



Maestría en Economía
Facultad de Ciencias Económicas
Universidad Nacional de La Plata

TESIS DE MAESTRIA

ALUMNO
Verónica Alaimo

TITULO
Estimaciones de la Subdeclaración de Ingresos en Base a Discrepancias en
el Consumo. Aplicaciones al Caso Argentino

DIRECTOR
Walter Sosa Escudero

FECHA DE DEFENSA
3/30/2001

**Estimaciones de la Subdeclaración de Ingresos
en base a Discrepancias en el Consumo.**
Aplicaciones al Caso Argentino.¹

Verónica Alaimo

Director: Dr. Walter Sosa Escudero

Maestría en Economía
Departamento de Economía
Universidad Nacional de La Plata

La Plata, 30 de marzo de 2001

¹ Agradezco a mi Director, Walter Sosa Escudero, por su confianza y apoyo; a Fernando Navajas, por creer siempre en la economista que hay en mí; y a Damián Bonari, por permitirme conciliar las horas de trabajo con las de estudio. Los comentarios de Leonardo Gasparini y de los asistentes al Seminario Interno del Departamento de Economía de la FCE-UNLP enriquecieron el trabajo. Una versión del capítulo 2 aparece en “La Economía Oculta en la Argentina” FIEL (2000).

Resumen ejecutivo

Este trabajo utiliza los datos de la Encuesta Nacional de Gasto de los Hogares realizada por el INDEC en 1996/97 para estimar la subdeclaración de ingresos de las familias del Gran Buenos Aires a partir de discrepancias en el consumo de las mismas.

Sus contribuciones principales son: (i) la utilización de un método con nula implementación en el caso argentino y la mejora de ciertos aspectos técnicos del mismo; y (ii) la obtención de estimaciones de un fenómeno importante como lo es el tamaño de la economía informal.

El método de *discrepancias en el consumo* se basa en un principio simple: si dos familias similares declaran ingresos semejantes, pero el consumo de una de ellas es notoriamente superior al de la otra, esto podría sugerir la presencia de subdeclaración de ingresos. Su aplicación requiere el uso de dos supuestos: (a) existe un grupo de familias que reporta correctamente su ingreso; y (b) todos los hogares declaran correctamente el gasto en algún tipo de bien.

El trabajo presenta dos versiones de este método, que tienen en común, entre otros, la función a estimar, el resultado obtenido y algunas de sus interpretaciones. En efecto, ambas estiman una ecuación que vincula al (logaritmo del) consumo con el (logaritmo del) ingreso, con ciertas características del hogar y la familia y con una variable binaria que separa a las familias en dos grupos (las que potencialmente subdeclaran ingresos y las que no lo hacen). Asimismo, coinciden en la obtención de una tasa de subdeclaración de ingresos *relativa* (de un grupo respecto al otro) y no absoluta.

Sin embargo, estas dos versiones del método de discrepancias en el consumo se diferencian en varios sentidos. La segunda versión supone que la volatilidad del ingreso permanente varía de un hogar a otro, en particular, que los hogares caracterizados como los que “reportan correctamente su ingreso” tienen un ingreso permanente con menor volatilidad que los que lo reportan con error. De este modo, esta versión obtiene una tasa de subdeclaración *promedio*, mientras que la primera fuerza este resultado. Por otra parte, mientras la primera versión emplea mínimos cuadrados, la segunda evita el problema de endogeneidad del ingreso, instrumentándolo y utilizando por ende mínimos cuadrados en dos etapas.

Los resultados empíricos obtenidos indican que bajo el método simple, el grupo de hogares que subdeclara ingresos, denominado “empleador”, oculta a las estadísticas oficiales aproximadamente el 23% de sus ingresos, mientras que con el modelo extendido la tasa asciende al 43%. Estos resultados tienen cierto grado de similitud con los encontrados por otros autores para el caso de Gran Bretaña. Sin embargo, debido a que las estimaciones obtenidas descansan en los supuestos realizados, no se pueden hacer comparaciones directas con respecto a mediciones para el caso argentino obtenidas con otros métodos (que también se basan en supuestos restrictivos).

Introducción

La subdeclaración de ingresos es un fenómeno que afecta la calidad de la información relevada por las estadísticas oficiales y por ende invalida las conclusiones que de ellas se derivan.

Los motivos que llevan a los agentes económicos a reportar incorrectamente sus ingresos varían dependiendo del órgano recolector de la información. Por ejemplo, cuando las familias y las empresas subdeclaran ingresos ante el ente recaudador el objetivo es evadir el pago de impuestos y ésta es una de las actividades usualmente definidas como parte de la “economía informal” de un país. Por otro lado, las familias también declaran con error ingresos ante otros organismos oficiales, por ejemplo los que realizan encuestas sociodemográficas (como las relevadas por el INDEC). Esto puede explicarse tanto por cuestiones vinculadas con la evasión impositiva (si no confían en la confidencialidad de la encuesta) como por otros factores, entre los que se encuentran (i) la posibilidad de que los ingresos del hogar sean variables en el tiempo, dificultando el cálculo de un ingreso mensual; (ii) la existencia de una remuneración no monetaria que la familia no considere un “ingreso” y por ende no lo declare como tal; (iii) la falta de interés del hogar en contestar la encuesta correctamente.

Este trabajo utiliza una encuesta de gasto de los hogares para aproximar la tasa de subdeclaración a través de *discrepancias en el consumo* de las familias. Aún cuando el modelo económico utilizado y su posterior implementación econométrica revistan cierta complejidad, el principio en el cual se basan es muy simple: si dos familias similares declaran ingresos semejantes, pero el consumo de una de ellas es notoriamente superior al de la otra, esto podría sugerir la presencia de subdeclaración de ingresos.

Si bien este método es usado por otros autores para estimar el tamaño de la economía informal, los resultados obtenidos deben interpretarse con cautela, debido a varias cuestiones. En primer lugar, los motivos de subdeclaración de ingresos en las encuestas de hogares pueden deberse a otras razones, no vinculadas con la economía informal. En segundo lugar, como se verá más adelante, el método utilizado no necesariamente arroja una tasa de subdeclaración absoluta, sino relativa, ya que parte del supuesto crítico de que un grupo de hogares no subdeclara sus ingresos. Teniendo

en cuenta esto, y con el objeto de precisar el alcance de la presente investigación, podría decirse que la misma intenta aproximar *una* de las actividades vinculadas con la economía informal, la subdeclaración de ingresos.

Por otra parte, un aporte adicional de este trabajo se vincula con el uso de las encuestas de hogares para -entre otros- el cálculo de indicadores de pobreza y coeficientes de desigualdad, los cuales se ven alterados debido al fenómeno de subdeclaración.

El capítulo 1 presenta una revisión de la definición de la economía informal, de las actividades involucradas y de los métodos frecuentemente utilizados para su estimación. Los capítulos 2 y 3 contienen el principal aporte de este trabajo: reportan modelos microeconómicos a través de los cuales se obtienen estimaciones de la subdeclaración presente en encuestas de gasto. En particular, el capítulo 2 consta de tres secciones: la primera discute un modelo simple y estimable económicamente con datos de encuestas de consumo; la siguiente describe la información a utilizar para la medición: la Encuesta de Gastos de los Hogares del Área Metropolitana Gran Buenos Aires, relevada por el INDEC en el año 1996; y la última presenta los resultados de la estimación econométrica del modelo descrito en la sección 1. El capítulo 3 se divide en dos secciones: la primera describe un modelo que presenta algunas ventajas respecto al utilizado en el capítulo anterior y la segunda reporta los resultados de la estimación econométrica correspondiente. Finalmente se presentan algunos comentarios a modo de conclusión.

Capítulo 1

Conceptos básicos, antecedentes y aplicaciones

La expresión “economía informal” en su uso más frecuente abarca todas aquellas actividades que contribuyen al producto bruto nacional oficialmente calculado, pero que no están registradas. Sin embargo, esta definición parece simplista ya que brinda muy poca información sobre el amplio espectro de situaciones que podrían asimilarse como actividades “informales”.

La definición de las actividades involucradas y el nombre que se les da es una de las controversias existentes en el tema. Los otros aspectos conflictivos giran en torno a los procedimientos de estimación o medición y el uso de estas estimaciones en el análisis económico. En este sentido, es importante apreciar que, debido a la naturaleza de las actividades definidas como informales, no existen métodos directos para su estimación. Por el contrario, existen diversos métodos basados en supuestos restrictivos que limitan el alcance de la estimación (qué se está midiendo) y la consecuente comparación de los resultados obtenidos por los distintos métodos.

Por ello, resulta de sumo interés para este trabajo tener en claro cuál es la definición de economía informal, qué aspectos de la misma se pueden medir con cada método y cuáles son sus respectivas limitaciones. Este capítulo describe los principales aspectos arriba planteados.

1.1. Definición del problema

La literatura en este campo utiliza indistintamente términos como economía informal, oculta, negra, subterránea, paralela, en efectivo, entre otros.² Esta abundancia de nombres genera dos problemas: por un lado, se utilizan distintos términos para referirse al mismo fenómeno; por otro, se llama de igual modo a cuestiones diferentes del problema. Por este motivo, ha habido en los últimos años esfuerzos por sistematizar y clasificar las distintas terminologías en circulación con el objeto de avanzar hacia las demás cuestiones también controvertidas del problema.

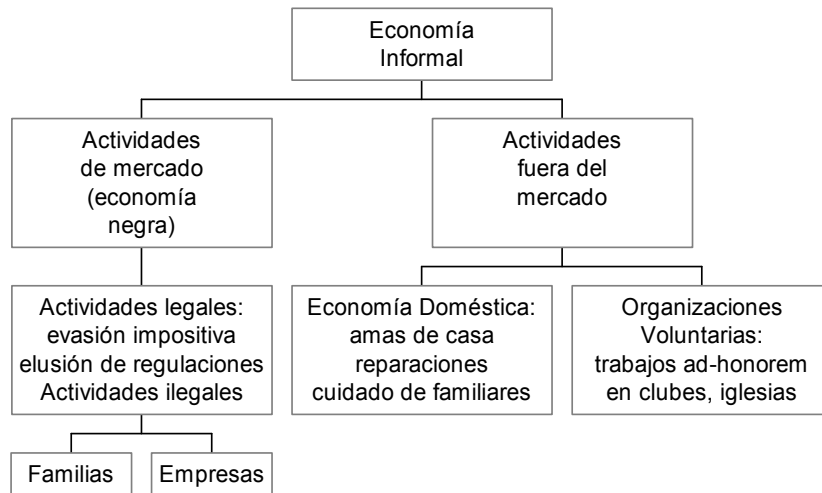
² Una revisión detallada puede encontrarse en Gasparini, L (2000) en “La Economía Oculta en la Argentina: Introducción”. FIEL.

En primer lugar, Smith (1986) distingue entre economía *formal* e *informal* y clasifica las actividades consideradas en ellas como dentro y fuera del mercado. De este modo, la economía *formal* abarca todas las transacciones de mercado registradas y aquellas fuera del mercado (principalmente las asociadas al gasto público de todos los niveles de gobierno). En cambio, dentro de la economía *informal* (shadow economy) se incluyen como actividades de mercado aquellas definidas como “black economy” (vinculadas con la evasión o ilegales) y como fuera del mercado las actividades productivas en los hogares o instituciones voluntarias.

Es decir que, según esta taxonomía, el término “shadow economy” abarca a todas las actividades económicas no registradas en las Cuentas Nacionales, incluyendo las actividades productivas realizadas en el hogar o en organizaciones voluntarias, así como también las actividades no reportadas ni registradas, asociadas con la evasión impositiva y / o previsional. En cambio, el término “black economy” es utilizado en un sentido más estricto ya que abarca a las actividades que implican evasión impositiva y / o previsional y a aquellas que están “ocultas”³ por ser ilegales.

Tabla 1.1.

Actividades involucradas en la economía informal



Fuente: Smith (1986).

Si bien podría decirse que, contrariamente a la economía informal, la formal involucra a todas las actividades registradas en las Cuentas Nacionales de un país, esta división no es siempre tan obvia. Las Cuentas Nacionales suelen realizar imputaciones de algunos ítems como la renta de la vivienda ocupada por sus dueños o los autos de una compañía que son utilizados por sus empleados, etc. Asimismo, las Cuentas Nacionales ajustan por evasión: en ese caso, sólo podría considerarse como economía informal aquella porción de la evasión que excediera el ajuste ya realizado. Smith no tiene en cuenta este aspecto y por ello no podría decirse que la economía informal está compuesta por las actividades no registradas en las Cuentas Nacionales.

1.2. Revisión metodológica

Debido a la naturaleza de las actividades catalogadas como informales, no es posible contar con métodos directos de medición ya que los agentes involucrados en ellas *deciden* ocultarlas por diversos motivos (evasión impositiva, ilegalidad, etc.). Existen varios métodos indirectos (también llamados de “indicadores”) que intentan estimar la magnitud de las actividades ocultas y todos ellos utilizan supuestos simplificadores que deben ser considerados al momento de comparar los resultados obtenidos con distintas metodologías.

a) Discrepancias entre gasto e ingreso nacional.

Se basa en las discrepancias entre las estadísticas de gasto e ingreso. En Cuentas Nacionales, el producto bruto nacional (PBI) calculado a través del gasto debería igualar la medición a partir de los ingresos. Cuando esto no ocurre, la discrepancia entre ambos valores suele considerarse como una aproximación de la economía oculta. Como se mencionó anteriormente, los cálculos de Cuentas Nacionales incluyen estimaciones sobre la evasión e imputan algunos ítems, como la renta de la vivienda de los propietarios ocupantes de las mismas. Por ello, los resultados extraídos de este método deben interpretarse en términos relativos ya que captan la economía informal que excede las estimaciones oficiales. Algunas aplicaciones de este método se encuentran en Franz, A. (1983) o Smith, J. D. (1985).

³ El término utilizado es “concealed”.

Thomas (1992) realiza una síntesis sobre este método y enfatiza sus deficiencias. No existen trabajos publicados para el caso Argentino.

La metodología de estimación de la presente investigación es un caso particular de ésta, ya que mide la discrepancia entre ingresos y gastos de las familias, a través de los resultados reportados por la Encuesta Nacional de Gasto de los Hogares (ENGH), realizada por el INDEC entre 1996 y 1997. Smith, Pissarides y Weber (1986) y Pissarides y Weber (1989) aplican este método utilizando la Family Expenditure Survey (FES) de Gran Bretaña. Esta alternativa se presenta brevemente al final de esta sección y se desarrolla con más detalle en el resto de la investigación.

b) Discrepancias entre la fuerza laboral oficial y real.

Este método toma como dada la participación total de la fuerza laboral y considera una caída en la participación de la misma en la economía oficial como un indicador de que aumentó la actividad en el sector informal. Estudios de este tipo existen para Italia y Estados Unidos.⁴

Las críticas a este método se basan fundamentalmente en dos aspectos: (i) la tasa de participación de la fuerza laboral varía, y (ii) la gente puede trabajar en ambos sectores (oficial e informal). No existen trabajos publicados para el caso Argentino.

c) El enfoque monetario.

Esta es una de las metodologías más utilizadas. En este caso, el tamaño de la economía informal se aproxima a partir de la demanda de dinero, suponiendo que las transacciones “ocultas” se realizan en efectivo. El método consiste en estimar la demanda de dinero incluyendo todos los determinantes convencionales posibles. De este modo, un incremento en la demanda de dinero no explicado por las variables independientes del modelo se debe a un movimiento en el mismo sentido de la economía informal. Este método ha sido utilizado por Gutman (1977), Feige (1979) y perfeccionado por Tanzi (1980, 1983), Bhattacharyya (1999). Guisarrri (1986) y Ahumada et al. en FIEL (2000) realizan estimaciones para la Argentina.

El método aplicado por Gutman (1977) se basa en cuatro supuestos centrales: a) la alta presión impositiva y las regulaciones son la única causa de la existencia de

⁴ Véase Del Boca (1981) u O'Neil (1983), respectivamente.

una economía informal; b) todas las transacciones de la economía informal se realizan utilizando circulante; c) la relación circulante sobre depósitos sólo es función del nivel de presión tributaria y del conjunto de disposiciones que regulan la actividad económica; y d) para algún momento del tiempo el tamaño de la economía informal es despreciable.

Quienes critican este método fundamentan que (i) no todas las transacciones involucradas en la economía informal se realizan en efectivo; (ii) al menos para el caso de Estados Unidos, estudiado por Tanzi, el dólar es una moneda con demanda internacional y por ende debería eliminarse del análisis la tenencia de dólares en el extranjero; y (iii) el supuesto de velocidad del dinero constante en ambos sectores de la economía es poco creíble: si existe mucha incertidumbre en las estimaciones de la velocidad de dinero en el sector formal, más aún habrá en estimaciones de esta variable en el sector informal. En este aspecto hacen especial hincapié Ahumada et al. en FIEL (2000): plantean que el supuesto de igualdad de la velocidad del dinero en ambos sectores (formal e informal) sólo genera estimaciones consistentes del tamaño de la economía informal cuando la elasticidad-ingreso de la demanda de circulante es igual a la unidad.⁵

d) El método de los insumos

Los métodos que se encuadran en este enfoque, en general, consideran que la electricidad es un insumo que permite realizar inferencias sobre el tamaño de la economía informal en un país o región. Uno de ellos es *el método de Kaufmann – Kaliberda*, que considera al consumo de electricidad como el mejor indicador de la economía en su conjunto y consiste en estimar el tamaño total de la economía a través del consumo de electricidad para luego sustraer el monto de PBI oficial.

Si bien existe evidencia internacional que confirma una estrecha relación entre ambas variables, con una elasticidad electricidad – PBI cercana a la unidad, las críticas a esta metodología se basan en que (i) no todas las actividades económicas requieren una cantidad considerable de electricidad, (ii) existen otras fuentes de energía, (iii) el uso de la electricidad ha mejorado a lo largo del tiempo debido al avance tecnológico y esto impacta de igual modo en el sector oficial y no oficial de la

⁵ Para mayor detalle, véase Ahumada et al. (2000) en “La Economía Oculta en la Argentina”, Capítulo I. FIEL.

economía, y (iv) se observan diferencias y cambios en la elasticidad electricidad – PBI tanto temporal como espacialmente (distintos países).

1.2.1. El método de discrepancias en el consumo

El enfoque a utilizar en este trabajo se basa en los estudios de Smith, Pissarides y Weber (1986) y Pissarides y Weber (1989), posteriormente reelaborados por Lyssioutoum, Pashardes y Stengos (1999).

Estos estudios realizan un análisis microeconómico de encuestas de gastos y son, esencialmente, estudios de *corte transversal* que permiten medir actividades informales a través de la contrastación del comportamiento de diferentes grupos de individuos en un momento en particular (el momento de realización de la encuesta).

La esencia del método, que se describe con detalle en el capítulo siguiente, se basa en comparar los patrones de ingreso y consumo de dos grupos de individuos, en donde se supone que uno de ellos tiene pocos incentivos a subdeclarar sus ingresos en la encuesta. Si además se considera un bien (o grupo de bienes) para el cual *ambos* grupos de individuos tienen pocos incentivos a subdeclarar su consumo, la presencia de discrepancias en el consumo de este bien entre estos grupos que no se puede explicar por factores observables, tales como el ingreso declarado u otras características familiares, es interpretada como proveniente de la subdeclaración de ingresos. En términos todavía más elementales, asumiendo que existen dos familias similares (igual número de hijos, edad, educación, etc.) A y B en donde A no subdeclara sus ingresos y B si lo hace. Suponiendo también que ambas familias declaran correctamente su consumo de alimentos. Si el ingreso real (no observado) de ambas es similar, el consumo de alimentos debería ser similar. Entonces, en las encuestas de gasto, la familia B (que subdeclara ingresos) aparecerá consumiendo relativamente más que la familia A. Bajo los supuestos realizados, la discrepancia en las proporciones de consumo de ambos grupos permitiría inferir la cantidad de ingreso subdeclarado.

Lyssioutoum, Pashardes y Stengos (1999) critican este método por dos motivos: (i) por un lado, consideran que el método no distingue las discrepancias en el consumo debidas a la subdeclaración de las vinculadas con la heterogeneidad de las preferencias; (ii) por otro lado, creen que la clasificación de los hogares de acuerdo a su principal fuente de ingreso es arbitraria ya que excluye del análisis ingresos que

pueden provenir de actividades vinculadas con economía informal simplemente porque no representan la principal fuente de ingreso del hogar.

1.3. Relevancia de los resultados obtenidos

Como es de esperar, la aplicabilidad de estos métodos indirectos, y la relevancia práctica de sus conclusiones, están sujetas a una fuerte controversia (ver Dixon, 1999). En un extremo, autores como Thomas (1999) se muestran completamente escépticos con respecto al estado de las artes en lo que se refiere al problema de medición de economía oculta. Por otro lado, autores como Giles (1999) o Bhattacharyya (1999) se muestran como más optimistas frente a las posibilidades de los métodos de medición.

Como se señalara anteriormente, los métodos indirectos requieren el uso de supuestos acerca del comportamiento de los agentes y de su relación con los datos existentes. De modo que cualquier interpretación de un análisis de este tipo depende crucialmente del tipo de supuestos implementados, y de la naturaleza de los datos utilizados para realizar la medición. Desde este punto de vista, la mayor parte del escepticismo se refiere a la relevancia de los supuestos implementados y a la calidad de la información disponible, más que a la validez de las metodologías estadísticas implementadas. En los términos planteados por Manski (1994), la discusión pasa por un problema de *identificación* (qué conclusiones se podrían sacar aún cuando se dispusiera de una ilimitada cantidad de observaciones) más que por uno de *inferencia estadística* (qué se puede aprender con una cantidad limitada de información).

Por otro lado, existen muy pocas experiencias de medición de economía informal para el caso Argentino (Guisarri (1986), FIEL (2000)). Esto se debe, principalmente, a las dificultades metodológicas anteriormente descritas, y a que el interés en este tipo de problemas es relativamente reciente. En lo que respecta al método de estimación aquí utilizado, la escasez de datos de encuestas de consumo para el caso Argentino representa la mayor limitación.

Capítulo 2

Un modelo simple de subdeclaración de ingresos basado en discrepancias en el consumo

Este capítulo presenta un modelo muy simple de subdeclaración que puede ser estimado económicamente utilizando datos de corte transversal, como los de la ENGH. El objetivo de la presentación consiste en destacar el tipo de supuesto necesario para poder identificar, y posteriormente cuantificar, las actividades ocultas. La estructura del modelo es deliberadamente simplista, y posteriormente se consideraran varias extensiones que permitirían levantar algunos supuestos. La exposición se basa en Smith et al. (1986).

2.1. El Modelo

El punto de partida consiste en suponer que existen dos grupos de individuos que se denotarán como A y E . Ambos grupos tienen una estructura de consumo similar, que se asumirá log-lineal. Más específicamente, se supone que C es el consumo en un bien particular por una familia, Y el ingreso familiar y Z un vector de características observables del hogar (diferentes del ingreso) que determinan el consumo (tamaño de la familia, educación, estructura etaria, etc.). La especificación econométrica supone que los valores esperados del consumo en logaritmos para ambos grupos estará dada por:

$$E(\ln C^A) = \alpha_A + \beta_A \ln Y + \delta Z \quad (1)$$

$$E(\ln C^E) = \alpha_E + \beta_E \ln Y + \delta Z \quad (2)$$

en donde $\alpha_A, \beta_A, \alpha_E, \beta_E$ y δ son parámetros desconocidos.

Esta estructura permite que el (logaritmo del) consumo de ambos grupos de individuos tenga diferentes interceptos y elasticidades ingreso, pero supone que los determinantes del consumo que no son el ingreso (Z) operan de la misma forma para ambos grupos de individuos. Adicionalmente, se espera que para el ingreso promedio, dos familias idénticas en cuanto a sus características Z consuman en promedio lo mismo, o sea, igualando (1) a (2) y llamando Y^* al ingreso promedio:

$$\alpha_A + \beta_A \ln Y^* = \alpha_E + \beta_E \ln Y^* \quad (3)$$

Se llama Y^R al ingreso *reportado* por las familias a la encuesta, y S a la tasa de subdeclaración, de modo que:

$$\ln Y^R = \ln Y - S \quad (4)$$

S , entonces, será la *tasa de subdeclaración*, la cual será el objeto de la estimación empírica. Los supuestos cruciales de identificación son los siguientes: 1) ambos grupos declaran correctamente el consumo del bien en cuestión y 2) uno de los grupos de individuos declara correctamente sus ingresos mientras que el otro lo subdeclara, y es posible identificar en la muestra si una familia pertenece a uno u otro grupo. En particular, se supone:

$S = 0$ si el individuo pertenece al grupo A

$S > 0$ si el individuo pertenece al grupo E

El siguiente paso consiste en expresar S en términos de magnitudes observables y / o estimables a partir de la información disponible. Expresando Y en términos del ingreso reportado y de la subdeclaración, y reemplazando en (2):

$$\begin{aligned} E(\ln C^E) &= \alpha_E + \beta_E (\ln Y^R + S) + \delta Z \\ &= \alpha^* + \beta_E \ln Y^R + \delta Z \end{aligned} \quad (5)$$

en donde:

$$\alpha^* = \alpha_E + \beta_E S \quad (6)$$

Por otro lado, despejando en la restricción (3) se obtiene:

$$\alpha_E = \alpha_A + (\ln Y^{*R} + S) (\beta_A - \beta_E)$$

y reemplazando en (6) y despejando S :

$$S = [(\alpha^* - \alpha_A) + \ln Y^{*R} (\beta_E - \beta_A)] / \beta_A$$

Es importante notar que, expresada de esta manera, la tasa de subdeclaración S es una magnitud que depende de valores observables y/o estimables. Las ecuaciones (1) y (5) son funciones de regresión que dependen enteramente de magnitudes observables, de modo que sus parámetros pueden ser consistentemente estimados por métodos estándar (como mínimos cuadrados ordinarios) para los grupos correspondientes, pero imponiendo la restricción de que los coeficientes de Z sean iguales en ambas ecuaciones. Esto produce estimaciones consistentes de α^* , α_A , β_E y β_A . Adicionalmente, Y^{*R} , el ingreso promedio reportado, es una magnitud directamente observable. Reemplazando en S por los valores estimados y observados se obtienen, consecuentemente, estimaciones consistentes de S .

En la práctica resulta conveniente proceder de la siguiente manera. Denotando con D a una variable binaria que toma valor 1 si un hogar pertenece al grupo E , y 0 si pertenece al grupo A , y considerando el siguiente modelo de regresión para todos los individuos de la muestra:

$$E(\ln C) = b_1 D + b_2 \ln Y + b_3 (D * \ln Y) + \delta Z \quad (7)$$

Es fácil verificar que los coeficientes de este modelo están relacionados con los del caso anterior de la siguiente forma:

$$\begin{aligned} b_1 &= \alpha^* - \alpha_A \\ b_2 &= \beta_A \\ b_3 &= (\beta_E - \beta_A) \end{aligned}$$

y, en particular,

$$S = [b_1 + \ln Y^{*R} b_3] / b_2 \quad (8)$$

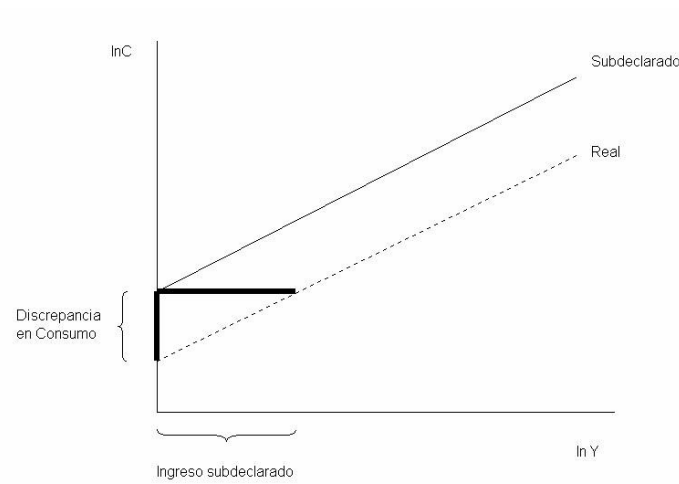
De modo que S puede ser estimada consistentemente a través de una regresión del logaritmo del consumo en D , $\ln Y$, $D * \ln Y$ y Z para *todos* los individuos de la muestra, y posteriormente reemplazando las estimaciones de b_1 , b_3 y b_2 , y el valor observado de $\ln Y^{*R}$ en (8).

Los Gráficos 2.1. y 2.2. muestran estas relaciones. Para simplificar el ejemplo, se asumirá que las familias son idénticas en lo que respecta a su conformación demográfica o, alternativamente, que $Z=0$. El Gráfico 2.1. se refiere a un caso extremo con dos grupos de individuos (A y E) con patrones de consumo

idénticos, en donde ambos declaran correctamente su consumo pero uno de ellos (el grupo E) subdeclara el ingreso. Las observaciones del grupo A se encontrarán dispersas alrededor de la recta denotada como “real” y las del grupo B sobre la recta denotada como “subdeclarado”. La estimación econométrica de (7) permite identificar la distancia vertical entre los patrones de consumo de ambos grupos (las *discrepancias en el consumo*), la cual puede ser fácilmente traducida en términos de la distancia horizontal, que provee una estimación de la tasa de subdeclaración. El Gráfico 2.2. muestra una situación más realista, en donde el grupo que subdeclara además tiene una menor elasticidad ingreso en el consumo. En este caso, el modelo y los datos identifican los patrones de ingreso consumo observados. Bajo el supuesto de que para el ingreso promedio ambos grupos tienen el mismo consumo, entonces es posible identificar la subdeclaración en este punto, representada en el gráfico por la línea sólida.

Gráfico 1.1.

Modelo simple de subdeclaración: Iguales elasticidades de consumo



Habiendo controlado por Z, las discrepancias en el consumo se deben fundamentalmente a subdeclaración de ingresos y, consecuentemente, de no ser el caso, lo que se está estimando es subdeclaración y cualquier otra cosa que haga que

2.2. Los Datos

Los datos utilizados corresponden a un conjunto de variables de la Encuesta Nacional de Gastos de los Hogares (ENGH) relevada entre febrero de 1996 y marzo de 1997 por el Instituto Nacional de Estadística y Censos (INDEC). Este estudio se concentra en la Región Metropolitana del Gran Buenos Aires.⁶

La ENGH releva una gran cantidad de variables referidas a la estructura de gasto e ingreso de los hogares, a las cuales adiciona información sobre distintas características demográficas, ocupacionales y educacionales de los hogares encuestados.

El modelo discutido en la sección anterior implica que, a los fines de la discusión, es necesario identificar en la muestra: 1) Grupos de hogares clasificados de acuerdo a las posibilidades de subdeclarar su ingreso. 2) Grupos de bienes cuyo consumo se declare en forma correcta. 3) Características de los hogares que permitan controlar por diferencias entre familias, de modo de evitar sesgos en la estimación de la subdeclaración, de acuerdo a lo discutido anteriormente.

El primer paso consistió, siguiendo la metodología descrita en Pissarides y Weber (1989), en seleccionar únicamente los hogares en que había dos adultos y el segundo era el cónyuge del jefe de familia.

Con respecto a los grupos de ingresos, en forma similar a los trabajos anteriores, se intentó clasificar a los individuos en asalariados y no asalariados. La encuesta permite distinguir ingresos por fuentes asalariada, cuentapropista y empleador.⁷ Una cuestión básica es que la unidad de referencia en donde se determina el consumo es el *hogar*, de modo que es necesario clasificar a los *hogares* de acuerdo a su fuente de ingreso. Aquí se siguió la misma línea metodológica que en Smith et al. (1986) y Pissarides y Weber (1989), que clasifican a un hogar como, por ejemplo, “asalariado” de acuerdo a la proporción del ingreso total obtenido a través de esta fuente.

En forma similar a estos trabajos, se probaron varias alternativas basándose en el jefe del hogar y en la totalidad del ingreso familiar. En el primer caso, se trabajó con dos posibilidades (i) teniendo en cuenta lo que cada jefe de hogar reportó como su ocupación principal y (ii) teniendo en cuenta el porcentaje de cada fuente de

⁶ Para una descripción detallada de la encuesta, su metodología y objetivos ver INDEC (1998)

ingreso en el total del jefe. En el segundo caso, se trabajó únicamente con la participación de cada fuente en el total del ingreso del hogar (más del 25%, más del 30%, 40% y 50%). Como los resultados obtenidos fueron similares a los de la opción elegida, los mismos no se comentan aquí por razones de espacio (en el Anexo II.1. se muestran los resultados para la definición del hogar de acuerdo a la fuente que representa más del 50% del ingreso familiar).

Tabla 2.1.
Hogares clasificados por categorías de ingreso
estructura porcentual a partir de frecuencias expandidas

	Criterios de categorización de los hogares en base al ingreso familiar					
	más del 25% por fuente E	más del 25% por fuente C	más del 25% por fuentes EyC	más del 50% por fuente E	más del 50% por fuente C	más del 50% por fuentes EyC
Empleadores (E)	6,0%			5,2%		
Cuentapropistas (C)		27,2%			19,3%	
Empleadores y Cuentapropistas (EyC)			32,8%			24,4%
Otros	94,0%	72,8%	67,2%	94,8%	80,7%	75,6%
Hogares (1)	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

(1) Corresponde a 3.402 observaciones (2.319.777 hogares expandidos).

Fuente: Elaboración propia en base a ENGH.

La Tabla 2.1. presenta la proporción de hogares clasificados como empleadores, cuentapropistas o ambos, de acuerdo a dos criterios basados en el ingreso total familiar. La clasificación adoptada para las estimaciones presentadas en las secciones siguientes define a un hogar como “empleador”, “cuentapropista” o “empleador y cuentapropista” si más del 25% del ingreso total familiar proviene de la fuente respectiva (en el último caso, se suma el ingreso proveniente de ambas fuentes). Por ejemplo, la segunda columna de la Tabla 2.1. indica que el 6% de los hogares obtienen más del 25% de su ingreso de la fuente empleador y el resto entraría en la categoría de otros, que en este caso significaría “no empleador”.

También se consideraron distintos grupos de bienes, clasificados de acuerdo a las posibilidades que las familias tienen de subdeclarar su consumo. Los estudios anteriores sugieren concentrar el análisis en bienes de baja elasticidad ingreso, a partir de los cuales, naturalmente, resulta complejo inferir el ingreso de las familias a partir de su consumo. Se consideraron: alimentos, alimentos y bebidas, alimentos, bebidas y tabaco, commitments, servicios y bienes no durables, etc.

⁷ También distingue entre ingresos por fuente rentista, jubilado, transferencias y autoconsumo.

El gasto en alimentos incluye: productos de panadería, harinas, arroz y cereales, pastas frescas o secas, achuras y menudencias, carne vacuna, porcina, ovina, aves, pescados frescos o congelados, otras carnes frescas, en conserva, congeladas, embutidos y fiambres, aceites, grasas, leche, productos lácteos, huevos, frutas frescas, secas, en conserva o congeladas, verduras, tubérculos y legumbres frescos, secos, en conserva o congelados, azúcar, dulces y miel, cacao y chocolates, caramelos y confituras, especias, aderezos y condimentos, comidas listas para consumir y otros productos alimenticios.

Al gasto en alimentos se agrega el consumo de bebidas alcohólicas y no alcohólicas, comidas y bebidas en bares en horas de esparcimiento y por otras razones para obtener la segunda canasta de consumo. Luego, se sumó el gasto en cigarrillos y tabaco para crear un tercer indicador de consumo de los hogares.

En todos los casos, los rubros incorporados en cada clasificación dependieron de la disponibilidad de datos de la ENGH. En este sentido, el gasto en “commitments”, que agrupa las erogaciones en intereses por pago de hipotecas, alquileres, reparación y mantenimiento de la vivienda de uso permanente, gasto en hotel o pensión y pagos por servicios de agua potable y electricidad, fue el que más dificultades generó por la escasez de información sobre estos ítems.

Por último se agrupó en una quinta variable el gasto en otros bienes no durables y servicios. Bajo este nombre se contabilizaron los gastos de las familias en: combustibles, lubricantes, lavado y engrase de vehículos, artículos para la limpieza del hogar, artículos varios para mantenimiento, elementos para primeros auxilios, material escolar y técnico, artículos de tocador, diarios y revistas, seguros, servicios domésticos, servicios de reparación y mantenimiento del hogar, internación y cirugías, servicios auxiliares para la salud, transporte público y por turismo, correo, teléfono, alojamiento, estadía y paquete turístico, servicios recreativos y culturales, servicios deportivos, servicios para el cuidado personal y otros.

La Tabla 2.2. muestra el peso de cada tipo de gasto en el ingreso familiar, clasificando a los hogares por fuente de ingreso. Los hogares cuentapropistas gastan el 31% de su ingreso total en alimentos, mientras que los hogares empleadores destinan a este tipo de bienes el 18% de su ingreso. Esta diferencia se achica a medida que nos movemos hacia la derecha en la Tabla 2.2., hasta llegar al gasto en servicios

y no durables en que los empleadores destinan una mayor parte de su ingreso que los cuentapropistas a este tipo de consumos.

Tabla 2.2.
Relación Gasto - Ingreso de los hogares por grupos de bienes y clasificación de los hogares
a partir de frecuencias expandidas

más del 25% del ingreso familiar por fuente...	Alimentos	Alimentos y Bebidas	Alimentos, Bebidas y Tabaco	Commitments (1)	Servicios y No Durables (2)
Empleadores (E)	18,0%	25,9%	26,8%	7,0%	21,2%
Otros	29,8%	38,3%	39,9%	9,4%	19,6%
Cuentapropistas (C)	31,1%	39,4%	41,2%	9,1%	19,5%
Otros	28,4%	36,9%	38,4%	9,3%	19,7%
E y C	28,8%	37,0%	38,7%	8,8%	19,8%
Otros	29,3%	37,8%	39,4%	9,5%	19,6%

(1) Incluye gastos en: alquiler (o valor locativo de la vivienda), hotel o pensión, reparación y mantenimiento de la vivienda, servicios de agua y electricidad.

(2) Incluye gastos en: combustibles, artículos de limpieza, servicio doméstico, servicios de reparación y mantenimiento del hogar, transporte público, correo, teléfono, servicios recreativos y culturales, etc.

Fuente: Elaboración propia en base a ENGH.

En cuanto a las características de los hogares, se utilizaron variables vinculadas con: (i) *Tamaño y composición de la familia*: cantidad de miembros, cantidad de adultos equivalentes y cantidad de adultos equivalente ajustados por economías de escala internas al hogar⁸, cantidad de hijos (incluyendo cantidad de hijos al cuadrado) número de chicos entre 0 y 4 años y número de chicos entre 5 y 16 años; si bien se probaron todas las opciones, aquí se presentan los resultados utilizando el total de miembros del hogar y las dos variables que reflejan la edad de los chicos de cada hogar. (ii) *Características de los jefes*: se consideró la edad de los jefes construyendo una variable que toma como valor medio 40 años, es decir, se calculó la distancia de la edad del jefe respecto a esa edad de referencia (y se usó la misma variable al cuadrado); también se intentó controlar por el nivel de instrucción de los jefes, pero los resultados aquí presentados no incluyen esta variable (como se verá en el capítulo 3, esta variable se utilizó para instrumentar el ingreso). (iii) *Características de la vivienda y otras*: número de habitaciones (excluyendo baños y cocina), número de baños, material de la casa (paredes, techo y piso), cochera, jardín,

⁸ Para la conversión de adultos equivalentes se utilizó la clasificación del INDEC y se tomó un coeficiente de economías de escala de 0,8, lo cual implica economías de escala moderadas.

teléfono, calefacción central, gas de red, agua caliente, distancia al transporte más cercano, hogares con pagos pendientes de la vivienda. Además, para los departamentos, se construyeron variables que consideran la cantidad de pisos del edificio, la disponibilidad de portero eléctrico y de servicio de seguridad. Si bien se probaron modelos con estas variables y a pesar de que Pissarides y Weber (1989) encuentran que algunas de estas variables tienen poder explicativo en el caso de Gran Bretaña, los resultados aquí reportados no incluyen variables que caracterizan a la vivienda. En la sección siguiente se muestra una tabla con algunas estadísticas descriptivas de las variables utilizadas en la regresión.⁹

De este modo, la muestra queda conformada por 3.402 hogares, que utilizando el factor de expansión de la ENGH, representan a 2.319.777 hogares de la Región Metropolitana del Gran Buenos Aires en 1996/97.

2.3. Resultados empíricos

La Tabla 2.3. muestra algunas estadísticas descriptivas para la muestra utilizada en la estimación.

⁹ En el Anexo I se presenta una tabla que incluye la descripción de todas las variables consideradas.

Tabla 2. 3.
Estadísticas descriptivas de la muestra utilizada en la regresión
a partir de frecuencias expandidas

	más del 25% del ingreso familiar por fuente...									
	Empleadores		No Empleadores *		Cuentapropistas		Empleadores y Cuentapropistas		No Empleadores ni Cuentapropistas *	
	Media	Desvío Estándar	Media	Desvío Estándar	Media	Desvío Estándar	Media	Desvío Estándar	Media	Desvío Estándar
<i>Consumos (\$)</i>										
Alimentos	357,0	191,4	257,0	163,2	260,7	154,1	278,0	165,6	255,7	166,8
Alimentos y Bebidas	551,9	329,1	356,6	243,6	354,3	227,7	389,9	260,0	357,8	250,1
Alimentos, Bebidas y Tabaco	571,6	330,9	371,2	250,4	370,2	234,6	406,1	265,9	371,9	256,7
Commitments (1)	177,8	371,0	108,9	209,9	104,8	183,2	118,3	231,2	110,5	219,4
Servicios y No Durables (2)	579,9	634,9	237,8	359,2	249,9	475,4	307,8	523,6	234,1	300,8
<i>Ingresos (\$)</i>										
Ingreso total del hogar	3.019,5	2.346,0	1.299,6	1.331,3	1.318,3	1.394,9	1.620,5	1.726,5	1.296,6	1.314,7
Ingreso por fuente que clasifica al hogar como tal	2.294,2	1.906,4	1.296,3	1.318,7	919,4	1.124,6	1.187,1	1.426,9	1.270,3	1.268,0
Ingreso per cápita familiar	847,9	775,4	369,7	396,4	362,7	405,2	449,1	524,5	373,6	395,5
<i>Características del hogar usadas en la regresión</i>										
Cantidad de miembros	4,03	1,52	4,05	1,80	4,26	1,78	4,22	1,74	3,96	1,81
Edad del jefe - 40 (age)	8,53	11,40	8,80	14,73	6,50	11,51	6,86	11,53	9,73	15,73
Age al cuadrado (agesq)	202,15	270,19	294,31	407,77	174,55	249,61	180,03	254,68	341,89	446,77
nº de chicos e/ 0 y 4 años	0,31	0,58	0,36	0,68	0,36	0,68	0,34	0,66	0,37	0,68
nº de chicos e/ 5 y 16 años	0,83	1,12	0,89	1,19	1,07	1,20	1,03	1,19	0,82	1,18

Fuente: Elaboración propia en base a ENGH

2.3.1. Estimación de modelos de consumo

Un primer punto metodológico a sortear es la posible no-normalidad de las variables de interés. Teniendo en cuenta que los datos provienen de encuestas de hogares, y dadas las características de la información relevada, es lógico conjeturar que la posible presencia de valores atípicos (*outliers*) sea un problema no menor para los métodos estándar, tales como el de mínimos cuadrados ordinarios (MCO). Bajo esta premisa, se optó por estimar los parámetros del modelo descrito en la sección 2.1. utilizando *métodos robustos*, que permiten obtener estimaciones consistentes y que además producen inferencias válidas aún bajo la presencia de outliers¹⁰.

A los efectos de explorar los inconvenientes del método de mínimos cuadrados (MCO), se estimaron, usando este método, varias versiones de la ecuación (7) similares a las presentadas en las Tablas 2.4. y 2.5. más adelante (las cuales pueden verse en el Anexo II.2.), y posteriormente se implementó el test de normalidad de Jarque y Bera (1987) sobre los residuos de los mismos, en todos los casos rechazando en forma contundente la hipótesis nula de normalidad. A modo de ejemplo, el estadístico de Jarque y Bera para el modelo de la última columna de la Tabla 2.4. arroja un valor igual a 2318,31, notoriamente mayor que los valores críticos usuales de la distribución Chi cuadrado con 2 grados de libertad (distribución asintótica del estadístico bajo la hipótesis nula). Esto muestra claramente la inadecuación del método de MCO para este tipo de problemas.¹¹

Los resultados de las estimaciones se resumen en las Tablas 2.4. y 2.5., las cuales presentan varias versiones del modelo correspondiente a la ecuación (7) en la sección 2.1. En la Tabla 2.4. se exploran distintas clasificaciones de hogares y en la Tabla 2.5. se explora el efecto de variar la canasta de consumo.

¹⁰ El método utiliza la rutina **rreg** de Stata 6.0 que realiza una reponderación basada en las distancias “D” de Cook. Posteriormente se itera utilizando ponderaciones de Huber y del tipo “biweight”. Ver Wilcox (1997) para una discusión detallada de estos procedimientos.

¹¹ También se realizaron estimaciones utilizando *regresiones cuantílicas* (Buchinsky, 1998), que proveen estimaciones robustas frente a outliers en la dirección de la variable explicada, obteniendo resultados muy similares a los producidos por el método robusto de Huber.

Tabla 2.4.**Estimaciones de varias versiones del modelo – variable dependiente: gasto en alimentos***Estimación robusta - gastos e ingreso en logaritmos*

	(1)	(2)	(3)	(4)	(5)
Ingreso	0,3741 (26,81)	0,3683 (33,37)	0,3629 (31,91)	0,3660 (31,30)	0,3627 (31,93)
Cantidad de miembros	0,0635 (6,88)	0,0639 (6,94)	0,6440 (6,99)	0,0640 (6,95)	0,0643 (6,98)
Edad del jefe - 40 (age)	0,0056 (4,77)	0,0056 (4,78)	0,0056 (4,74)	0,0056 (4,75)	0,0056 (4,79)
Age al cuadrado (agesq)	-0,0010 (-2,72)	-0,0010 (-2,77)	-0,0010 (-2,82)	-0,0010 (-2,88)	-0,0010 (-2,94)
nº de chicos e/ 0 y 4 años	0,0019 (0,11)	0,0016 (0,09)	0,0002 (0,01)	0,0003 (0,02)	0,0001 (0,00)
nº de chicos e/ 5 y 16 años	0,0318 (2,64)	0,0314 (2,61)	0,0310 (2,58)	0,0316 (2,62)	0,0311 (2,59)
Empleadores o cuentapropistas (no asalariados) EoC	0,1266 (0,82)	0,0258 (1,36)			
EoC x ingreso	-0,0145 (-0,66)				
Cuentapropistas			0,0104 (0,52)		
Empleadores			0,0866 (2,26)	0,4912 (1,32)	0,0839 (2,21)
Empleadores x ingreso				-0,0533 (-1,11)	
Constante	2,5507 (25,64)	2,5899 (31,93)	2,6276 (31,59)	2,6111 (30,85)	2,6328 (31,92)

Nota: debajo de cada coeficiente se reporta entre paréntesis el estadístico z correspondiente
Fuente: Elaboración propia en base a ENGH.

La Tabla 2.4. presenta estimaciones de un modelo log lineal de consumo de alimentos en función del (logaritmo del) ingreso, un indicador de grupo de ingresos y un conjunto representativo de variables socioeconómicas determinantes del consumo (cantidad de miembros, edad del jefe y su cuadrado, número de hijos entre 0 y 4 años y entre 5 y 16 años).

Inicialmente se estimó una versión de la ecuación (7) utilizando como grupo de ingresos potencialmente subdeclarados a los “no asalariados”, definidos de acuerdo al criterio descrito en la sección anterior. Se incluyeron variables binarias de intercepto y pendiente (columna 1), resultando ambas no significativas. Esto se debe fundamentalmente a la exagerada correlación muestral entre ambas (0,998), por lo que se procedió a dejar sólo la variable binaria de intercepto (columna 2),

aumentando su significatividad notoriamente, aunque estadísticamente sólo al 17% de confianza.

La columna 3 diferencia entre cuentapropistas y empleadores, y en este caso, la variable binaria por hogar empleador es significativa mientras que la de cuentapropista no lo es. Este resultado es llamativo, en particular en comparación con los resultados de Smith et al. (1986). En dicho trabajo, los autores diferencian entre categorías laborales del tipo *white-blue collar*, estimando modelos distintos para ambos grupos, lo cual es notoriamente complejo de replicar en base a la información de la ENGH. Implementando esta distinción, los resultados de Smith et al. (1986) sugieren que la significatividad de la variable binaria “no asalariado” es notoriamente mayor para el caso del grupo “white collar”, lo cual es en parte consistente con los resultados de la columna 3. De todos modos, la no significatividad de la variable “cuentapropista” sugiere que, bajo el esquema de la sección 2.1., a los efectos de la subdeclaración, la distinción relevante es entre empleadores y no empleadores. Consecuentemente, las columnas 4 y 5 estiman el modelo correspondiente a la ecuación (7) tomando como grupo de referencia al de “empleadores” solamente, y la misma apreciación vale en relación a la alta colinealidad entre las variables binarias por intercepto y pendiente.

Es importante recordar que la inclusión del resto de las variables sociodemográficas tiene como objetivo evitar sesgos por omisión de variables en la estimación de la elasticidad ingreso y de la variable binaria del grupo de referencia. Se probaron varias alternativas a fines de controlar las diferencias en el consumo de acuerdo a distintas medidas de estructura familiar. Tal como se describe en la sección 2.2., se utilizaron distintas nociones de tamaño de familia, así como una gran variedad de características sociodemográficas y de las viviendas habitadas, que podrían explicar las diferencias en el consumo entre familias. Ninguna de estas especificaciones alternativas difiere significativamente de los resultados presentados en la Tabla 2.4., y por razones de espacio se presentan estos resultados en el Anexo II.3.

La Tabla 2.5. muestra varias estimaciones del modelo utilizando la misma estructura de variables explicativas que el modelo de la columna 5, pero para distintos grupos de bienes: alimentos y bebidas, alimentos, bebidas y tabaco, commitments y servicios y bienes no durables. Como se mencionara anteriormente, el objetivo

consiste en explorar la sensibilidad de los resultados obtenidos al uso de canastas de bienes alternativas con distinta elasticidad ingreso.

Tabla 2.5.
Estimaciones de varias versiones del modelo - variable binaria: empleadores
Estimación robusta - gastos e ingreso en logaritmos

	Alimentos y bebidas	Alimentos, bebidas y tabaco	Commitments	Servicios & No Durables
Ingreso	0,4780 (44,38)	0,4758 (44,49)	0,6286 (26,72)	0,8677 (54,72)
Cantidad de miembros	0,0617 (7,07)	0,0639 (7,36)	-0,1151 (-6,20)	0,0205 (1,61)
Edad del jefe - 40 (age)	0,0030 (2,68)	0,0027 (2,45)	-0,0116 (-4,80)	-0,0003 (-0,21)
Age al cuadrado (agesq)	-0,0001 (-3,37)	-0,0002 (-3,92)	0,0002 (1,93)	-0,0002 (-2,95)
nº de chicos e/ 0 y 4 años	-0,0224 (-1,33)	-0,0278 (-1,66)	0,0762 (2,09)	-0,0095 (-0,39)
nº de chicos e/ 5 y 16 años	0,0023 (0,20)	-0,0035 (-0,31)	0,0361 (1,48)	-0,0415 (-2,51)
Empleadores	0,0816 (2,27)	0,0717 (2,00)	-0,1047 (-1,37)	0,1105 (2,12)
Constante	2,2139 (28,30)	2,2758 (29,30)	-0,0493 (-0,29)	-0,9161 (-7,95)

Nota: debajo de cada coeficiente se reporta entre paréntesis el estadístico z correspondiente
Fuente: Elaboración propia en base a ENGH.

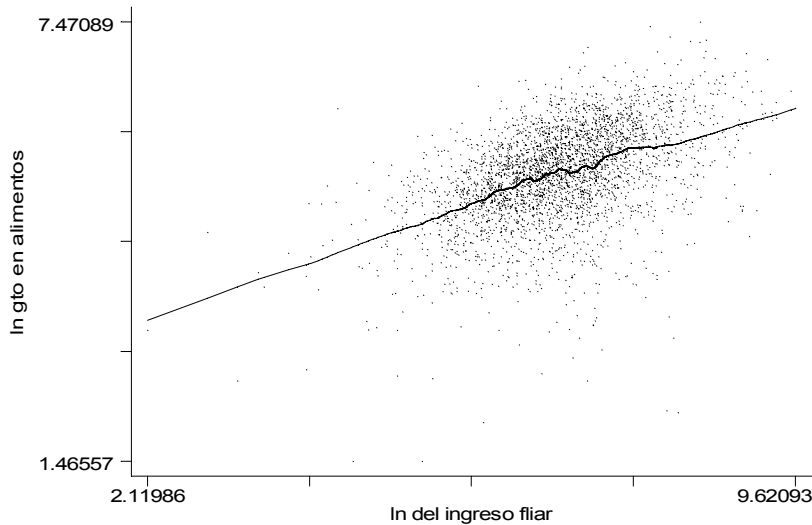
La elasticidad ingreso del consumo de alimentos (0,3627) es estimada con gran precisión y arroja valores superiores a los estimados por Smith (1986) para Gran Bretaña (obtiene un valor cercano a 0,26). Como era de esperar, para las otras canastas de bienes la elasticidad ingreso es notoriamente superior: 0,47 para alimentos y bebidas y alimentos, bebidas y tabaco, 0,68 para commitments y 0,86 para servicios y otros bienes no durables. La variable binaria empleador es siempre significativa y con el signo esperado, salvo en el caso de commitments, en donde no es significativamente distinta de cero.

A fines de explorar si el supuesto de log-linealidad es apropiado, se estimaron regresiones no-paramétricas simples entre los logaritmos del ingreso y consumo, las cuales sugieren que, para este caso, el mismo no es demasiado restrictivo. El Gráfico 2.3. muestra los resultados de estimar en forma no paramétrica una función simple de

consumo de alimentos para los no empleadores utilizando métodos no paramétricos estándar, lo cual sugiere que la aproximación lineal es razonable¹².

Gráfico 2.3.

Log linealidad de la función consumo de alimentos



2.3.2. Estimaciones de la subdeclaración

La Tabla 2.6. muestra las estimaciones de la tasa de subdeclaración que surgen de los distintos modelos de consumo estimados anteriormente. Las mismas se obtienen de reemplazar los coeficientes estimados en la fórmula (8).

¹² Se utilizó un estimador *lowess* con ancho de banda 0.1 (ver Pagan y Ullah, 1999 para más detalles). Esta aproximación es decididamente simplista ya que ignora el efecto de las otras variables explicativas. Como es bien conocido en la literatura, la estimación general no paramétrica de un modelo de regresión con varias variables explicativas requiere una enorme cantidad de datos (el problema de “maldición de la dimensionalidad”). Lyssiotou et al. (1999) estiman un modelo parcialmente lineal para sortear esta dificultad.

Tabla 2.6.**Estimaciones de la tasa de subdeclaración para varios modelos**

	Alimentos	Alimentos y Bebidas	Alimentos, Bebidas y Tabaco	Commitments	Servicios y No Durables
Elasticidad ingreso (b2)	0,3627	0,4780	0,4758	0,6286	0,8677
Dummy (b1)	0,0839	0,0816	0,0717	-0,1047	0,1105
<i>S (b1 / b2)</i>	<i>0,231</i>	<i>0,171</i>	<i>0,151</i>	<i>-0,167</i>	<i>0,127</i>
Error Estándar	0,0975	0,0685	0,0707	0,1260	0,0638
Intervalo de Confianza 95%					
<i>Límite inferior</i>	<i>0,041</i>	<i>0,029</i>	<i>0,004</i>	<i>-0,406</i>	<i>0,010</i>
<i>Límite superior</i>	<i>0,419</i>	<i>0,312</i>	<i>0,289</i>	<i>0,087</i>	<i>0,269</i>

Fuente: Elaboración propia en base a ENGH.

La estimación puntual de la subdeclaración para el modelo de alimentos sugiere que los hogares clasificados como empleadores ocultan el 23% de sus ingresos. Los resultados para otras canastas de bienes arrojan valores entre 13% y 17%. Es importante remarcar que la posibilidad de identificar la subdeclaración en los ingresos se basa crucialmente en utilizar canastas de baja elasticidad ingreso, de modo que los valores que surgen del modelo de alimentos posiblemente provean las estimaciones más confiables de la subdeclaración.¹³

Smith et al. (1986) obtienen valores mayores para la estimación basada en el grupo de alimentos (36%), pero concluyen que la subdeclaración promedio para las distintas canastas consideradas en el caso de Gran Bretaña es de 18%. Este promedio de estimaciones de tasas de subdeclaración para el caso del Área Metropolitana Gran Buenos Aires (descartando la de commitments, que no es significativa) arroja una cifra sorprendentemente similar, igual a 17%.

Una cuestión importante, y no explorada en Smith et al. (1986) y los trabajos posteriores, consiste en evaluar la relevancia estadística de estas magnitudes. La tasa estimada de subdeclaración, de acuerdo a la fórmula (8) es una transformación no-lineal de los coeficientes estimados, lo que dificulta la aplicación de métodos estándar de inferencia. Consecuentemente, se optó por aplicar una técnica de remuestreo (*bootstrap*) a fines de aproximar la variabilidad muestral de las estimaciones

¹³ También se realizaron estimaciones de la tasa de subdeclaración utilizando variables binarias por pendiente en vez de por intercepto, obteniéndose resultados muy similares.

anteriores¹⁴. La parte inferior de la Tabla 2.6. muestra los errores estándar calculados para las tasas de subdeclaración y los intervalos de confianza al 95%. Salvo en el caso de commitments, en todos los casos el intervalo de confianza no incluye al cero, por lo que las tasas de subdeclaración resultan estadísticamente significativas, aunque son estimadas con poca precisión, por ejemplo, para la caso de alimentos, el intervalo de confianza va de 4% a 41,8%.

2.3.3. Interpretación de los resultados obtenidos

Como se discutiera en el capítulo 1, los resultados de cualquier estimación indirecta de la economía informal reflejan en parte la magnitud del fenómeno a medir y en parte los supuestos adoptados para poder llevar a cabo dicho ejercicio. De modo que cualquier medición de la economía informal debe ser tomada con cierta precaución.

Un primer punto a remarcar es que las mediciones obtenibles a partir del método utilizado en este trabajo se refieren a un aspecto particular de la economía informal, relacionado con el volumen de ingresos que un grupo de agentes “esconde” de las estadísticas oficiales. Esto dificulta la comparación con otros resultados existentes para el caso argentino, que miden conceptos distintos.

En cuanto a la relevancia de los supuestos en los que se basa la estimación, posiblemente el más importante se refiere a que sea posible explicar la mayor parte de las heterogeneidades familiares en el consumo, de modo que las diferencias en los patrones de consumo entre los grupos considerados (captadas por la variable binaria de pertenencia a un grupo particular de ingresos) se deban a la subdeclaración y no a otras diferencias no observables entre el consumo de los grupos. En este trabajo se experimentó con una considerable cantidad de características sociodemográficas que permitirían controlar por estas diferencias, aunque es todavía posible explorar otras que no son directamente observables en la ENGH.

Un supuesto importante de identificación es que exista un grupo “base” que no subdeclara, de modo que la tasa de subdeclaración pueda interpretarse en términos absolutos. Trivialmente, si el grupo base también subdeclara, el método estimado

¹⁴ Se realizó un bootstrap no paramétrico con 300 replicaciones de la rutina **rreg**. Los intervalos de confianza se calcularon de acuerdo al método de percentiles. Ver Sosa Escudero

permite identificar cuánto el otro grupo subdeclara por sobre el monto ya subdeclarado por el grupo base. De ser este el caso, las tasas de subdeclaración estimadas tendrían una interpretación en términos relativos: cuánto más subdeclara un grupo con respecto al otro, resultando imposible identificar la subdeclaración total.

Otro punto relevante se refiere a la representatividad de la muestra utilizada. Como es habitual en este tipo de encuestas complejas, la tasa de no respuesta no es trivial, y es perfectamente posible que aquellos hogares que más tienen para ocultar sean los que tengan mayor incentivos a negarse a completar la encuesta. En principio, la posibilidad de extrapolar resultados muestrales a la población descansa en el supuesto implícito de que la tasa de subdeclaración de los individuos en la muestra es similar a la de los individuos fuera de la misma. Si no se contara con ninguna otra información adicional acerca de este comportamiento, la tasa de subdeclaración real fluctuaría de acuerdo a que la subdeclaración de los individuos que no están en la muestra sea notoriamente mayor o menor que la de los que si lo están.¹⁵ El comentario anterior sugiere que la tasa de subdeclaración de las familias que no están en la muestra es mayor a la de las que si están, de modo que las tasas de subdeclaración estimadas con los datos disponibles serían límites inferiores de las tasas reales. Por otro lado, Smith et al. (1986) argumentan que el crecimiento en la actividad informal se da a través de actividades relativamente pequeñas, distribuidas en un gran número de hogares, de modo que los resultados se verían afectados si la tasa de no respuesta fuese exorbitantemente grande. Teniendo en cuenta la complejidad de la ENGH, la tasa de respuesta es llamativamente elevada (entre 60% y 80%, aunque con cifras cercanas al 50% en el Gran Buenos Aires, ver Indec (1998, pp. 27) similares a las presentes en la encuesta utilizada por Smith et al. (1986)), lo cual amortiguaría este potencial problema.

y Gasparini (2000) para una descripción detallada de estos métodos y una aplicación al caso de la variabilidad de los coeficientes de Gini en la Encuesta Permanente de Hogares.

¹⁵ Los extremos corresponden al caso en que todas las familias que no están en la muestra subdeclaran o que ninguna lo hace. El enfoque de cotas de Manski (*Manski bounds*, Manski, 1994) se basa en esta apreciación.

Capítulo 3

Extensiones del modelo básico

Este capítulo presenta una extensión del “modelo básico” tratado en el capítulo 2 basada en Pissarides y Weber (1989). Con respecto a dicho modelo, esta versión avanza en dos sentidos: (i) permite que la subdeclaración varíe por individuos, de modo que el modelo estima una *tasa promedio* de subdeclaración; y (ii) modela explícitamente la relación entre ingreso corriente e ingreso permanente admitiendo diferencias en las volatilidades entre los ingresos del grupo que subdeclara y del que no lo hace. Asimismo, la estimación se lleva a cabo utilizando mínimos cuadrados en dos etapas (MC2E), lo cual elimina el problema derivado de la endogeneidad del ingreso. Como se verá más adelante, éste se instrumenta utilizando algunas características de los jefes de familia, como por ejemplo la edad y el nivel de instrucción de los mismos.

3.1. El modelo

En la sección 2.1. la tasa de subdeclaración se plantea en la ecuación (4) como la diferencia de los logaritmos del ingreso real y del reportado por las familias. Aquí, la tasa de subdeclaración se denota con la letra k y, a diferencia del modelo básico, es una variable aleatoria que muestra el tamaño de la subdeclaración de ingresos del individuo i , k_i . Es decir,

$$Y_i = k_i * Y_i^R \quad (9)^{16}$$

Al igual que en el modelo anterior, el supuesto básico es que A reporta correctamente su ingreso ($k_i^A = 1$) mientras que E subdeclara ($k_i^E > 1$).

Adicionalmente, se sigue suponiendo que ambos grupos reportan correctamente su consumo, pero en este caso se plantea explícitamente que el mismo es una función de su ingreso permanente, y está dado por:

$$\ln C_i = Z_i \alpha + \beta \ln Y_i^P + \varepsilon_i \quad (10)$$

Y_i^p es el ingreso permanente y se relaciona con el ingreso real a través del parámetro p_i , que mide los desvíos con respecto al ingreso permanente, de modo tal que

$$Y_i = p_i Y_i^p \quad (11)$$

Al igual que k_i , p_i es una variable aleatoria que varía para cada individuo i . El supuesto clave en este sentido es la mayor volatilidad del ingreso real de los empleadores respecto al de los asalariados. Aplicando logaritmos a (9) y (11) se obtiene que el ingreso permanente puede expresarse en función su volatilidad, del ingreso reportado y de la tasa de subdeclaración.

$$\ln Y_i^p = \ln Y_i^R - \ln p_i + \ln k_i \quad (12)$$

Dado que p_i y k_i no son observables se asume que ambos tienen una distribución log normal con:

$$\ln p_i = \mu_p + u_i, \text{ con } E(u_i) = 0 \quad (13)$$

$$\ln k_i = \mu_k + v_i, \text{ con } E(v_i) = 0 \quad (14)$$

Dado que se supone que el grupo A no subdeclara, $k^A = 1$ ($\ln k^A = 0$), en términos de la ecuación (14) esto implica que $\mu_k^A = 0$ y $\sigma_v^A = 0$

Por otra parte, si p_i es log normal con $E(p_i) = p^-$, entonces:

$$\ln p^- = \mu_p + \frac{1}{2} \sigma_u^2 \quad (15)$$

Si el ingreso de empleadores y asalariados es igual en promedio y difieren únicamente en la volatilidad de los mismos, se tiene que

$$E(p^{-E}) = E(p^{-A})$$

$$\mu_p^E + \frac{1}{2} \sigma_u^{2E} = \mu_p^A + \frac{1}{2} \sigma_u^{2A}$$

¹⁶ Alternativamente, puede escribirse como (4) haciendo $\ln Y_i^R = \ln Y_i - \ln k_i$

$$\mu_p^E - \mu_p^A = -1/2 (\sigma_u^2{}^E - \sigma_u^2{}^A) \quad (16)$$

siendo $(\sigma_u^2{}^E - \sigma_u^2{}^A) > 0$ porque p es más volátil para E.

Reemplazando (13) y (14) en (12) y ésta a su vez en (10), el modelo estimable para E y A es:

$$\ln C_i = Z_i \alpha + \beta \ln Y_i^R - \beta (\mu_p^E - \mu_p^A) - \beta (u_i - v_i) + \varepsilon_i \quad (17)$$

Siguiendo con la lógica del modelo básico, se asume que los determinantes del consumo no difieren por grupo de ingreso, es decir, que $\alpha^E = \alpha^A = \alpha$ y $\beta^E = \beta^A = \beta$

Entonces, los parámetros de (17) para E y A pueden ser estimados a través de una única ecuación en la que se introduce la variable binaria D_i , que adopta un valor igual a 1 si el individuo pertenece al grupo que subdeclara ingresos y 0 en caso contrario.

$$\ln C_i = Z_i \alpha + \beta \ln Y_i^R + \gamma D_i + \eta_i \quad (18)$$

Al estimar (18) deben tenerse en cuenta dos cuestiones:

- (i) esta ecuación es heterocedástica ya que la varianza de η_i cambia con el valor que adopta D_i ;
- (ii) γ estima la diferencia en los modelos para E y A. Dado que se asume que los asalariados no subdeclaran ($k^A = 1$), la constante para este grupo es $-\beta (\mu_p^A - \mu_k^A) = -\beta \mu_p^A$, mientras que para los empleadores, la constante es $-\beta (\mu_p^E - \mu_k^E)$. Entonces

$$\begin{aligned} \gamma &= -\beta (\mu_p^E - \mu_k^E) + \beta \mu_p^A \\ \gamma &= \beta (\mu_p^A - \mu_p^E + \mu_k^E) \end{aligned}$$

reemplazando por la ecuación (16) se llega a:

$$\gamma = \beta (1/2 (\sigma_u^2{}^E - \sigma_u^2{}^A) + \mu_k^E) \quad (19)$$

$$\mu_k^E = \gamma / \beta - 1/2 (\sigma_u^2 E - \sigma_u^2 A) \quad (20)$$

El objetivo es estimar la subdeclaración promedio que está dada por k^E . Como los únicos que subdeclaran son los E, se busca estimar la subdeclaración promedio de este grupo de individuos, k^E , que por el supuesto de log normalidad:

$$\ln k^E = \mu_k^E + 1/2 \sigma_v^2 \quad (21)$$

reemplazando (20) en (21)

$$\ln k^E = \gamma / \beta + 1/2 (\sigma_v^2 - \sigma_u^2 E + \sigma_u^2 A) \quad (22)$$

El punto es estimar (22). El primer término del miembro derecho se obtiene a partir de (18). El segundo término requiere algunos cálculos y supuestos adicionales. Como se mencionó, uno de los avances de esta versión es la estimación a través de MC2E, eligiendo algunas variables como instrumentos del ingreso de las familias. El ingreso reportado es estimado entonces a partir de la siguiente ecuación, donde X_i es el conjunto de variables instrumentales.

$$\ln Y_i^R = Z_i \delta_1 + X_i \delta_2 + \zeta_i \quad (23)$$

De esta ecuación puede deducirse el segundo término de (22): dado que las varianzas de u y v son desconocidas, deben hacerse inferencias a partir de la varianza residual del ingreso estimado.

Teniendo en cuenta la ecuación (12), se observa que el residuo de (23), ζ_i , está compuesto por tres errores: las variaciones no explicadas del ingreso permanente, las variaciones del ingreso real respecto del permanente (u_i) y las variaciones del ingreso real respecto del reportado (v_i). Si se asume que las variaciones no explicadas del ingreso en (23) tienen la misma varianza para E y A y se tiene en cuenta que A no subdeclara puede decirse que:

$$\text{var} \zeta_i^E - \text{var} \zeta_i^A = \text{var} (u - v)^E - \text{var} u^A \quad (24)$$

reexpresando esto en función de la notación usada más arriba:

$$\text{var}\zeta_i^E - \text{var}\zeta_i^A = \sigma_u^2{}^E + \sigma_v^2{}^E - 2 \text{cov}(uv)^E - \sigma_u^2{}^A \quad (25)$$

Para estimar el segundo término de (22) a través de (25) se necesita conocer la correlación entre u y v para el grupo que potencialmente subdeclara sus ingresos, es decir, la correlación entre la volatilidad del ingreso real respecto al permanente y la volatilidad de la subdeclaración de ingresos. En particular, si la covarianza entre ambas es cero, puede encontrarse un rango pequeño para la subdeclaración promedio, pero si la correlación es distinta de cero surgen algunas dificultades.

El valor de esta covarianza depende de los supuestos que se hagan sobre la propensión de los individuos a subdeclarar ingresos. Es factible suponer que si un “no asalariado” tiene un año excepcionalmente bueno, estará menos propenso a declarar dicho ingreso que si proviniera de una fuente de ingresos regular. En cambio, si esa persona tiene un año excepcionalmente malo, posiblemente esté dispuesto a declararlo ante las autoridades impositivas. Comportamientos de este tipo llevan a una covarianza positiva entre u y v .

Una covarianza nula requiere que sin importar el ingreso de una persona que potencialmente subdeclara sus ingresos, su tendencia sea a subdeclarar siempre el mismo porcentaje a las autoridades impositivas. En este punto cabe recalcar que las razones que llevan a un individuo a subdeclarar ingreso frente a un encuestador pueden diferir de las razones de subdeclaración frente a las autoridades impositivas. En particular, es posible que algunas personas declaren incorrectamente su ingreso (en más o en menos) simplemente porque no están interesados en responder con sinceridad, porque no tienen tiempo de calcular el monto exacto de sus ingresos, entre otros factores. Por eso, puede ser razonable asumir que la covarianza entre u y v sea cero en este caso.

Bajo el supuesto de no correlación, de (22) y (25') debe obtenerse una estimación de la subdeclaración promedio.

$$\ln k^{-E} = \gamma / \beta + 1/2 (\sigma_v^2 - \sigma_u^2{}^E + \sigma_u^2{}^A) \quad (22)$$

$$var\zeta_i^E - var\zeta_i^A = \sigma_v^2 E + \sigma_u^2 E - \sigma_u^2 A \quad (25')^{17}$$

En el caso extremo en que los individuos del grupo E subdeclararan el mismo porcentaje de sus ingresos ($\sigma_v^2 E = 0$), (25') podría reemplazarse en (22) para obtener un límite inferior de la subdeclaración promedio:

$$\ln k^{-E} = \gamma / \beta - 1/2 (var\zeta_i^E - var\zeta_i^A) \quad (27)$$

Por otro lado, recordando el supuesto de que la volatilidad de los ingresos del grupo E es mayor que la del grupo A, podría llevarse al extremo en que las varianzas de ambos grupos sean iguales ($\sigma_u^2 E = \sigma_u^2 A$). En este caso se llega a un límite superior de la tasa de subdeclaración promedio dada por:

$$\ln k^{-E} = \gamma / \beta + 1/2 (var\zeta_i^E - var\zeta_i^A) \quad (28)$$

De este modo se concluye que el modelo así planteado estima una tasa de subdeclaración promedio del grupo E respecto al grupo A a partir de las ecuaciones (18) y (23). Los resultados se reemplazan en (27) y (28) y aplicando antilogaritmo se obtiene un rango para k^{-E} , por el cual debe multiplicarse el ingreso promedio reportado para llegar al ingreso real del grupo E.

3.2. Resultados empíricos

3.2.1. Estimación de modelos de consumo

Siguiendo el esquema utilizado en el capítulo 2, las Tablas 3.1 a 3.3 presentan los principales resultados obtenidos de las estimaciones de (18).¹⁸

La Tabla 3.1 muestra los resultados de estimar el modelo instrumentando el ingreso familiar por la edad y el nivel de instrucción del jefe del hogar. Las columnas 1 y 4 utilizan variables de intercepto y pendiente, con resultados disímiles. Con respecto a la variable de pendiente, cabe aclarar que mientras el (logaritmo del)

¹⁷ Supone $cov(uv)^E = 0$.

¹⁸ El método utiliza la rutina `ivreg` de Stata 6.0, con el agregado del comando de robustez, que arroja estimaciones robustas de los desvíos estándar de los coeficientes, pero no de los coeficientes en sí mismos.

ingreso es instrumentado, la variable de pendiente ($D * Y$) utiliza el (logaritmo del) ingreso reportado en la encuesta.

A diferencia de lo observado en el capítulo anterior, la columna 1 muestra que las variables binarias de intercepto y pendiente son ambas significativas cuando el grupo que potencialmente subdeclara ingresos son los “no asalariados”, aunque la variable binaria no tiene el signo esperado. En cambio, la columna 4 muestra que cuando el grupo que potencialmente subdeclara son los empleadores, si bien la variable binaria tiene el signo esperado, ésta y la variable de pendiente no son significativamente distintas de cero.

Tabla 3.1.

Estimaciones de varias versiones del modelo – variable dependiente: gasto en alimentos / instrumentos: edad (age) y nivel de instrucción (jnivins) del jefe

Estimación por MC2E - gastos e ingreso en logaritmos

	(1)	(2)	(3)	(4)	(5)
Ingreso (instrumentado)	0,2753 (8,97)	0,3083 (14,49)	0,3005 (13,40)	0,3028 (12,84)	0,3016 (13,44)
Cantidad de miembros	0,0911 (9,13)	0,0887 (9,05)	0,0894 (9,10)	0,0894 (9,07)	0,0895 (9,11)
n° de chicos e/ 0 y 4 años	-0,0847 (-5,00)	-0,0836 (-4,94)	-0,0849 (-5,00)	-0,0852 (-5,02)	-0,0851 (-5,02)
n° de chicos e/ 5 y 16 años	0,0050 (0,39)	0,0058 (0,45)	0,0057 (0,44)	0,0065 (0,50)	0,0064 (0,50)
Empleadores o cuentapropistas (no asalariados) EoC	-0,4975 (-2,01)	0,0390 (1,86)			
EoC x ingreso	0,0775 (2,17)				
Cuentapropistas			0,0206 (0,93)		
Empleadores			0,1154 (2,53)	0,3633 (0,84)	0,1095 (2,42)
Empleadores x ingreso				-0,0330 (-0,58)	
Constante	3,1394 (15,51)	2,9229 (20,69)	2,9743 (20,00)	2,9642 (19,00)	2,9719 (19,99)

Fuente: Elaboración propia en base a ENGH.

La columna 2 muestra lo que ocurre cuando sólo se utiliza la variable binaria de intercepto. La columna 3 analiza en forma separada el comportamiento de los hogares empleadores y de los cuentapropistas: al igual que en el modelo anterior, los cuentapropistas no muestran un patrón sistemático de respuesta distinto al de los que no lo son, mientras que esto sí se observa para los hogares empleadores. La columna

5 muestra el resultado de estimar la ecuación (18) utilizando únicamente la variable binaria de empleador, la cual presenta el signo esperado y es significativamente distinta de cero. Este modelo se toma como base para las comparaciones con las Tablas 3.2 y 3.3.

Tabla 3.2.

Estimaciones de varias versiones del modelo – variable binaria: empleadores / instrumentos: age y jnivins

Estimación por MC2E - gastos e ingreso en logaritmos

	variable dependiente			
	Alimentos y Bebidas	Alimentos, Bebidas y Tabaco	Commitments	Servicios y No Durables
Ingreso (instrumentado)	0,4587 (21,90)	0,4627 (22,35)	0,9253 (22,25)	1,1079 (35,25)
Cantidad de miembros	0,0745 (8,11)	0,0761 (8,37)	-0,1841 (-10,32)	-0,0098 (-0,72)
nº de chicos e/ 0 y 4 años	-0,0659 (-4,14)	-0,0673 (-4,28)	0,2412 (7,72)	0,0425 (1,80)
nº de chicos e/ 5 y 16 años	0,0009 (0,07)	-0,0007 (-0,06)	0,1190 (5,06)	0,0058 (0,33)
Empleador	0,0736 (1,74)	0,0640 (1,52)	-0,3535 (-4,30)	-0,0341 (-0,55)
Constante	2,2617 (16,31)	2,2712 (16,57)	-1,8937 (-6,85)	-2,6039 (-12,49)

Fuente: Elaboración propia en base a ENGH.

La Tabla 3.2 utiliza el mismo grupo de variables que las de la columna 5 de la tabla anterior, pero selecciona como variable dependiente distintos tipos de consumos. Las columnas 1 y 2 muestran qué sucede cuando se tiene en cuenta, además del gasto en alimentos, el gasto en bebidas y en bebidas y tabaco, respectivamente. En ambos casos se observa un sensible aumento de la elasticidad ingreso respecto a este tipo de bienes, que pasa de 0,3016 en el caso tomado como base a 0,4587 y 0,4627, respectivamente. Por último, la tabla muestra que al estimar la función consumo de otros gastos (commitments y servicios y no durables) la elasticidad ingreso aumenta y la variable binaria deja de tener el signo esperado (aunque en el caso del gasto en servicios y bienes no durables la misma resulta no ser significativamente distinta de cero).

Finalmente, la Tabla 3.3 estima distintas versiones del modelo de la columna 5 de la Tabla 3.1, cambiando las variables que instrumentan el ingreso. La tercer

columna muestra que la variable *age* no es un buen instrumento del ingreso, mientras que sí lo es el nivel de instrucción del jefe (columna 2). Cuando se utilizan las variables *age*, *agesq* y *jnivins* como instrumentos del ingreso, la elasticidad ingreso es levemente superior a la observada en el modelo utilizado como base y el coeficiente de la variable binaria es inferior con un menor grado de significatividad.

Tabla 3.3.

Estimaciones de varias versiones del modelo – variable dependiente: gasto en alimentos / variable binaria: empleadores

Estimación por MC2E - gastos e ingreso en logaritmos

	Instrumentos		
	age, agesq, jnivins	jnivins	age
Ingreso (instrumentado)	0,3154 (14,43)	0,3094 (13,73)	0,0471 (0,58)
Cantidad de miembros	0,0876 (8,95)	0,0884 (9,00)	0,1239 (8,25)
n° de chicos e/ 0 y 4 años	-0,0820 (-4,85)	-0,0834 (-4,92)	-0,1424 (-5,60)
n° de chicos e/ 5 y 16 años	0,0084 (0,66)	0,0075 (0,59)	-0,0314 (-1,73)
Empleador	0,0974 (2,16)	0,1027 (2,27)	0,3336 (3,95)
Constante	2,8826 (19,89)	2,9214 (19,57)	4,6245 (8,69)

Fuente: Elaboración propia en base a ENGH.

3.2.2. Estimaciones de la subdeclaración

Las Tablas 3.4 y 3.5 muestran las tasas de subdeclaración de ingresos *promedio* derivadas de una selección de los modelos antes presentados. La estimación se obtiene a partir de la ecuación (26). Como se mencionó en la sección 3.1., los coeficientes γ y β se obtienen de las estimaciones de (18) presentadas en las tablas 3.1. a 3.3. Para el cálculo de las varianzas del ingreso para los grupos E y A, se corrieron distintas versiones de la ecuación (23), acotando la muestra primero al grupo E y luego al A¹⁹.

¹⁹ En el Anexo II.4. se presentan las tablas con las estimaciones de estos modelos.

Tabla 3.4.**Estimaciones de la tasa de subdeclaración para varios modelos***Variable binaria: empleadores - Instrumentos: edad y nivel de instrucción del jefe*

	Alimentos	Alimentos y Bebidas	Alimentos, Bebidas y Tabaco	Commitments	Servicios y No Durables
Ingreso (instrumentado)	0,3016	0,4587	0,4627	0,9253	1,1079
Empleador	0,1095	0,0736	0,0640	-0,3535	-0,0341
varianza ingreso empleador (E)	0,4376	0,4376	0,4376	0,4376	0,4376
varianza ingreso no empleador (A)	0,4303	0,4303	0,4303	0,4303	0,4303
TASA DE SUBDECLARACIÓN					
<i>límite inferior</i>					
log de la subdeclaración (lnk)	0,3593	0,1568	0,1347	-0,3856	-0,0345
factor de subdeclaración (k)	1,4324	1,1698	1,1442	0,6800	0,9661
tasa de subdeclaración (k - 1)%	43%	17%	14%	-32%	-3%
<i>límite superior</i>					
log de la subdeclaración (lnk)	0,3666	0,1641	0,1420	-0,3783	-0,0272
factor de subdeclaración (k)	1,4429	1,1784	1,1525	0,6850	0,9732
tasa de subdeclaración (k - 1)%	44%	18%	15%	-31%	-3%

Fuente: Elaboración propia en base a ENGH.

Para la Tabla 3.4 se utilizó la misma regresión del ingreso, es decir que las varianzas del ingreso de uno y otro grupo coinciden. Lo que varía en cada caso es el gasto considerado como variable dependiente en la ecuación (18). Si bien se verifica que la varianza del ingreso de los empleadores es mayor que la de los asalariados, la diferencia entre ambas es pequeña, reduciendo el rango de variación de la tasa de subdeclaración promedio en todos los casos. La primer columna muestra el resultado para el gasto en alimentos, que se considera el más relevante debido a que es el que presenta menor elasticidad ingreso (0,30). Allí se observa una tasa de subdeclaración promedio que oscila entre el 43% y el 44%, cifra muy superior al 23% encontrado con el modelo básico del capítulo 2. Sin embargo, los resultados de las columnas 2 y 3 son similares a los encontrados anteriormente (Tabla 2.6.).

En la Tabla 3.5. se muestra el resultado de cambiar los instrumentos del ingreso, es decir que cada columna implica una regresión diferente de (23), con diferentes varianzas para el ingreso. Cuando el ingreso se instrumenta a través de la edad del jefe y su cuadrado y de su nivel de instrucción o únicamente a través de su

nivel de instrucción, se verifica el supuesto de mayor varianza del ingreso de los empleadores. En el último caso, en que el ingreso se instrumenta a través de la edad del jefe, el resultado se invierte y la tasa de subdeclaración obtenida no puede interpretarse²⁰.

Tabla 3.5.

Estimaciones de la tasa de subdeclaración para varios modelos
Variable binaria: empleadores - variable dependiente: gasto en alimentos

	Instrumentos		
	age, agesq, jnivins	jnivins	age
Ingreso (instrumentado)	0,3154	0,3094	0,0471
Empleador	0,0974	0,1027	0,3336
varianza ingreso empleador (E)	0,4398	0,4361	0,5765
varianza ingreso no empleador (A)	0,4204	0,4320	0,6043
TASA DE SUBDECLARACIÓN			
<i>límite inferior</i>			
log de la subdeclaración (lnk)	0,2990	0,3297	7,0980
factor de subdeclaración (k)	1,3486	1,3906	1209,5545
tasa de subdeclaración (k - 1)%	35%	39%	120855%
<i>límite superior</i>			
log de la subdeclaración (lnk)	0,3184	0,3338	7,0702
factor de subdeclaración (k)	1,3750	1,3963	1176,4216
tasa de subdeclaración (k - 1)%	37%	40%	117542%

Fuente: Elaboración propia en base a ENGH.

Si se realiza un promedio con los cinco casos más relevantes, se obtiene un rango de variación promedio que oscila entre el 30% y 31%, porcentaje superior al 18% promedio del capítulo 2, pero inferior a los resultados encontrados en Pissarides y Weber (1989), donde la tasa de subdeclaración promedio es del 55%.

²⁰ En la Tabla 3.3. ya se había visto que el este no era un buen instrumento del ingreso. Además, se observa que se invierte la relación entre los coeficientes del ingreso y de la variable binaria, elevando la tasa de subdeclaración promedio a cifras no interpretables.

3.2.3. Interpretación de los resultados obtenidos

Los resultados obtenidos en este capítulo deben analizarse teniendo en cuenta las mejoras introducidas en términos de especificación y estimación del modelo. No obstante, cabe aclarar que existe una ventaja potencial de los resultados del capítulo anterior respecto a los actuales, vinculada con el método de estimación utilizado. En el capítulo 2, se estimaron los parámetros de la ecuación (7) utilizando métodos robustos, para evitar la presencia de *outliers*, problema frecuente en las encuestas de hogares. Sin embargo, al utilizar aquí el método de mínimos cuadrados en dos etapas, no es posible continuar con este procedimiento, es decir que en este caso las estimaciones de los parámetros de la ecuación (18) no son robustas.

Las Tablas 3.6. y 3.7 sintetizan los resultados expuestos en las secciones 2.3. y 3.2. del presente trabajo y los obtenidos por Smith et al. (1986) y Pissarides y Weber (1989), respectivamente.

Tabla 3.6.
Comparación de los resultados del capítulo 2 y 3
variable binaria: empleador

	Modelo básico (capítulo 2)		Modelo extendido (capítulo 3)*	
	Alimentos	Alimentos, bebidas y tabaco	Alimentos	Alimentos, bebidas y tabaco
Ingreso	0,3627 (31,93)	0,4758 (44,49)	0,3016 (13,44)	0,4627 (22,35)
Empleador	0,0839 (2,21)	0,0717 (2,00)	0,1095 (2,42)	0,0640 (1,52)
varianza ingreso empleador (E)			0,4376	0,4376
varianza ingreso no empleador (A)			0,4303	0,4303
TASA DE SUBDECLARACIÓN	23%	15%	43% / 44%	14% / 15%

* Instrumentos del ingreso: edad y nivel de instrucción del jefe

Fuente: tablas 2.4., 2.5., 3.1., 3.2. y 3.4.

Tabla 3.7.

Comparación entre los resultados de Smith et al. (1986) y los de Pissarides y Weber (1989)

variable dependiente: gasto en alimentos - variable binaria: self-employed

	Smith, Pissarides y Weber (1986)		Pissarides y Weber (1989)	
	White-collar *	Blue-collar **	White-collar	Blue-collar
Ingreso	0,2300 (11,40)	0,3000 (17,70)	0,2700 (2,78)	0,2350 (2,90)
Self-employed (SE)	0,0960 (3,30)	0,0890 (3,50)	0,0920 (1,92)	0,1070 (2,55)
varianza ingreso SE			0,2500	0,1460
varianza ingreso no SE			0,0650	0,0600
TASA DE SUBDECLARACIÓN	42%	30%	28% / 54%	51% / 64%

* No coincide con el resultado de la tabla 12.6 de Smith et al (1986) que es igual a 36%, posiblemente porque esa sea un promedio de 2 regresiones (con y sin pendiente).

** Como no muestran resultados para la variable dependiente alimentos, aquí se exponen para “alimentos, bebidas y tabaco”.

Fuente: Smith, Pissarides y Weber (1986) y Pissarides y Weber (1989)

Como se mencionó en el capítulo anterior, la distinción entre *white* y *blue-collar* que hacen los autores para el caso británico dificulta la comparación con el caso bajo estudio. Si se compara el caso de white-collar con el resultado de la tabla 3.6. se observa dos diferencias: (i) Para el caso británico, la tasa de subdeclaración del modelo básico se ubica dentro del rango de la tasa de subdeclaración promedio obtenida con la versión extendida, mientras que en el caso argentino, la tasa promedio obtenida en el capítulo 3 casi duplica a la del capítulo 2, si se analizan las discrepancias en el consumo de alimentos (cuando se incorpora el consumo de bebidas y tabaco, las tasas de subdeclaración de uno y otro modelo son coincidentes). (ii) En el caso británico, el modelo básico reporta una elasticidad ingreso del gasto en alimentos menor que en la versión extendida, mientras que lo contrario se observa para el caso argentino.

La falta de diferencias sistemáticas entre uno y otro modelo impide la extracción de afirmaciones concluyentes. Sin embargo, cabe destacar que tanto en Pissarides y Weber (1989) como en el presente trabajo se verifica el supuesto de mayor volatilidad del ingreso del grupo que potencialmente subdeclara respecto del que no lo hace. Sin embargo, las diferencias encontradas en el caso británico son

mayores a las del caso argentino, repercutiendo sobre la extensión del rango de subdeclaración de ingresos.

Para finalizar, cabe reiterar que los resultados expuestos no pueden interpretarse como una aproximación absoluta de la subdeclaración existente en encuestas de gasto ni tampoco puede concluirse que ésta sea una aproximación del tamaño de la economía informal en el Gran Buenos Aires. Específicamente, este método permite calcular una tasa de subdeclaración promedio de un grupo de hogares respecto de otro. Asimismo, como se indica en el capítulo 1, la subdeclaración de ingresos por parte de las familias es una de las actividades asociadas con la economía informal y al mismo tiempo es posible que las familias reporten incorrectamente sus ingresos a las encuestas de hogares por razones ajenas a la evasión impositiva (aunque es poco probable que exista un patrón sistemático de las familias en este sentido).

Conclusiones

El objetivo del trabajo fue estimar la tasa de subdeclaración de ingresos para el caso argentino, en particular, para el Gran Buenos Aires en base a los datos de la Encuesta Nacional de Gasto de los Hogares realizada por el INDEC en 1996/97. Esta tarea es relevante por varias cuestiones: (i) Existen muy pocas cifras de economía oculta para el caso argentino, ninguna utilizando datos microeconómicos. (ii) La economía oculta es difícil de medir directamente. (iii) La estimación usa un método relativamente sofisticado, llamado “Discrepancias en el Consumo”.

El capítulo 1 sintetiza los principales puntos de debate académico sobre el tema, debido fundamentalmente a la multiplicidad de actividades asociadas con la economía informal y a la imposibilidad de medirla directamente, lo cual deriva en la existencia de diversos métodos indirectos cuyos resultados no pueden compararse fácilmente.

Este trabajo se basa en el método de discrepancias en el consumo, que consiste en traducir diferencias en el consumo de familias similares en una tasa de subdeclaración de un grupo de familias respecto a otro. Para ello, descansa en dos supuestos fundamentales: es posible clasificar a las familias en dos grupos, uno que potencialmente subdeclara sus ingresos y otro que no lo hace y todas las familias reportan correctamente el gasto en algún tipo de bien. Este método, su implementación a través de un modelo sencillo y sus principales limitaciones son tratados en detalle en el capítulo 2, donde también se describen los datos utilizados para tal fin. El modelo que mejor se ajusta indica que la tasa de subdeclaración de las familias empleadoras respecto a las que no lo son es del 23%, aunque el promedio de los modelos arroja una tasa que oscila en el entorno del 17%.

El capítulo 3 extiende el modelo permitiendo que la volatilidad del ingreso varíe entre individuos, al igual que la tasa de subdeclaración. Bajo el supuesto de que el ingreso de las familias que subdeclaran es más volátil que el de las que no subdeclaran, obtiene una tasa de subdeclaración *promedio*. Asimismo, este modelo se estima a través de mínimos cuadrados en dos etapas, lo cual significa una mejora respecto a la estimación por mínimos cuadrados ordinarios, porque elimina la endogeneidad del ingreso en la función consumo. En este caso, el modelo similar al

del capítulo 2 estima una tasa de subdeclaración mayor, del 43%, aunque el promedio de las distintas estimaciones varía en el entorno del 30%.

Este método, al igual que otras alternativas indirectas de medición de la economía informal, tiene limitaciones originadas en los supuestos restrictivos que permiten llevar a cabo la estimación y en la disponibilidad de los datos necesarios.

En primer lugar, si el grupo que se supone no subdeclara sus ingresos sí lo hace, la tasa de subdeclaración obtenida sólo puede interpretarse como la sobretasa de subdeclaración de un grupo respecto al otro. En segundo lugar, el método utilizado para clasificar a las familias entre uno y otro grupo es arbitrario y puede no reflejar correctamente el factor que lleva a una familia a subdeclarar sus ingresos. En tercer lugar, en todo momento se supone que los hogares reportan correctamente su ingreso o lo subdeclaran. Esto es así porque se asume que las razones que llevan a declarar incorrectamente el ingreso se vinculan únicamente con la evasión impositiva. Sin embargo, las respuestas dadas a encuestas de gastos son confidenciales y alguien puede “sobredeclarar” sus ingresos.

En cuarto lugar, no todas las personas encuestadas responden la encuesta. Si se considera que los motivos que llevan a no responder están asociados con la subdeclaración de ingresos, podría decirse que los resultados obtenidos representan un piso de subdeclaración. En quinto lugar, es posible que las familias también reporten con error sus gastos. Para ser consistentes con el supuesto de subdeclaración de ingresos, podría asumirse que algunas familias tiendan también a subdeclarar sus consumos. Por este motivo se seleccionaron bienes con baja elasticidad ingreso (como los alimentos), aunque esto no excluye la posibilidad de que estén subdeclarados. En ese caso, la tasa de subdeclaración de ingresos obtenida también estaría indicando un límite inferior por esta fuente. En sexto lugar, si bien se probó un amplio conjunto de variables explicativas para captar las diferencias entre familias, es posible que haya diferencias no explicadas que afecten el consumo de las mismas y que esto esté siendo captado por la variable binaria. En ese caso, la tasa estimada estaría sobreestimada ya que incluiría también el efecto de otras cuestiones que diferencian el consumo de dos familias aparentemente similares. Por último, debe tenerse en cuenta que la subdeclaración de ingresos es una de las actividades asociadas con la economía informal. Por ello, es muy difícil comparar los resultados

de este trabajo con los obtenidos con otros métodos de estimación, también basados en supuestos que limitan su generalización.

A pesar de las limitaciones arriba enunciadas, este trabajo puede considerarse un primer paso en un método no explorado para el caso argentino. Existen varias direcciones de investigación futura en este campo. Por un lado, pueden aplicarse métodos semiparamétricos o no paramétricos que permitan abordar preguntas empíricas más detalladas. Por otro lado, una mejor apertura de los datos que permita distinguir entre actividades white o blue-collar haría posible una mejor identificación del factor causante de la subdeclaración. Otra posible dirección de investigación sería la incorporación de las otras regiones relevadas en la ENGH de 1996/97 o de la encuesta de gasto realizada en el Gran Buenos Aires en 1985, haciendo posible esta última una comparación temporal de la tasa de subdeclaración. Tal vez el camino más arduo que podría seguirse es la compatibilización de este método con otros alternativos, haciendo posible la comparación de resultados obtenidos en distintos momentos o en distintas zonas de la Argentina.

Anexo

D) Descripción de variables utilizadas en el trabajo

La Encuesta Nacional de Gasto de los Hogares implementada por el INDEC entre febrero de 1996 y marzo de 1997 fue la primera encuesta de este tipo con alcance nacional urbano. Basada en una muestra de 114 ciudades, representa a 28 millones de habitantes, el 96% de la población urbana. La anterior encuesta de gastos e ingresos de los hogares fue realizada en 1985-86 en el área geográfica de Capital Federal y 19 partidos del conurbano bonaerense, que tenía una población de 10 millones de habitantes.

Una de las razones principales que determinaron la realización de esta encuesta fue la actualización de los coeficientes de ponderación del índice de precios al consumidor. Sin embargo, la encuesta capta información sobre diversos aspectos de las condiciones de vida de los hogares. De este modo, los datos relevados permiten relacionar los ingresos y gastos con variables sociodemográficas.²¹

A continuación se describen las variables construidas con los microdatos de la ENGH (1996/97) para el Área Metropolitana del Gran Buenos Aires y que fueron utilizadas en este trabajo.

Variable	Descripción
clave	clave del hogar
expan	factor de expansión
cantmiem	cantidad de miembros en el hogar
aef	cantidad de adultos equivalentes en el hogar
aefee	cantidad de adultos equivalentes corregidos por economías de escala en el hogar
ingreso	ingreso total familiar (sumatoria de ingreso por fuente de todos los percetores)
ipcf	ingreso per cápita familiar = ingreso / cantmiem
ipae	ingreso por adulto equivalente = ingreso / aef
ipaece	ingreso por adulto equivalente y economías de escala = ingreso / aefee
lning	ln del ingreso
lnipcf	ln del ipcf
lnipae	ln del ipae
lnipaece	ln del ipaece
lngtoal	ln del gtoalim
lngcl	ln del gcl
lngcltab	ln gto en alimentos, bebidas y tabaco
lngtoal	ln gto en alimentos

²¹ Para una descripción detallada de la encuesta, su metodología y objetivos ver INDEC (1998).

Variable	Descripción
lnrent	ln renta
lnrent1	ln rentrep
lnrent2	ln rentrep2
lncomb1	ln comb1
lncomb2	ln comb2
lncombto	ln combtot
lnnodur	ln nodur
lnserv	ln serv
lnnodser	ln nodurserv
gtoalim	gasto del hogar en alimentos
gc1	gasto del hogar en alimentos y bebidas
gc2	gasto del hogar en indumentaria y calzado
gc3	gasto del hogar en vivienda
gc4	gasto del hogar en equipamiento y esparcimiento
gc5	gasto del hogar en atención médica y gastos para la salud
gc6	gasto del hogar en transporte y comunicaciones
gc7	gasto del hogar en esparcimiento
gc8	gasto del hogar en educación
gc9	gasto del hogar en bienes y servicios varios
gastot	gasto total del hogar
gc1tab	gc1 + tabaco
tabaco	gto en cigarrillos y tabaco
renta	gasto en alquiler, valor locativo, hotel o pensión
repaviv	gasto en reparacion y mantenimiento de vivienda
rentrep	renta + repaviv
rentrep2	renta + repaviv + agelect
agelect	gasto en agua potable y electricidad
comb1	gasto en combustibles, lavado, engrase, lubricantes, etc.
comb2	gasto en gas, kerosene, carbon, etc.
combtot	com1 + comb2
nodur	gasto en otros bienes no durables
serv	gasto en servicios
nodurser	nodur + serv
jasal	variable binaria =1 para jefe asalariado como ocupación principal
jctaprop	variable binaria =1 para jefe cuentapropista como ocupación principal
jempl	variable binaria =1 para jefe empleador como ocupación principal
jjubil	variable binaria =1 para jefe jubilado (puede además figurar en las tres anteriores)
jrent	variable binaria =1 para jefe rentista (puede además figurar en las tres primeras)
jnoasal	variable binaria =1 si jasal=0
ingAj	ingreso del jefe por fuente asalariado
ingCj	ingreso del jefe por fuente cuentapropia
ingEj	ingreso del jefe por fuente empleador
ingJj	ingreso del jefe por fuente jubilado
ingRj	ingreso del jefe por fuente rentista
ingresoj	ingreso total del jefe
ingjefeA	proporción de ingreso del jefe por fuente asalariado
ingjefeC	proporción de ingreso del jefe por fuente cuentapropia
ingjefeE	proporción de ingreso del jefe por fuente empleador
ingjefeJ	proporción de ingreso del jefe por fuente jubilado

Variable	Descripción
ingjefeR	proporción de ingreso del jefe por fuente rentista
jasal2	variable binaria =1 para ingjefeA >0.25
jctap2	variable binaria =1 para ingjefeC >0.25
jempl2	variable binaria =1 para ingjefeE >0.25
jjubil2	variable binaria =1 para ingjefeJ >0.25
jrent2	variable binaria =1 para ingjefeR >0.25
ingA	ingreso familiar por fuente asalariado
ingC	ingreso familiar por fuente cuentapropia
ingE	ingreso familiar por fuente empleador
ingJ	ingreso familiar por fuente jubilado
ingR	ingreso familiar por fuente rentista
ingEoC	ingreso familiar por fuente empleador y cuentapropia
ingfliaA	proporción de ingreso familiar por fuente asalariado
ingfliaC	proporción de ingreso familiar por fuente cuentapropia
ingfliaE	proporción de ingreso familiar por fuente empleador
ingfliaJ	proporción de ingreso familiar por fuente jubilado
ingfliaR	proporción de ingreso familiar por fuente rentista
ingfliaEoC	proporción de ingreso familiar por fuente empleador y cuentapropia
fliaA	variable binaria =1 para ingfliaA >0.25
fliaC	variable binaria =1 para ingfliaC >0.25
fliaE	variable binaria =1 para ingfliaE >0.25
fliaJ	variable binaria =1 para ingfliaJ >0.25
fliaR	variable binaria =1 para ingfliaR >0.25
fliaEoC	variable binaria =1 para ingfliaEoC >0.25
jnivins	nivel de instrucción del jefe
age	edad del jefe - 40
agesq	age al cuadrado
room	cantidad de habitaciones
Room2	cantidad de habitaciones de uso exclusivo
Car	nº de autos del hogar (truncado en 2)
Oom	hogares con pago pendiente por la vivienda
Ch	hogares con calefacción central
Nch	nº de hijos en el hogar (no importa la edad sino la clasificación de "hijos" del jefe de familia)
nchsq	nch al cuadrado
Kids1	nº de chicos entre 0 y 4 años (no importa si son hijos o nietos del jefe de familia o tienen otro parentesco)
kids2	nº de chicos entre 5 y 16 años (no importa si son hijos o nietos del jefe de familia o tienen otro parentesco)
dosadult	variable binaria = 1 si en el hogar hay dos adultos y son cónyuges
YfliaE	lning * fliaE

II) Resultados de regresión

II.1) Regresiones basadas en la clasificación de los hogares de acuerdo a la fuente de ingresos que representa más del 50% del ingreso familiar

Estimaciones de varias versiones del modelo - variable dependiente: gasto en alimentos

Estimación robusta - gastos e ingreso en logaritmos

	(1)	(2)	(3)	(4)	(5)
Ingreso	0,3703 (28,06)	0,3681 (33,43)	0,3643 (32,14)	0,3659 (31,56)	0,3629 (32,10)
Cantidad de miembros	0,0649 (7,02)	0,0651 (7,07)	0,0654 (7,10)	0,0646 (7,01)	0,0647 (7,03)
Edad del jefe - 40 (age)	0,0055 (4,72)	0,0055 (4,73)	0,0055 (4,70)	0,0056 (4,75)	0,0056 (4,80)
Age al cuadrado (agesq)	-0,0001 (-2,60)	-0,0001 (-2,62)	-0,0001 (-2,68)	-0,0001 (-2,88)	-0,0001 (-2,95)
nº de chicos e/ 0 y 4 años	0,0007 (0,04)	0,0006 (0,03)	-0,0004 (-0,02)	-0,0004 (-0,03)	-0,0004 (-0,02)
nº de chicos e/ 5 y 16 años	0,0302 (2,51)	0,0300 (2,49)	0,0296 (2,46)	0,0309 (2,57)	0,0305 (2,54)
Empleadores o cuentapropistas (no asalariados) EoC	0,1028 (0,64)	0,0546 (2,66)			
EoC x ingreso	-0,0069 (-0,30)				
Cuentapropistas			0,0387 (1,72)		
Empleadores			0,1024 (2,53)	0,5258 (1,34)	0,0941 (2,34)
Empleadores x ingreso				-0,0566 (-1,11)	
Constante	2,5672 (27,19)	2,5818 (31,85)	2,6096 (31,41)	2,6107 (31,04)	2,6306 (32,01)

Fuente: Elaboración propia en base a ENGH.

Estimaciones de varias versiones del modelo - variable binaria: empleador

Estimación robusta - gastos e ingreso en logaritmos

	Alimentos y Bebidas	Alimentos, Bebidas y Tabaco	Commitments	Servicios y No Durables
Ingreso	0,4785 (44,62)	0,4762 (44,72)	0,6243 (26,66)	0,8684 (55,02)
Cantidad de miembros	0,0621 (7,11)	0,0643 (7,40)	-0,1149 (-6,18)	0,0209 (1,65)
Edad del jefe - 40 (age)	0,0030 (2,69)	0,0027 (2,46)	-0,0117 (-4,84)	-0,0003 (-0,20)
Age al cuadrado (agesq)	-0,0001 (-3,39)	-0,0002 (-3,93)	0,0002 (1,96)	-0,0002 (-2,96)
nº de chicos e/ 0 y 4 años	-0,0228 (-1,36)	-0,0282 (-1,69)	0,0761 (2,09)	-0,0102 (-0,42)
nº de chicos e/ 5 y 16 años	0,0018 (0,15)	-0,0040 (-0,35)	0,0360 (1,47)	-0,0422 (-2,55)
Empleadores	0,0866 (2,27)	0,0770 (2,03)	-0,0591 (-0,73)	0,1180 (2,13)
Constante	2,2101 (28,36)	2,2727 (29,37)	-0,0240 (-0,14)	-0,9215 (-8,03)

Fuente: Elaboración propia en base a ENGH.

II.2) Regresiones a partir de mínimos cuadrados ordinarios

Estimaciones de varias versiones del modelo - variable dependiente: gasto en alimentos

Estimación por MCO - gastos e ingreso en logaritmos

	(1)	(2)	(3)	(4)	(5)
Ingreso	0,3628 (23,49)	0,3595 (29,43)	0,3568 (28,31)	0,3612 (27,89)	0,3561 (28,29)
Cantidad de miembros	0,0665 (6,51)	0,0667 (6,54)	0,0669 (6,56)	0,0666 (6,53)	0,0669 (6,56)
Edad del jefe - 40 (age)	0,0054 (4,16)	0,0054 (4,16)	0,0054 (4,14)	0,0054 (4,18)	0,0055 (4,20)
Age al cuadrado (agesq)	-0,0001 (-1,95)	-0,0001 (-1,98)	-0,0001 (-2,00)	-0,0001 (-2,11)	-0,0001 (-2,18)
nº de chicos e/ 0 y 4 años	-0,0222 (-1,13)	-0,0223 (-1,14)	-0,0232 (-1,18)	-0,0235 (-1,20)	-0,0238 (-1,21)
nº de chicos e/ 5 y 16 años	0,0345 (2,59)	0,0344 (2,58)	0,0342 (2,57)	0,0351 (2,63)	0,0344 (2,59)
Empleadores o cuentapropistas (no asalariados) EoC	0,0923 (0,54)	0,0342 (1,63)			
EoC x ingreso	-0,0084 (-0,34)				
Cuentapropistas			0,0248 (1,12)		
Empleadores			0,0644 (1,52)	0,7338 (1,78)	0,0582 (1,39)
Empleadores x ingreso				-0,0880 (-1,65)	
Constante	2,5701 (23,35)	2,5920 (28,87)	2,6113 (28,34)	2,5901 (27,63)	2,6244 (28,71)

Fuente: Elaboración propia en base a ENGH.

Estimaciones de varias versiones del modelo - variable binaria: empleador

Estimación por MCO - gastos e ingreso en logaritmos

	Alimentos y Bebidas	Alimentos, Bebidas y Tabaco	Commitments	Servicios y No Durables
Ingreso	0,4802 (40,62)	0,4792 (41,04)	0,6578 (28,41)	0,8960 (51,57)
Cantidad de miembros	0,0605 (6,32)	0,0633 (6,68)	-0,1193 (-6,52)	0,0207 (1,49)
Edad del jefe - 40 (age)	0,0034 (2,81)	0,0030 (2,49)	-0,0119 (-5,01)	-0,0021 (-1,18)
Age al cuadrado (agesq)	-0,0001 (-2,94)	-0,0001 (-3,43)	0,0002 (1,95)	-0,0001 (-1,65)
nº de chicos e/ 0 y 4 años	-0,0358 (-1,94)	-0,0444 (-2,44)	0,0791 (2,21)	-0,0337 (-1,26)
nº de chicos e/ 5 y 16 años	0,0102 (0,82)	0,0038 (0,31)	0,0355 (1,48)	-0,0459 (-2,53)
Empleadores	0,0452 (1,14)	0,0370 (0,95)	-0,1238 (-1,65)	0,1350 (2,36)
Constante	2,1599 (25,16)	2,2147 (26,12)	-0,1299 (-0,77)	-1,1540 (-9,14)

Fuente: Elaboración propia en base a ENGH.

II.3) Regresiones con distintas nociones de tamaño de familia, con diversas características sociodemográficas y de las viviendas habitadas

Estimaciones de varias versiones del modelo - variable dependiente: gasto en alimentos

Tamaño del hogar: definido a partir del número de hijos

Estimación robusta - gastos e ingreso en logaritmos

	(1)	(2)	(3)	(4)	(5)
Ingreso	0,3712 (26,79)	0,3649 (33,28)	0,3603 (31,91)	0,3637 (31,35)	0,3602 (31,95)
n° de hijos (nch) - sin importar edad	(0,1362) (8,66)	(0,1368) (8,70)	(0,1367) (8,70)	(0,1366) (8,70)	(0,1367) (8,71)
nch al cuadrado	-0,0114 (-5,14)	-0,0115 (-5,17)	-0,0115 (-5,17)	-0,0115 (-5,18)	-0,0115 (-5,18)
Edad del jefe - 40 (age)	0,0067 (5,96)	0,0067 (5,98)	0,0067 (5,97)	0,0067 (5,94)	0,0068 (6,00)
Age al cuadrado (agesq)	-0,0001 (-1,99)	-0,0001 (-2,05)	-0,0001 (-2,11)	-0,0001 (-2,11)	-0,0001 (-2,17)
n° de chicos e/ 0 y 4 años	0,0166 (1,02)	0,0167 (1,02)	0,0159 (0,98)	0,0159 (0,98)	0,0158 (0,97)
n° de chicos e/ 5 y 16 años	0,0355 (2,89)	0,0354 (2,88)	0,0354 (2,88)	0,0358 (2,92)	0,0354 (2,89)
Empleadores o cuentapropistas (no asalariados) EoC	0,1292 (0,85)	0,0197 (1,05)			
EoC x ingreso	-0,0158 (-0,72)				
Cuentapropistas			0,0059 (0,30)		
Empleadores			0,0736 (1,94)	0,5173 (1,40)	0,0720 (1,91)
Empleadores x ingreso				-0,0583 (-1,22)	
Constante	2,6214 (26,68)	2,6643 (33,53)	2,6980 (33,03)	2,6767 (32,17)	2,7008 (33,34)

Fuente: Elaboración propia en base a ENGH.

Estimaciones de varias versiones del modelo - variable binaria: empleador

Tamaño del hogar: definido a partir del número de hijos

Estimación robusta - gastos e ingreso en logaritmos

	Alimentos y Bebidas	Alimentos, Bebidas y Tabaco	Commitments	Servicios y No Durables
Ingreso	0,4782 (44,50)	0,4769 (44,65)	0,6247 (26,66)	0,8649 (54,60)
n° de hijos (nch) - sin importar edad	0,1026 (6,85)	0,1000 (6,71)	-0,1966 (-6,08)	0,0663 (3,05)
nch al cuadrado	-(0,0060) (-2,85)	-(0,0058) (-2,77)	(0,0168) (3,61)	-(0,0054) (-1,77)
Edad del jefe - 40 (age)	0,0038 (3,54)	0,0037 (3,46)	-0,0151 (-6,52)	-0,0005 (-0,33)
Age al cuadrado (agesq)	-0,0001 (-2,92)	-0,0001 (-3,60)	0,0002 (1,90)	-0,0001 (-2,29)
n° de chicos e/ 0 y 4 años	-0,0091 (-0,58)	-0,0112 (-0,72)	0,0072 (0,21)	-0,0146 (-0,64)
n° de chicos e/ 5 y 16 años	0,0039 (0,34)	0,0013 (0,11)	0,0043 (0,17)	-0,0514 (-3,01)
Empleadores	0,0698 (1,94)	0,0596 (1,67)	-0,0878 (-1,15)	0,1077 (2,07)
Constante	2,2943 (29,71)	2,3583 (30,73)	-0,1464 (-0,87)	-0,9016 (-7,93)

Fuente: Elaboración propia en base a ENGH.

Estimaciones de varias versiones del modelo - variable dependiente: gasto en alimentos y bebidas

Estimación robusta - gastos e ingreso en logaritmos

	(1)	(2)
Ingreso	0,4523 (26,80)	0,4447 (27,74)
Cantidad de miembros	0,0700 (8,00)	0,0704 (8,21)
Edad del jefe - 40 (age)	0,0012 (0,94)	0,0015 (1,20)
Age al cuadrado (agesq)	-0,0001 (-0,97)	-0,0001 (-1,22)
n° de chicos e/ 0 y 4 años	-0,0313 (-1,77)	-0,0333 (-1,91)
n° de chicos e/ 5 y 16 años	-0,0109 (-0,83)	-0,0101 (-0,78)
Empleadores	0,0518 (1,41)	0,0618 (1,70)
Agua corriente en la vivienda	0,0050 (0,20)	
Agua caliente en la cocina	0,0936 (2,49)	0,0709 (2,41)
Ascensor en el edificio	-0,0535 (-0,83)	
Casa	0,0187 (0,52)	
Cochera	0,0109 (0,47)	
Cocina	-0,0564 (-0,74)	
Combustible en la cocina	-0,0228 (-0,60)	
Distancia al transporte público	0,0279 (0,67)	0,0255 (0,62)
Depto en edificio de 2 pisos	0,0004 (0,01)	
Material de la Vivienda	0,0002 (0,01)	0,0029 (0,18)
Cantidad de baños	0,0310 (1,65)	0,0317 (1,75)
Cantidad de habitaciones	0,0092 (0,83)	0,0087 (0,81)
Seguridad en el edificio	-0,0562 (-0,96)	
Servicio de portero eléctrico	0,0203 (0,32)	
Sistema de calefacción	0,0115 (0,42)	
Teléfono	-0,0319 (-1,20)	-0,0403 (-1,55)
Constante	2,3137 (16,77)	2,3337 (21,37)

Fuente: Elaboración propia en base a ENGH.

Estimaciones de varias versiones del modelo - variable dependiente: gasto en alimentos y bebidas
Estimación robusta - gastos e ingreso en logaritmos

	(1)	(2)
Ingreso	0,4523 (26,80)	0,4447 (27,74)
Cantidad de miembros	0,0700 (8,00)	0,0704 (8,21)
Edad del jefe - 40 (age)	0,0012 (0,94)	0,0015 (1,20)
Age al cuadrado (agesq)	-0,0001 (-0,97)	-0,0001 (-1,22)
n° de chicos e/ 0 y 4 años	-0,0313 (-1,77)	-0,0333 (-1,91)
n° de chicos e/ 5 y 16 años	-0,0109 (-0,83)	-0,0101 (-0,78)
Empleadores	0,0518 (1,41)	0,0618 (1,70)
Agua corriente en la vivienda	0,0050 (0,20)	
Agua caliente en la cocina	0,0936 (2,49)	0,0709 (2,41)
Ascensor en el edificio	-0,0535 (-0,83)	
Casa	0,0187 (0,52)	
Cochera	0,0109 (0,47)	
Cocina	-0,0564 (-0,74)	
Combustible en la cocina	-0,0228 (-0,60)	
Distancia al transporte público	0,0279 (0,67)	0,0255 (0,62)
Depto en edificio de 2 pisos	0,0004 (0,01)	
Material de la Vivienda	0,0002 (0,01)	0,0029 (0,18)
Cantidad de baños	0,0310 (1,65)	0,0317 (1,75)
Cantidad de habitaciones	0,0092 (0,83)	0,0087 (0,81)
Seguridad en el edificio	-0,0562 (-0,96)	
Servicio de portero eléctrico	0,0203 (0,32)	
Sistema de calefacción	0,0115 (0,42)	
Teléfono	-0,0319 (-1,20)	-0,0403 (-1,55)
Constante	2,3137 (16,77)	2,3337 (21,37)

Fuente: Elaboración propia en base a ENGH.

II.4) Regresiones del ingreso utilizadas para el cálculo de la tasa de subdeclaración en el modelo extendido del capítulo 3

Estimaciones de las ecuaciones de ingreso

Variable dependiente: logaritmo del ingreso familiar

	Para tablas 3.1. y 3.2		Para tabla 3.3					
	Empleador (fliaE=1)	No empleador (fliaE=0)	Empleador (fliaE=1)	No empleador (fliaE=0)	Empleador (fliaE=1)	No empleador (fliaE=0)	Empleador (fliaE=1)	No empleador (fliaE=0)
Edad del jefe - 40 (age)	0,0031 (0,58)	-0,0035 (-3,75)	0,0037 (0,44)	0,0072 (4,71)			-0,0040 (-0,67)	-0,0106 (-9,79)
Age al cuadrado (agesq)			(0,00) (-0,09)	(0,00) (-8,71)				
nivel de instrucción del jefe (jnivins)	0,1635 (7,95)	0,2161 (35,93)	0,1634 (7,92)	0,2131 (35,79)	0,1615 (7,98)	0,2209 (37,50)		
Cantidad de miembros	(0,05) (1,05)	(0,20) (17,59)	(0,05) (0,99)	(0,16) (13,51)	(0,06) (1,28)	(0,19) (17,18)	(0,04) (0,77)	(0,17) (12,58)
n° de chicos e/ 0 y 4 años	-0,1799 (-1,85)	-0,3023 (-13,90)	-0,1773 (-1,75)	-0,2277 (-9,83)	-0,2092 (-2,52)	-0,2660 (-13,63)	-0,2133 (-1,92)	-0,3436 (-13,35)
n° de chicos e/ 5 y 16 años	-0,0258 (-0,41)	-0,2051 (-12,97)	-0,0258 (-0,41)	-0,1918 (-12,22)	-0,0418 (-0,74)	-0,1866 (-12,40)	-0,0606 (-0,84)	-0,2177 (-11,62)
Constante	6,5134 (29,77)	5,2255 (101,70)	6,5189 (28,72)	5,3980 (99,03)	6,5409 (30,66)	5,1685 (105,11)	7,7051 (42,10)	6,5501 (154,41)
Varianza del ingreso	0,4376	0,4303	0,4398	0,4204	0,4361	0,4320	0,5765	0,6043

Fuente: Elaboración propia en base a ENGH.

Referencias

- Ahumada, H., Canavese, A., Canavese, P. y González Alvaredo, F., 2000, El tamaño de la economía oculta. Revisión del método de estimación basado en la demanda de circulante con una ilustración para la Argentina. FIEL.
- Bhattacharyya, D., 1990, An econometric model of estimating the “hidden” economy, United Kingdom (1960-1984): estimates and tests, *Economic Journal*, 100, pp. 703-717.
- Bhattacharyya, D., 1999, On the economic rationale of estimating the hidden economy, *The Economic Journal*, 109, F348-F359.
- Blundell y Duncan, 1989, Non-parametric regression.
- Cagan, P., 1958, The demand for currency relative to the total money supply, *Journal of Political Economy*, 66, 303-329.
- Dixon, H., 1999, Controversy: On the use of the ‘hidden economy’ estimates, *The Economic Journal*, 109, F335-F337.
- Feige, E., 1979, How big is the irregular economy, *Challenge*, 22, 5-13.
- Frey, B. y Weck-Hannemann, 1984, The hidden economy as an unobserved variable, *European Economic Review*, 26, 33-53.
- Giles, D., 1997, The hidden economy and the tax gap in New Zealand: a latent variable analysis, Discussion Paper 97-8, Department of Economics, University of Victoria.
- Giles, D., 1999, Measuring the hidden economy: implicatinos for econometric modelling, *The Economic Journal*, 109, F370-F380.
- Instituto Nacional de Estadística y Censos (INDEC), 1998, *Encuesta Nacional de los Gastos de los Hogares 1996/97. Región Metropolitana del Gran Buenos Aires, Resultados Provisorios. Vol. 1.*
- Jarque, C. y Bera, A., 1987, A test for normality of observations and regression residuals, *International Statistical Review*, 55, pp. 163-172.
- Lyssioutou, P., Pashardes, P. y Stengos, T., 1999, Consumer demand based estimates of the black economy: parametric and nonparametric approaches, mimeo, University of Cyprus.
- Manski, C., 1994, *Identification Problems in the Social Sciences*, Harvard University Press, Cambridge.
- Pissarides, C. y Weber, G., 1989, An expenditure-based estimate of Britain’s black economy, *Journal of Public Economics*, 39, pp. 17-32.

- Schneider, F., 1997, Empirical results for the size of the shadow economy of Western European countries over time, Working Paper 9710, Linz University.
- Schneider, F. y Enste, D., 2000, Shadow economies: size, causes, and consequences, *Journal of Economic Literature*, en prensa.
- Smith, S., 1986, *Britain's Shadow Economy*, Clarendon Press, Oxford.
- Sosa Escudero, W. y Gasparini, L., 2000, A note on the statistical significance of changes in inequality, mimeo, Universidad Nacional de La Plata.
- Tanzi, V., 1983, The underground economy in the Unites States: annual estimates 1930-1980, *IMF Staff Papers*, 30, 283-305
- Tanzi, V., 1999, Uses and abuses of estimates of the underground economy, *The Economic Journal*, 109, pp. 338-347.
- Thomas, J., 1999, Quantifying the black economy: 'Measurement without theory' yet again?, *The Economic Journal*, 109, pp. F381-F389.
- Wilcox, R., *Introduction to Robust Estimation and Hypothesis Testing*, Academic Press, San Diego.