabajo se realizó para todas las ondas entre mayo de 1993 y mayo de 2002, y para todos los aglomerados.

Los ajustes realizados a la EPH generan estimaciones del ingreso laboral medio sustantivamente mayores y distribuciones notablemente más desiguales. El incremento de la desigualdad se debe a que la corrección impacta mayormente en los últimos deciles, lo cual provoca un mayor incremento en los indicadores de

desigualdad con mayor ponderación en esa parte de la distribución (Atkinson con parámetro 0,5).

Estos resultados son relativamente estables durante toda la serie. En concreto, el índice de Gini se incrementa entre un 8% y un 14%, mientras que el ingreso medio aumenta entre un 10% y un 21%. Consecuentemente, la evolución de las variables no presenta diferencias significativas, salvo en el período entre octubre de 1994 y mayo de 1996.

Los métodos utilizados aquí se basan en un análisis y comparación de fuentes de microdatos. La comparación de los datos (entre los que declaran y los que no declaran; entre las relaciones ingreso-gasto de autónomos y asalariados; y entre la distribución de dos fuentes comparables) implicó un ajuste sólo porque la diferencia demostró ser significativa para cada una de las correcciones. Asimismo, las correcciones difieren entre categorías, regiones y características del trabajador. En particular, es diferente para cada nivel de ingreso.

El ajuste presentado aquí se diferencia de los ajustes macro (Llach y Montoya, 1999; Gasparini, Marchionni y Sosa Escudero, 2000). Mientras estos últimos aplican un coeficiente para todos los ingresos de una fuente, aquí se corrige a nivel de cada observación de acuerdo a sus características, lo cual resulta en cambios en la distribución dentro de cualquier categoría o grupo.

También existe una diferencia en el nivel de ajuste, en especial entre las distintas categorías ocupacionales. Según el método presentado aquí, el ajuste aproximado en el ingreso de los patrones es de 70%, mientras que en los cuenta propias es de 25% y en los asalariados registrados y no registrados es de entre 3% y 4%. Por su parte, Llach y Montoya (1999) ajustan por medio de la corrección macro en un 28,2% a los asalariados y un 83,3% a los autónomos. Aunque menos comparable, cabe destacar que Gasparini, Marchionni y Sosa Escudero (2000) corrigen con un coeficiente de subdeclaración según fuente que, para el ingreso per cápita ajustado de la EPH, toma valores de 1.42, 1.92, 1.80, 1.96 y 2.12, según el quintil de ingreso. Debe tenerse en cuenta que estos coeficientes incluyen ajustes para ingresos no laborales.

En cuanto a la distribución del ingreso, en los trabajos citados se mide el ingreso per cápita familiar ajustado. Por ello, los resultados no son comparables. Igualmente, cabe destacar que en el caso de Llach y Montoya (1999) el coeficiente de Gini se incrementa en un 6% con la corrección macro para los años 1990-1998<sup>12</sup>. En el caso de Gasparini, Marchionni y Sosa Escudero (2000), la corrección no incrementa el Gini para los mismos años, pero sí el Atkinson(3) en un 5%.<sup>13</sup>

Por lo tanto, más allá de las diferencias, en todos los casos las correcciones tienden a incrementar el ingreso medio y a aumentar la desigualdad.

Las características de este método han permitido una estimación más adecuada de los salarios medios y los ingresos mixtos medios, lo que posibilita una mejor comparación con otras fuentes. En particular, en base al ingreso corregido se han calculado las cuentas de Generación del Ingreso de manera consistente con las otras cuentas del sistema de Cuentas Nacionales, sin mediar ajustes arbitrarios.

Igualmente, el trabajo futuro continuará mejorando los métodos aquí presentados. En particular, se requiere una estimación actualizada de la subdeclaración, que no suponga un coeficiente constante durante toda la serie. En este punto, especificaciones alternativas también podrían contribuir a una estimación más detallada de la subdeclaración por fuente y nivel de ingreso. Por otro lado, se debería minimizar la cantidad de observaciones remanentes con no respuesta parcial. Asimismo, el seguimiento del nivel de la economía oculta, de la informalidad y de las regulaciones del mercado laboral son indispensables para realizar ajustes en los métodos para diferentes años.

## ANEXO METODOLÓGICO

### **No respuesta. Imputación de ingresos laborales para los que no responden.**

Alrededor de un 10% de los ocupados recibe ingresos laborales pero no declara el monto en la Encuesta Permanente de Hogares. Para corregir los posibles sesgos generados por este problema, se ha estimado un modelo para imputar un ingreso en cada observación sin dato. El método es similar al utilizado en Gasparini, Marchionni y Sosa Escudero, 2000) aunque con una especificación diferente del modelo y correcciones para cada una de las regiones del país.

La lógica de la imputación es asignarle a un trabajador un ingreso equivalente al de otro trabajador con las mismas características. Las características que determinan el ingreso laboral son múltiples. Entre ellas deben considerarse, en primer lugar, las específicas del trabajador, principalmente las que se relacionan con su capital humano general (utilizable en una gran gama de puestos de trabajo) y específico (utilizable sólo en el puesto actual). Como variables más relevantes, se pueden citar educación, experiencia en general (inserción en el mercado de trabajo), antigüedad en el puesto de trabajo, además de edad, sexo, etc.

En segundo lugar, las características de la empresa o el establecimiento que contrata al trabajador también influyen en el salario pactado. En particular, son relevantes el sector de actividad y el tamaño de la firma.

Por último, deben considerarse las características del puesto de trabajo, es decir, la forma específica en que un trabajador se inserta en una empresa. Estas incluyen la forma contractual específica, la tarea que realiza, las horas trabajadas habitualmente, etc.

Para incluir todas estas variables, debe disponerse de una herramienta como la regresión multivariada. Por ello, se realizó una estimación de un modelo de determinación de ingresos laborales al estilo de Mincer (1974) en donde se estima el logaritmo del ingreso laboral horario en base a variables de características del trabajador, del establecimiento y del puesto. La elección de ingreso laboral horario es habitual en la literatura, al igual que su logaritimización.

La ecuación finalmente elegida ha sido:

$$\ln(IH) = c + \beta_1 T + \beta_2 E + \beta_3 P + \beta_4 D$$

en donde  $IH$  es el ingreso laboral horario,  $T$  es un conjunto (vector) de variables propias del trabajador, como la edad ( $d_{edad}$ ), la edad al cuadrado ( $d_{edadsq}$ ), los años de educación ( $e_{anios}$ ), los años de educación al cuadrado ( $e_{aniosq}$ ), dummies para educación superior ( $e_{dumsup}$ ) y educación universitaria ( $e_{dumuni}$ ), una dummy para mujer ( $d_{mujer}$ ) y otra para jefe de hogar ( $d_{jefe}$ );  $E$  es el conjunto de variables de características del establecimiento, como tamaño ( $tamano$ ), rama (una dummy para cada rama identificada,  $r_7$ , etc. según la clasificación de rama de la EPH), y una dummy para sector público ( $spubl$ );  $P$  es la serie de variables que caracterizan al puesto, como meses de antigüedad ( $antigue$ ), antigüedad al cuadrado ( $antiguesq$ ), calificación de la tarea (una dummy para profesional,  $prof$ , y otra para técnica,  $tecn$ );  $D$  es una serie de variables dummy para identificar diferencias medias por aglomerados en las regiones ( $a_1$ , etc.); las  $\beta$  son los coeficientes asociados a cada variable.

A partir de la estimación de los determinantes del ingreso laboral, en base a los trabajadores que lo reportaron, se obtienen coeficientes que permiten estimar el ingreso para cualquier trabajador a partir de sus características, en particular de los ocupados que no declararon ingreso laboral. Sin embargo, podría haber un sesgo entre los trabajadores que declaran ingresos y los que no lo declaran. Estos últimos podrían tener ingresos relativamente superiores que generen incentivos a no declararlos. Si este fuera el caso, el término de error estaría afectado por el sesgo de la censura de los datos. Para corregir este posible sesgo se aplica la corrección propuesta por Heckman (1979). Este método estima la corrección en los errores atribuibles a un sesgo de selección a partir de la probabilidad de no declarar ingresos.

En concreto, el modelo corregido por Heckman es el mismo que el expuesto arriba, más una última variable llamada *delta1*.

El modelo probit asociado a la probabilidad de declarar ingresos está basado en el mismo tipo de variables que el modelo de ingresos, es decir, las relacionadas con el trabajador, el establecimiento y el puesto, incluyendo algunas adicionales.

Las estimaciones se realizan para cada categoría ocupacional (empleadores, cuentrapropistas, asalariados informales y asalariados formales), teniendo en cuenta que los determinantes del ingreso no necesariamente afectan de la misma forma a cada categoría. En particular, el diferencial de ingreso de educación puede presentar formas funcionales diferentes en cada categoría. Asimismo, se estiman las regresiones para cada una de las seis regiones de la EPH.

Aunque tomar logaritmos del ingreso horario tiende a normalizar su distribución, igualmente persisten datos extremos. Para atenuar la incidencia de los *outliers* en la estimación de los determinantes del ingreso se utilizó un método robusto de estimación que permite obtener estimaciones consistentes a pesar de la existencia de datos extremos. Este método intenta evitar la no normalidad de los errores que puede generarse por la característica de las distribuciones de las variables del modelo. En concreto, la rutina utilizada realiza una reponderación en base a una iteración de estimaciones, en donde se reduce el peso relativo de los datos con errores más elevados. Finalmente, cada observación utiliza dos ponderadores: el que provee la encuesta y el del método robusto.

Una vez estimados los coeficientes de la regresión, se aplican a las observaciones que no presentan declaración de ingreso pero sí del resto de las variables relevantes. El resultado es una estimación del logaritmo del ingreso laboral horario para cada observación. A partir de las declaraciones de horas trabajadas, se estima el ingreso mensual.

Para resumir, la imputación de ingresos se realiza en la siguiente secuencia para cada categoría ocupacional:

- Se estima un modelo probit entre los que declaran y no declaran ingresos.
- En base a este modelo, se calcula la inversa del cociente de Mill, la variable *delta1*, para la aplicación de la corrección de Heckman.
- Se estima de manera robusta la ecuación del logaritmo del ingreso laboral horario, incluyendo la variable calculada anteriormente.
- Se imputa, en base a los coeficientes estimados, un nivel de ingreso laboral horario a cada ocupado que no declara ingresos.
- Se calcula el ingreso laboral mensual del trabajador en base a la declaración de horas trabajadas.

Como ejemplo, se presentan las salidas de las regresiones para el año 1994. Los datos para la estimación incluyen las ondas de octubre de 1993 hasta mayo de 1995, con el objetivo de agrandar la muestra, especialmente relevante para las estimaciones referidas a empleadores. La selección del modelo finalmente aplicado, se basó en numerosas pruebas de especificaciones alternativas aplicadas a la serie de datos entre 1993 y 2001. Se priorizó el modelo que fuera significativo, con un R2 relativamente elevado, con significatividad conjunta de cada grupo de variables y que fuera suficientemente general para ser aplicado en todos los años. También se comprobó que la variable asociada a la corrección de Heckman resulta significativa (lo cual implica que es relevante el sesgo introducido por la censura en los datos) y que el método de regresión robusta normaliza considerablemente los errores de la estimación.

**Cuadro A.1 Resultados del modelo de regresión para la imputación de ingresos. GBA, año 1994**

Variable dependiente: Logaritmo del ingreso horario

	Empleadores		Trabajadores por cuenta propia		Asalariados informales		Asalariados formales	
	Observaciones= 603		Observaciones= 2855		Observaciones= 2871		Observaciones= 6712	
	F( 18, 584) = 23.64		F( 21, 2833) = 83.73		F( 22, 2848) = 62.52		F( 22, 6689) = 317.40	
	Prob > F = 0.0000		Prob > F = 0.0000		Prob > F = 0.0000		Prob > F = 0.0000	
	R-squared = 0.4215		R-squared = 0.3830		R-squared = 0.3257		R-squared = 0.5107	
	Adj R-squared = 0.4037		Adj R-squared = 0.3784		Adj R-squared = 0.3204		Adj R-squared = 0.5091	
	Coefficiente	P> t	Coefficiente	P> t	Coefficiente	P> t	Coefficiente	P> t
d_edad	.042453	0.007	.066273	0.000	.060756	0.000	.01421	0.000
d_edadsq	-.00056	0.001	-.000645	0.000	-.00079	0.000	-.000218	0.000
d_mujer	.096088	0.198	-.076353	0.018	-.05395	0.014	-.187537	0.000
d_jefe	.699912	0.000	-.110929	0.004	.141318	0.000	.124406	0.000
antigue	.001688	0.012	-.000179	0.511	.003286	0.000	.0028	0.000
antigsq	-2.85E-06	0.086	3.10E-06	0.000	-4.11E-06	0.000	-5.10E-06	0.000
tamano	.021582	0.509	.079641	0.004	.036372	0.000	.03870	0.000
e_anios	.036317	0.402	-.027041	0.081	-.022659	0.048	.065607	0.000
e_aniosq	.001931	0.373	.001728	0.031	.002465	0.000	-.001690	0.000
e_dumsup	-.851051	0.010	-.002899	0.971	.002109	0.979	.056815	0.017
e_dumuni	-.12220	0.403	.2451	0.001	.043355	0.548	.077318	0.009
r7	...	...	...	...	-.49625	0.003	.311276	0.000
r8	.06453	0.480	.104010	0.003	-.141970	0.002	.278407	0.000
r10	.045611	0.527	.240074	0.000	-.127534	0.000	.217382	0.000
r11	-.61053	0.000	.71786	0.000	.072013	0.135	-.362538	0.000
r14	...	...	.531428	0.002	-.155330	0.050	.098760	0.000
r15	.553911	0.000	.116061	0.019	-.032275	0.423	-.03826	0.063
r21	...	...	-.316799	0.000	.371481	0.000	-.097204	0.039
spubl	...	...	.719371	0.000	.117907	0.051	-.039052	0.004
prof	-.552873	0.026	.977482	0.000	.4766	0.000	.396867	0.000
tecn	-.521024	0.010	.313350	0.000	.104906	0.004	.188269	0.000
ax	...	...	...	...	...	...	...	...
delta1	4.25091	0.005	-4.85992	0.000	3.1232	0.000	4.32441	0.000
_cons	-1.63630	0.003	1.30319	0.000	-1.20117	0.000	-.851532	0.000

Como se puede apreciar, todos los modelos son significativos. En general los signos son los esperados. En todas las categorías el perfil de ingresos por edad tiene la forma de U invertida, la educación tiene un impacto positivo, aunque la forma difiere (para asalariados formales el rendimiento de la educación es casi lineal, pero con saltos cuando la educación es terciaria o universitaria, para los empleadores es creciente<sup>14</sup>; y para los asalariados informales y los cuentapropistas el rendimiento de la educación primaria y secundaria es casi nulo)<sup>15</sup>. La antigüedad toma la forma habitual de U invertida para todas las categorías, salvo para los cuentapropistas. La mayor parte de las ramas de actividad seleccionadas aporta diferencias significativas, sobre todo para las categorías de asalariados formales y cuentapropistas. En todos los casos la variable asociada a la corrección de Heckman es significativa, aunque en un caso es negativa.

Igualmente, se trata de un modelo para realizar una predicción, por lo cual lo relevante es la significatividad conjunta de las variables. En este sentido, se ha verificado la significatividad conjunta de los grupos de variables: educación, edad y antigüedad.

Las implicancias en la corrección del nivel de ingreso promedio se muestran en el cuerpo del texto. Aunque se trata de un mismo modelo para todos los años, el mismo parece ser relevante y suficientemente general para varios contextos y categorías. La aplicación de este método, puede estar subestimando el verdadero ajuste, ya que se trata de asignarle a una observación sin dato el ingreso *promedio* que tienen otros trabajadores del mismo grupo. De esta manera, si hubiera características no observables que incrementaran el ingreso de los que no declaran, no estarían contabilizadas, incluso teniendo en cuenta el ajuste por Heckman.

### **Subdeclaración. Estimación de un coeficiente de subdeclaración para los autónomos.<sup>16</sup>**

El punto central de esta corrección es estimar un coeficiente de subdeclaración de ingresos. Si bien la subdeclaración puede ser un fenómeno generalizado, aquí se supone que se concentra en un grupo de trabajadores (en este caso, algunos autónomos). La estimación de la subdeclaración implica analizar la diferencia entre lo que declara un trabajador perteneciente a este grupo en relación a lo que hubiera declarado un trabajador con las mismas características dentro del otro grupo (asalariados). En otras palabras, si hubiera una subdeclaración de *todos* los trabajadores, aquí se corregiría la subdeclaración *adicional* del grupo considerado.

A diferencia del caso anterior, aquí no basta con disponer de datos de ingresos del trabajador. Para calcular la subdeclaración se incluyen datos de gastos y se explota la relación funcional entre ambas variables. Las diferencias en la relación ingreso-gasto para los dos grupos que no estén explicadas por características de la familia son interpretadas como una diferencia producida por la subdeclaración de ingresos.

Dos supuestos básicos explican este ejercicio. El primero establece que sólo un grupo de trabajadores subdeclara ingresos, debido a que tiene (mayores) incentivos y posibilidades. El segundo supone que la declaración de gastos es correcta para todos los grupos.

El grupo de trabajadores autónomos suele tener mayores incentivos a subdeclarar ingresos que los asalariados. En primer lugar, los encuestados suelen relacionar la declaración de ingresos con cuestiones impositivas, lo que motiva la subdeclaración por temor a posibles problemas con el fisco. Esto no sucede con los asalariados, quienes declaran su ingreso neto luego de las retenciones al impuesto a las ganancias. En segundo lugar, el ingreso de los trabajadores autónomos es relativamente variable, por lo que una declaración fuera del "rango" habitual es difícil de identificar. En el caso de los asalariados, el hecho de poseer los comprobantes salariales (recibos de sueldo), susceptibles de ser solicitados por el encuestador, limita las posibilidades de declarar incorrectamente el ingreso. Por estos motivos, los autónomos tienen tanto incentivos como posibilidades de subdeclarar ingresos.

Por otro lado, se considera que la declaración de gastos es correcta a partir de que no existen incentivos a subdeclararlos. Esto ocurre en particular con aquellos gastos que son frecuentes, de montos relativamente bajos y no considerados "de lujo".<sup>17</sup> Por ello, la declaración de gastos en alimentos es un candidato natural para realizar la regresión.<sup>18</sup>

Sin embargo, el ingreso laboral es una variable propia de la persona, mientras que el gasto se realiza a nivel del hogar. Por ello, para vincular ingresos y gastos hace falta realizar supuestos adicionales. Una alternativa es identificar a un hogar como "autónomo" cuando sus ingresos provienen mayormente de esta fuente.

A partir de estos supuestos se compara la relación entre el gasto y los ingresos de hogares de ambos grupos con las mismas características. Si por cada nivel de ingreso los hogares autónomos declaran un gasto superior al que declaran los hogares asalariados, entonces los autónomos están subdeclarando el ingreso percibido.<sup>19</sup>

Más formalmente, la ecuación estimada es:

$$GAI = \alpha_0 + \alpha_1 X + \beta_0 Y + \beta_1 X Y + \beta_2 Z$$

donde  $GAI$  es el logaritmo de los gastos en alimentos,  $Y$  es el logaritmo del ingreso del hogar ( $\ln\_ing\_t$  en la salida de la regresión),  $X$  es una variable dummy para el grupo de los hogares autónomos ( $nd\_auton$ ) y  $Z$  son un conjunto de variables de características del hogar, como cantidad de miembros ( $cant\_miem$ ), cantidad de niños entre 1 y 5 años ( $cant\_1y5$ ), cantidad de niños entre 5 y 10 años ( $cant\_5y10$ ), los  $\alpha$  y  $\beta$  son los coeficientes asociados a las variables correspondientes, mientras que el  $\alpha_0$  es una constante o gasto autónomo. En particular, esta ecuación supone una elasticidad ingreso constante y diferente para los asalariados ( $\beta_0$ ) y los autónomos ( $\beta_0 + \beta_1$ ).<sup>20</sup>

Suponiendo que en la situación promedio (en el ingreso promedio y para las mismas características del hogar) ambos gastos son iguales, se verifica que:

$$\alpha_0^* + \beta_0^* \bar{Y}^* = \alpha_0^* + \alpha_1^* + (\beta_0^* + \beta_1^*) \bar{Y}^*$$

donde  $Y^*$  es el ingreso real y los coeficientes son las estimaciones de elasticidad y consumo autónomo reales. Sabiendo que, por definición,  $Y = Y^* - S$ , donde  $Y$  es el ingreso observado y  $S$  es la tasa de subdeclaración, entonces

$$\alpha_1^* = -\beta_1^* (\bar{Y} + S)$$

Ahora bien, dado que en la especificación del modelo  $S$  no es observable, el término  $\alpha_1$  incluye el diferencial de gastos debido a la subdeclaración. Es decir:

$$\alpha_0 + \alpha_1 = \alpha_0^* + \alpha_1^* + (\beta_0^* + \beta_1^*) S$$

Se supone que el resto de los coeficientes estimados son iguales a los parámetros reales. Por lo tanto, se puede utilizar la ecuación previa y reemplazarla en la última. Así,

$$\alpha_1 = -\beta_1^* (\bar{Y} + S) + (\beta_0^* + \beta_1^*) S = -\beta_1^* \bar{Y} + \beta_0^* S$$

y finalmente:

$$S = \frac{\alpha_1 + \beta_1^* \bar{Y}}{\beta_0^*} = \frac{\alpha_1 + \beta_1 \bar{Y}}{\beta_0}$$

Así, la estimación de  $S$  (la subdeclaración promedio) se puede computar a partir de parámetros observables y el ingreso medio reportado.

Para realizar las estimaciones se han utilizado los datos de la Encuesta Nacional de Gastos de los Hogares, realizada en 1996 para las zonas urbanas del país. Se consideró dentro del grupo que subdeclara ingresos a todos los empleadores y a los trabajadores por cuenta propia que realizan tareas técnicas y profesionales, excluyendo a los trabajadores del sector público, servicio doméstico, agro y pesca. Se han eliminado también las observaciones que sólo reciben ingresos por jubilación y los hogares con jefes de más de 65 años. Para que sea considerado autónomo, el hogar debe percibir más del 25% de sus ingresos de ocupaciones autónomas.<sup>21</sup> La canasta de alimentos es la finalmente seleccionada. Nuevamente, la estimación es robusta y el ponderador utilizado combina el de la encuesta y el del método robusto.

Los resultados de la regresión se muestran en el cuadro siguiente.

Cuadro A.2. Regresiones de la relación ingreso gasto.

Variable dependiente: Logaritmo del gasto en alimentos

	Gran Buenos Aires		Interior	
	Number of obs = 1745		Number of obs = 9160	
	F( 8, 1736) = 209.28		F( 8, 9151) = 1222.32	
	Prob > F = 0.0000		Prob > F = 0.0000	
	R-squared = 0.4909		R-squared = 0.5166	
	Adj R-squared = 0.4886		Adj R-squared = 0.5162	
	Coeficiente	P> t	Coeficiente	P> t
<b>ln_ing_t</b>	<b>.490534</b>	<b>0.000</b>	<b>.491931</b>	<b>0.000</b>
cant_mie	.09383	0.000	.086446	0.000
jedad	.004748	0.000	.004022	0.000
jedadsq	-.000094	0.066	-.000075	0.001
cant_1y5	-.025316	0.256	-.00916	0.054
cant_5y10	.0139042	0.328	-.000230	0.954
<b>nd_auton</b>	<b>.850883</b>	<b>0.020</b>	<b>.662335</b>	<b>0.000</b>
<b>nd_ingau</b>	<b>-.102119</b>	<b>0.029</b>	<b>-.084938</b>	<b>0.000</b>
_cons	1.50200	0.000	1.5443	0.000

En este caso, los dos modelos resultan significativos y todas las variables asociadas al cálculo de la subdeclaración son significativas al 5%.<sup>22</sup> La elasticidad ingreso de los asalariados es de 0,49 en ambos casos y la de los autónomos es significativamente



menor, lo cual puede deberse a la mayor variabilidad de sus ingresos que hace que el ingreso de un período no sea considerado como permanente, reduciendo la respuesta del gasto a los cambios del ingreso. El gasto es significativamente superior para los hogares autónomos, controlando por todas las otras características, lo cual implica una subdeclaración significativa. Ésta se estima, utilizando la fórmula final, en 32,9% para el GBA, y en 22,9% para el interior.<sup>23</sup>

Una vez calculada la proporción de subdeclaración de los trabajadores autónomos, se aplica este coeficiente a todos los ingresos declarados por trabajadores de ese grupo. Se supone que la subdeclaración es relativamente constante a lo largo del tiempo y del nivel de ingreso. El incremento de los niveles de ingreso de los diferentes deciles puede apreciarse en el texto.

En conclusión, la subdeclaración demostró ser significativa en casi todas las especificaciones probadas, lo que justifica la corrección de ingresos de los autónomos. El modelo elegido ha sido el que presentó mejores características, con una subdeclaración promedio de 32.9% y 22.9% para GBA e interior, respectivamente. La estimación de Alaimo y Sosa Escudero (2000) para el GBA es de 23%, calculada sólo para empleadores, mientras que en Smith (1986) calcula para autónomos "white collar" un ocultamiento de 34% del ingreso para Gran Bretaña.

### **No encuesta. Comparación con otras fuentes.**

La no respuesta total es un problema típico de las encuestas a hogares y que afecta los resultados obtenidos. No existe un método de corrección previsto por la encuesta que compense el hecho de que un hogar (completo) no responda. Sólo se produce un ajuste en la ponderación, que da cuenta de la caída de la muestra en cada estrato muestral.

Al igual que los casos anteriores, este problema puede introducir sesgos en la estimación, ya que los hogares que no responden pueden tener características particulares (como estar ubicados en determinados barrios o tener todos los miembros del hogar ocupados y no disponibles para ser encuestados, etc). Es decir, si la pérdida de datos no es al azar, y en particular dependen del ingreso o de las variables que lo determinan, se introduce un sesgo en la estimación.<sup>24</sup>

Aquí podría aplicarse un sistema de corrección similar al utilizado en la corrección por no declaración de ingresos. Sin embargo, las bases de datos carecen de las características suficientes para realizar inferencias del hogar que no responde. Por ello, se ha optado por comparar los datos con otra fuente de información que complementa a la encuesta a hogares. Dicha fuente es el Sistema Integrado de Jubilaciones y Pensiones (SIJP).

La ventaja del SIJP como fuente de información es que se dispone de un universo completo de todos los asalariados formales privados, con su ingreso laboral declarado. No existe un truncamiento en la cantidad de trabajadores ni en el ingreso declarado en este sub-universo, ya que, por definición, todo puesto formal privado debe ser declarado al SIJP y estar en la base de datos.<sup>25</sup> Los incentivos a subdeclarar salarios se atenúan si se tiene en cuenta que los salarios son un costo, y como tales, se deducen del pago de ganancias.<sup>26</sup> Por otro lado, el trabajador sufre una retención del salario en concepto de ganancias, lo que implica que tampoco permitirá una sobredeclaración.

Se busca entonces comparar las *distribuciones* del salario neto de los asalariados formales del sector privado no agropecuario de la EPH y del SIJP. Se supone que las diferencias entre ambas fuentes se deben al problema de la no respuesta. Este supuesto se justifica en que se está comparando un universo contra una muestra en la cual una proporción (significativa) ha sido sustraída no aleatoriamente.<sup>27</sup>

Para realizar la comparación entre ambas fuentes se requiere un trabajo previo de compatibilización, tanto en cantidad de observaciones como en montos declarados, dado que se parte de dos conceptos diferentes. En la EPH se trata del salario de bolsillo habitual de la ocupación principal del trabajador; mientras el SIJP reporta la

remuneración total de un CUIL-CUIT. Por ello, se han sustraído de la remuneración total del SIJP los aportes a cargo del trabajador, las indemnizaciones pagadas, los pagos extraordinarios por productividad y la retención de ganancias cuando correspondiere, y se han sumado las asignaciones familiares cuando estas no estuvieran declaradas. De esta manera, se llega a un ingreso de bolsillo mensual.<sup>28</sup>

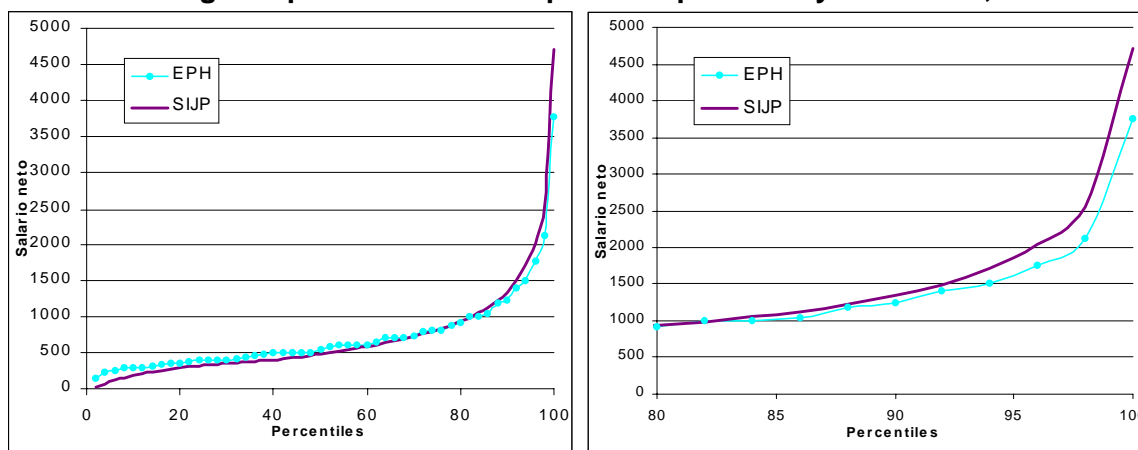
Para evitar datos no habituales, en el caso del SIJP para cada CUIL se ha calculado la mediana de un trimestre. Asimismo, si un CUIL tuviera dos declaraciones (un mismo trabajador con dos puestos diferentes), se ha seleccionado la de mayor remuneración, suponiendo que se trata de la ocupación principal.

De esta forma, se llega a conceptos comparables en ambas fuentes. A pesar de que las diferencias subsisten, estas se atenúan en la comparación de la distribución y, sobre todo, en los salarios más elevados.<sup>29</sup>

Las distribuciones comparadas corresponden a una onda de la EPH y el trimestre correspondiente en el SIJP, para cada una de las seis regiones de la EPH.<sup>30</sup> Para comparar la distribución se han agrupado los datos en 50-tilas.

Se muestran aquí las distribuciones del salario neto habitual de los asalariados formales para el año 2001 para el GBA.

**Gráfico A.1. Ingreso promedio de cada percentil para EPH y SIJP. GBA, 2001.**



*Fuente: Elaboración propia en base a EPH-INDEC y SIJP.*

La media de ambas distribuciones es similar, siendo 725 pesos para la EPH y 720 pesos para el SIJP. Sin embargo, ambas distribuciones se cortan, siendo la EPH significativamente mayor en los primeros percentiles y menor para los últimos. En particular, para el último cincuentil, el SIJP presenta ingresos un 25% mayores, mientras que para el último decil la diferencia es de 18.5%. El corte entre ambas distribuciones se verifica, en este caso, alrededor del percentil 80. Sin embargo, en dicho percentil la diferencia es mucho más baja (un 3% en este caso y cero para otras regiones).

Otras comparaciones de la media y la distribución de ambas fuentes de información han reportado resultados similares.<sup>31</sup> La divergencia en la parte baja de la distribución puede deberse a una subdeclaración del salario para puestos de baja calificación en el SIJP. Esta explicación es natural, y es extender a la informalidad del salario (parte del salario de bolsillo no declarado) la observación de la alta informalidad en este tipo de puestos. Por otro lado, la divergencia en los últimos percentiles puede deberse a subdeclaración o truncamiento de la muestra en la EPH, según se ha comentado previamente.

Es por ello que se ha optado por realizar las correcciones para los últimos cincuentiles. Así, para cada uno de los últimos cincuentiles y comenzando por último, todas las observaciones de la EPH con ingresos netos superiores a la media del 50-til se multiplican por un coeficiente de ajuste, de manera tal de igualar ambas distribuciones.

La corrección se realiza si la diferencia en las distribuciones es superior a 5%. La elección de este corte, aunque arbitraria, se basa en la observación de las distribuciones para todas las regiones en diferentes momentos del tiempo. Asimismo, el coeficiente se aplica no sólo a los asalariados formales, sino a todos los ingresos netos. Esta decisión se justifica por dos motivos. En primer lugar, la diferencia en las distribuciones podría deberse a una subdeclaración de ingresos en la EPH, en cuyo caso no alcanza con la corrección realizada previamente para los trabajadores autónomos, y deberían corregirse los ingresos del resto de los trabajadores, quienes también pueden haber subdeclarado ingresos<sup>32</sup>. En segundo lugar, podría deberse al efecto de la no respuesta total, la cual se distribuiría también en hogares compuestos por trabajadores de otras categorías. De esta forma, debería también incrementarse el peso y el ingreso de las otras categorías.

En conclusión, la comparación con otra fuente de ingresos manifiesta que la EPH no releva correctamente los ingresos laborales más elevados. Por ello, resulta conveniente realizar un ajuste de los mismos, para corregir este desfase.

Anexo Cuadros

Cuadro A.1. Ajuste por categoría ocupacional, por deciles e ingreso medio. Octubre 1998.

	DECIL	PATRONES	CUENTA PROPIA	ASAL. REGISTRADOS	ASAL. NO REGISTRADOS
1° CORRECCION	1	11,74%	20,26%	1,60%	11,24%
	2	12,75%	6,90%	0,97%	4,65%
	3	16,17%	4,53%	0,39%	3,25%
	4	4,07%	4,13%	2,28%	4,12%
	5	15,84%	4,69%	0,59%	3,83%
	6	0,86%	1,96%	1,06%	1,91%
	7	10,97%	1,65%	1,54%	1,75%
	8	3,00%	1,65%	0,63%	0,31%
	9	6,21%	2,36%	1,23%	-0,25%
	10	1,93%	2,55%	1,21%	-0,26%
		<b>Ingreso Medio</b>	<b>5,67%</b>	<b>2,84%</b>	<b>1,15%</b>
2° CORRECCION	1	30,95%	1,16%	0,00%	0,00%
	2	30,81%	1,96%	0,00%	0,00%
	3	30,97%	2,12%	0,00%	0,00%
	4	25,11%	2,19%	0,00%	0,00%
	5	30,89%	7,36%	0,00%	0,00%
	6	28,30%	2,64%	0,00%	0,00%
	7	33,05%	5,41%	0,00%	0,00%
	8	35,14%	9,70%	0,00%	0,00%
	9	26,49%	14,36%	0,00%	0,00%
	10	26,98%	20,10%	0,00%	0,00%
		<b>Ingreso Medio</b>	<b>29,15%</b>	<b>12,47%</b>	<b>0,00%</b>
3° CORRECCION	1	0,00%	0,00%	0,00%	0,00%
	2	0,00%	0,00%	0,00%	0,00%
	3	1,94%	0,00%	0,00%	0,00%
	4	10,88%	0,00%	0,00%	0,00%
	5	8,28%	0,00%	0,00%	0,00%
	6	17,05%	0,00%	0,00%	0,00%
	7	8,02%	0,00%	0,00%	0,00%
	8	15,61%	0,00%	-0,05%	0,00%
	9	27,16%	2,57%	1,66%	0,00%
	10	37,78%	17,48%	11,75%	5,06%
		<b>Ingreso Medio</b>	<b>21,22%</b>	<b>7,53%</b>	<b>3,63%</b>
TOTAL	1	46,33%	21,66%	1,60%	11,24%
	2	47,49%	9,00%	0,97%	4,65%
	3	55,09%	6,76%	0,39%	3,25%
	4	44,35%	6,41%	2,28%	4,12%
	5	64,17%	12,39%	0,59%	3,83%
	6	51,47%	4,65%	1,06%	1,91%
	7	59,51%	7,16%	1,54%	1,75%
	8	60,92%	11,51%	0,58%	0,31%
	9	70,84%	20,08%	2,91%	-0,25%
	10	78,33%	44,70%	13,10%	4,79%
		<b>Ingreso Medio</b>	<b>65,44%</b>	<b>24,38%</b>	<b>4,82%</b>

**Cuadro A.2. Índices de desigualdad. Mayo 1993 – Mayo 2002**

	May-93	Oct-93	May-94	Oct-94	May-95	Oct-95	May-96	Oct-96	May-97	Oct-97	May-98	Oct-98	May-99	Oct-99	May-00	Oct-00	May-01	Oct-01	May-02
<b>GINI original</b>	0,3999	0,3975	0,4111	0,4003	0,3960	0,4038	0,4011	0,4037	0,4118	0,4123	0,4229	0,4313	0,4175	0,4162	0,4270	0,4295	0,4381	0,4459	0,4434
<b>GINI 1° corrección</b>	0,3975	0,4041	0,4062	0,4029	0,4093	0,4110	0,4163	0,4081	0,4221	0,4131	0,4230	0,4319	0,4259	0,4233	0,4287	0,4340	0,4318	0,4410	0,4459
<b>GINI 2° corrección</b>	0,4216	0,4197	0,4244	0,4226	0,4272	0,4287	0,4332	0,4331	0,4399	0,4335	0,4396	0,4490	0,4427	0,4399	0,4512	0,4517	0,4548	0,4611	0,4569
<b>GINI 3° corrección</b>	0,4409	0,4375	0,4320	0,4311	0,4482	0,4544	0,4680	0,4594	0,4699	0,4631	0,4620	0,4725	0,4592	0,4548	0,4688	0,4691	0,4707	0,4761	0,4776
<b>ATK(0,5) original</b>	0,1299	0,1276	0,1369	0,1292	0,1269	0,1325	0,1310	0,1329	0,1384	0,1386	0,1455	0,1515	0,1420	0,1415	0,1497	0,1515	0,1576	0,1637	0,1624
<b>ATK(0,5) 1° corrección</b>	0,1275	0,1316	0,1336	0,1303	0,1354	0,1369	0,1406	0,1355	0,1447	0,1386	0,1454	0,1516	0,1474	0,1453	0,1500	0,1543	0,1530	0,1599	0,1634
<b>ATK(0,5) 2° corrección</b>	0,1437	0,1423	0,1456	0,1436	0,1478	0,1494	0,1525	0,1528	0,1573	0,1530	0,1574	0,1644	0,1596	0,1573	0,1662	0,1671	0,1696	0,1747	0,1716
<b>ATK(0,5) 3° corrección</b>	0,1580	0,1554	0,1513	0,1500	0,1636	0,1690	0,1794	0,1731	0,1808	0,1757	0,1750	0,1833	0,1726	0,1687	0,1802	0,1809	0,1825	0,1868	0,1879
<b>ATK(2) original</b>	0,4100	0,4144	0,4359	0,4075	0,4185	0,4286	0,4279	0,4387	0,4474	0,4550	0,4643	0,4785	0,4605	0,4643	0,4904	0,4982	0,5112	0,5347	0,5483
<b>ATK(2) 1° corrección</b>	0,4120	0,4168	0,4252	0,4082	0,4319	0,4383	0,4438	0,4418	0,4587	0,4547	0,4617	0,4750	0,4701	0,4693	0,4898	0,5066	0,5023	0,5264	0,5402
<b>ATK(2) 2° corrección</b>	0,4404	0,4331	0,4434	0,4334	0,4514	0,4574	0,4613	0,4678	0,4782	0,4781	0,4796	0,4950	0,4894	0,4893	0,5143	0,5270	0,5288	0,5476	0,5496
<b>ATK(2) 3° corrección</b>	0,4626	0,4537	0,4520	0,4431	0,4753	0,4865	0,5013	0,4973	0,5120	0,5107	0,5044	0,5207	0,5073	0,5053	0,5328	0,5449	0,5451	0,5626	0,5703
<b>ATK(3) original</b>	0,5261	0,5343	0,5572	0,5203	0,5424	0,5519	0,5531	0,5679	0,5745	0,5842	0,5906	0,6048	0,5891	0,5954	0,6261	0,6348	0,6462	0,6704	0,6870
<b>ATK(3) 1° corrección</b>	0,5304	0,5320	0,5440	0,5187	0,5522	0,5612	0,5650	0,5688	0,5834	0,5836	0,5864	0,5993	0,5966	0,5964	0,6227	0,6425	0,6378	0,6628	0,6757
<b>ATK(3) 2° corrección</b>	0,5551	0,5448	0,5569	0,5418	0,5677	0,5763	0,5781	0,5886	0,5986	0,6030	0,6007	0,6160	0,6128	0,6135	0,6419	0,6587	0,6588	0,6788	0,6819
<b>ATK(3) 3° corrección</b>	0,5737	0,5622	0,5641	0,5501	0,5875	0,6001	0,6111	0,6125	0,6259	0,6290	0,6204	0,6362	0,6269	0,6260	0,6559	0,6721	0,6710	0,6897	0,6970

**Cuadro A.3. Participación de ingresos medios por deciles**

	Oct-96				Oct-98				Oct-00			
Decil	Ingreso original	con 1° corrección	con 2° corrección	con 3° corrección	Ingreso original	con 1° corrección	con 2° corrección	con 3° corrección	Ingreso original	con 1° corrección	con 2° corrección	con 3° corrección
1	2%	2%	2%	2%	2%	2%	2%	1%	1%	1%	1%	1%
2	4%	4%	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%
3	5%	5%	5%	4%	4%	4%	4%	4%	4%	4%	4%	4%
4	6%	6%	5%	5%	5%	5%	5%	5%	6%	5%	5%	5%
5	7%	7%	6%	6%	6%	6%	6%	6%	7%	7%	6%	6%
6	8%	8%	8%	7%	8%	8%	7%	7%	8%	8%	8%	7%
7	10%	10%	9%	9%	9%	9%	9%	8%	9%	9%	9%	9%
8	12%	12%	11%	11%	12%	12%	11%	11%	12%	12%	12%	11%
9	16%	16%	16%	15%	16%	16%	16%	15%	16%	16%	16%	16%
10	32%	32%	35%	38%	34%	34%	36%	39%	33%	34%	35%	38%
<b>TOTAL</b>	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Nota: Los cuadros de este anexo son de elaboración propia en base a EPH-INDEC. Total de aglomerados.

## Referencias Bibliográficas

Camelo, H. (1993): "Comparación de los ingresos medios de la EPH con las Cuentas Nacionales", CEPAL, Buenos Aires.

Cowell, F.(1995): *Measuring Inequality*, Prentice Hall / Harvester Wheatsheaf.

Gasparini, Leonardo y Walter Sosa Escudero (2000) "Assessing Aggregate Welfare: Growth and Inequality in Argentina". Universidad Nacional de La Plata, Documento de Trabajo nro. 21. Marzo de 2000.

Gasparini, Marchionni y Sosa Escudero (2000): "La distribución del Ingreso en Argentina y la Provincia de Bs.As". Cuaderno de Economía N° 49, Marzo 2000.

Grosskoff, R. (1998): "Comparación de las Estadísticas de Ingresos Provenientes de Encuestas de Hogares con Estimaciones Externas", CEPAL, en 2º Taller Regional MECOVI: Medición del Ingreso en las Encuestas de los Hogares, Buenos Aires, Argentina, CEPAL.

Heckman, James (1979) "Sample Selection Bias as a Specification Error" *Econometrica*, vol. 47 issue 1.

Herrero, D. (2001): "Comparación armonizada de las estimaciones de población e ingresos del SIJP y la EPH (sector privado del Gran Buenos Aires ingresos devengados/percibidos en agosto/setiembre de 1997)", INDEC-Mecovi.

Llach, .J.J., Montoya, S. (1999): "En pos de la equidad. La pobreza y la distribución del ingreso en el área metropolitana de Buenos Aires: diagnóstico y alternativas políticas", IERAL, Buenos Aires.

Medina, F. (2001): "Consideraciones sobre el índice de Gini para medir la concentración del ingreso", en *Serie estudios estadísticos y prospectivos*, n° 9, CEPAL, Santiago de Chile.

Minardi, Graciela (2002) Informe Final Proyecto Mecovi: Calidad en Encuestas a Hogares.

Pissarides C., Weber G. (1989): "An expenditure-based estimate of Britain's black economy", en *Journal of Public Economics*, n° 39.

Roca E., Pena, H. (2001) "La declaración de ingresos en las encuestas de hogares", en *Estudios del Trabajo*, n° 22, ASET, Buenos Aires.

Smith, S. (1986) *Britain's shadow economy*. Oxford, Clarendon Press, 1986.

Sosa Escudero, Walter y Verónica Alaimo (2000). "La economía oculta en la Argentina: evidencia basada en Encuestas de Gasto", en *La economía oculta en la Argentina*, FIEL.

## NOTAS

<sup>1</sup> Para el caso de los asalariados, y dejando de lado la Seguridad Social, su ingreso laboral se corresponde con la retribución al trabajo. Sin embargo, para los cuenta propia y patrones, el ingreso se divide entre retribución al trabajo y retribución al capital, y se denomina ingreso mixto, dado que no es posible dividir ambos componentes de la declaración. Aquí no haremos hincapié en esta diferencia y denominaremos al ingreso de ambas categorías como ingreso laboral. Por otra parte sería bueno aclarar aquí que probablemente la declaración de ingresos por parte de los no asalariados a las encuestas de hogares no sea la totalidad del ingreso mixto, sino un ingreso luego de descontar las inversiones, acumulación de existencias o pago de préstamos relacionados con su negocio

<sup>2</sup> Para una explicación más detallada de la metodología empleada para las correcciones, ver el Anexo Metodológico.

<sup>3</sup> Un ajuste similar ha sido aplicado en Gasparini, Marchionni y Sosa Escudero (2000). Ver Anexo Metodológico para una explicación más detallada del método

<sup>4</sup> Sosa Escudero y Verónica Alaimo (2000) aplican este método para el Gran Buenos Aires. Este trabajo es básicamente una ampliación de los resultados al total país, con algunas diferencias menores en la aplicación del método, como ser la definición del grupo que subdeclara, que en nuestro caso se limitó a los empleadores y los cuenta propia con actividades de calificación profesional o técnica. Para una explicación más detallada del método ver Anexo Metodológico.

<sup>5</sup> Se selecciona un bien para el cual existan pocos incentivos a distorsionar la declaración de su gasto.

<sup>6</sup> Minardi (2002) realizó un análisis sobre las características de quienes no responden la Encuesta de Condiciones de Vida para el área metropolitana apareando los no respondientes con su cédula del Censo 2001, donde se comprueba que, en general, se trata de hogares que podrían asociarse a los altos ingresos.

<sup>7</sup> También es posible que exista una subdeclaración de ingresos salariales. Si fuera una subdeclaración en toda la muestra, ambas distribuciones serían similares pero con promedios y modas diferentes (el SIJP más alto). Sin embargo, la moda del SIJP es menor a la de la EPH, pero en los deciles más altos sus ingresos son mayores. Por ello, en el caso de haber subdeclaración, la misma estaría concentrada en el decil más alto.

<sup>8</sup> De esta manera no se corrige el truncamiento pero se llega a un resultado, en promedio, similar al SIJP, sin afectar la distribución de puestos de la EPH. Se probaron otros métodos alternativos que han demostrado resultados similares.

<sup>9</sup> El supuesto es que este problema es tanto o más importante para las otras categorías ocupacionales. Es decir, que la distribución de todos los ingresos se encuentra truncada en un nivel similar.

<sup>10</sup> Un cuadro con los valores de los indicadores de desigualdad para toda la serie se expone en el Anexo de Cuadros.

<sup>11</sup> Sin embargo, dado que existe un amplio consenso en que se trata de un problema inherente a todas las encuestas a hogares, no es recomendable su corrección a fin de realizar comparaciones internacionales.

<sup>12</sup> También realizan una corrección de las ponderaciones, pero que no afecta significativamente a la distribución.

<sup>13</sup> Debe tenerse en cuenta que, como ya se ha explicitado, estos autores también realizan una corrección de la no respuesta.

<sup>14</sup> El coeficiente negativo en la variable de educación universitaria no debe confundirse con una menor rentabilidad de este tipo de educación frente a la secundaria, ya que los coeficientes asociados a los años de educación tienen un efecto mucho mayor.

<sup>15</sup> Estos análisis son sólo ejemplificatorios, ya que la especificación del modelo no tiene el objetivo de aislar el efecto de cada variable sino realizar una predicción del conjunto de variables. Así, mientras un modelo explicativo debería evitar la posible multicolinealidad entre la calificación de la tarea y la educación, en este caso este aspecto no resulta tan relevante, sino el efecto conjunto de la educación y la calificación.

<sup>16</sup> Esta sección se basa fundamentalmente en Sosa Escudero y Alaimo (2000), quienes, utilizando la metodología planteada por Smith (1986) y Pissarides y Weber (1989), aplicaron el método aquí expuesto para el Gran Buenos Aires. Nuestro trabajo lo ha extendido para el total del país, introduciendo algunos cambios menores en el método.

<sup>17</sup> Por su evidente relación con el ingreso, estos últimos gastos también serán subdeclarados si el encuestado busca ocultar ingreso consistentemente. Asimismo, algunos bienes pueden también tener implicancias impositivas (patrimoniales).

<sup>18</sup> Los resultados de Smith (1986) y Sosa Escudero y Alaimo (2000) avalan esta elección. En ambos trabajos se concluye que las estimaciones más confiables se basan en esta canasta de gasto. Asimismo, se han probado otras canastas de consumo y las mismas han generado mediciones significativas de la subdeclaración, aunque menores.

<sup>19</sup> Debe considerarse que, de no haber subdeclaración, el gasto de los autónomos probablemente sería menor, ya que deben destinar parte de su ingreso neto al pago de impuestos y seguridad social ya descontados en los ingresos asalariados (formales).

<sup>20</sup>  $\beta_1$  es el coeficiente asociado a la variable  $nd\_ingau$  de la regresión.



---

<sup>21</sup> Se han probado también otras definiciones alternativas, pero esta ha sido la que mejores resultados ha provisto, además de estar justificada en que se trata del grupo más relacionado con una posible subdeclaración. (Ver Smith 1986).

<sup>22</sup> Cabe aclarar que se han probado varias especificaciones alternativas, entre las que se encuentran las basadas en el supuesto de igual elasticidad para ambos grupos, variables como cantidad de adultos equivalentes del hogar, diferentes definiciones de hogar autónomo, diferentes canastas de gastos, para cada una de las regiones. Asimismo, se han realizado estimaciones que incluían diferentes elasticidades para *fuentes* de ingresos alternativas, lo cual no exigía clasificar al hogar sino a la fuente de ingreso. Las características del modelo elegido en cuanto a significatividad de las variables incluidas en el cálculo de la subdeclaración y significatividad del modelo resultaron ser superiores.

<sup>23</sup> La media del logaritmo del ingreso es de 6.75 para GBA y 6.47 para el interior.

<sup>24</sup> Ver Minardi, (2002)

<sup>25</sup> Si bien pueden existir moras en las declaraciones, los trabajos de actualización realizados en Cuentas Nacionales permiten afirmar que la incidencia de estas no es significativa para el período analizado.

<sup>26</sup> Asimismo, las empresas no tienen incentivos a subdeclarar salarios elevados (en general ubicados en los últimos deciles), dado que la remuneración imponible presenta un tope, por encima de lo cual no se pagan cargas sociales.

<sup>27</sup> Las divergencias en las distribuciones pueden deberse también al problema de subdeclaración. En cualquier caso, la forma de corrección también daría cuenta de este sesgo si existiera.

<sup>28</sup> El trabajo de Herrero (2001) explicita más detalladamente las diferencias conceptuales entre ambas fuentes y da cuenta de un proceso de compatibilización similar al que se ha realizado para este trabajo. También se puede citar el trabajo de Roca y Pena (2001) como antecedente. Debe aclararse también, que el trabajo previo sobre esta fuente por parte de la oficina de Cuentas Nacionales que ha permitido una estimación confiable del ingreso asalariado formal privado y su distribución.

<sup>29</sup> Por ejemplo, si un asalariado tiene una ocupación principal informal y un segundo puesto formal no sería observado en la EPH y sí en el SIJP. Aunque esta divergencia no es menor, no afecta en las comparaciones de los asalariados con ingresos mayores, que es donde se centra la comparación.

<sup>30</sup> Para obtener regiones comparables se han considerado sólo las declaraciones al SIJP correspondientes a los aglomerados relevados por la EPH. Aquí, nuevamente, persistieron diferencias en la comparación ya que no todos los aglomerados relevados por la EPH pueden ser correctamente identificados en el SIJP (la localización se clasifica por zona de declaración definida *ad-hoc*).

<sup>31</sup> Herrero (2001) y Roca y Pena (2001).

<sup>32</sup> Cabe recordar que la corrección para los autónomos se trataba de una subdeclaración adicional de este grupo por sobre los asalariados, por lo que si hubiera subdeclaración de los asalariados, debe corregirse el ingreso de los autónomos en una relación similar.