



Blaconá, María Teresa*

Bussi, Javier**

**Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística. Consejo de Investigaciones, Universidad Nacional de Rosario. E-mail: mblacona@fcecon.unr.edu.ar*

*** Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario.*

ESTUDIO MULTIVARIADO DE LAS SERIES DE TIEMPO TASA DE DESOCUPACIÓN DE GRAN BUENOS AIRES Y GRAN ROSARIO, 1974-2005¹

1 INTRODUCCIÓN

El estudio de la tasa de desocupación es un tema de interés a nivel mundial. A través de su análisis, es posible proveer información para la implementación de políticas destinadas a promover el bienestar general a través de la oferta de trabajo. La tasa de desocupación en Argentina no escapa a esta tendencia mundial, siendo objeto de estudio por las razones antes mencionadas, pero también debido a la preocupante presencia de un comportamiento creciente en los últimos veinte años, que alcanza valores máximos históricos con tasas de alrededor del 20%.

Las tasas de desocupación corresponden a la información provista por la EPH-INDEC, para los distintos aglomerados urbanos del país. Las tasas se elaboraban con información que se relevaba y publicaba en forma bianual. A partir del año 2003 se registra la información en forma continua y se publica en forma trimestral. Por lo tanto, las series de tasa de desocupación de los distintos aglomerados cuentan con información bianual desde 1974 a 2002 e información trimestral a partir del año 2003.

Por lo tanto, es de interés estudiar las series de las tasas de desocupación en la nueva periodicidad, con el fin de determinar su comportamiento, reflejado en la tendencia y la estacionalidad, y también para poder realizar pronósticos en forma trimestral.

Entre los distintos aglomerados urbanos del país, Gran Rosario (GR) y Gran Buenos Aires (GBA) son áreas metropolitanas que han manifestado un marcado incremento de la desocupación, con comportamientos similares a través del tiempo con respecto a la tendencia y cambios de nivel.

En trabajos anteriores se analizó ambas series por separado a través de los modelos de espacio de estados (MEE) mediante un enfoque univariado (Blaconá y otros, 2005; Blaconá, M.T.; Bussi, J., 2005). Estos trabajos permitieron estimar la tendencia, la estacionalidad y la realización de pronósticos combinando ambas periodicidades presentes en las series.

Es de interés entonces ahora, extender el análisis mediante la utilización de un enfoque multivariado, a través del uso de los modelos de espacios de estados, lo cual permite además combinar la información provista por ambas series simultáneamente.

¹ Este trabajo se realizó en el marco del proyecto: Estimación y pronósticos de series de tiempo multivariadas.



En la sección 2 se presenta como se desagregan las series de tiempo tasa de desocupación de GBA y GR, para poder contar con información trimestral de ambas en el período completo (1974 – 2005), en la sección 3 se realiza una breve descripción de los MEE multivariados (MEM), en la sección 4 se presentan los resultados para las series tasa de desocupación de GBA y GR y por último en la sección 5 se realizan las consideraciones finales.

2 ESTIMACIÓN DE LAS SERIES DE TIEMPO TRIMESTRALES TASA DE DESOCUPACIÓN DE GBA Y GR EN EL PERÍODO 1974-2005

Como se menciona en la sección anterior las series tasa de desocupación de GBA y GR se publica en forma bianual desde 1974 al 2002, mientras que a partir del 2003 se lo hace en forma trimestral. Por esta razón, para realizar el análisis bivariado de las series trimestrales en el período completo 1974-2005, se decide desagregar las series usando los modelos MEE, según las estimaciones encontradas en los análisis de las series en forma univariada (Blaconá y otros, 2005; Blaconá, M.T.; Bussi, J., 2005).

En los trabajos antes mencionados, por convención se tomaron las observaciones de la primera onda del período 1974-2002 como equivalentes a las del segundo trimestre y las de la segunda onda como las del cuarto trimestre.

Para realizar la estimaciones de los datos faltantes, es decir, primer y tercer trimestre en el período 1974-2002, se utiliza el MEE de la siguiente manera.

El modelo básico de espacio de estados, que también se conoce como modelo lineal Gaussiano de espacio de estados, se puede expresar como

$$\begin{aligned} \mathbf{y}_t &= \mathbf{Z}_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, & \boldsymbol{\varepsilon}_t &\sim \text{NID}(\mathbf{0}, \mathbf{H}_t), \\ \boldsymbol{\alpha}_t &= \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \mathbf{R}_t \boldsymbol{\eta}_t, & \boldsymbol{\eta}_t &\sim \text{NID}(\mathbf{0}, \mathbf{Q}_t), \end{aligned} \quad (2.1)$$

donde \mathbf{y}_t es un vector de orden $p \times 1$ de observaciones, $\boldsymbol{\alpha}_t$ es un vector de orden $m \times 1$ no observable denominado vector de estado, las matrices $\mathbf{Z}_t (p \times m)$, $\mathbf{T}_t (m \times m)$ y $\mathbf{R}_t (m \times g)$ son conocidas y $\boldsymbol{\eta}_t$ es un vector aleatorio de orden $g \times 1$. \mathbf{R}_t se denomina matriz de selección y está formada de ceros y unos de acuerdo a si los $\boldsymbol{\eta}$ son determinísticos o aleatorios.

A la primera ecuación de (2.1) por lo general se la llama ecuación de medida y a la segunda, ecuación de transición.

Se considera el caso correspondiente a una serie con dos periodicidades, siendo la periodicidad inicial la que corresponde al período más largo entre dos observaciones registradas. En consecuencia, la periodicidad final corresponde al registro de observaciones en forma más frecuente.

El modelo general resulta:



$$y_t^v = \mathbf{z}'\alpha_t + \varepsilon_t, \quad \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2 \quad (2.2.1)$$

$$\alpha_t = \mathbf{T}\alpha_{t-1} + \mathbf{R}\eta_t, \quad \text{Var}(\eta_t) = \mathbf{Q}$$

En el caso en que se deba analizar una serie con distintas periodicidades, se debe describir más detalladamente el MEE que es necesario formular.

Se incluye en la notación el supraíndice v con el fin de destacar que la serie no es observada en todos los períodos de tiempo $t = 1, 2, 3, \dots, T$. Las observaciones que se indican de la forma $y_{t,\tau}$, $\tau = 1, \dots, T_1$, están registradas cada δ períodos

$$y_t = y_{\delta\tau}^v \quad \tau = 1, \dots, T_1.$$

En estos casos para los $t \neq \delta\tau$, y_t^v no es observada. Las restantes observaciones de T_1+1 a T se presentan en la periodicidad más frecuente sin datos faltantes.

El MEE cuando existen observaciones faltantes debido al cambio de periodicidad, puede ser formulado de dos maneras. Una forma es plantear el modelo de acuerdo a los períodos más cortos (denominada de ahora en más FEE($t,1$)) con $t = 1, \dots, T$. Otra forma de plantear el modelo es de acuerdo a los períodos más extensos, denominada FEE(τ,δ).

Ambas formas de plantear el MEE permiten que sean estimados los hiperparámetros, pero solo la FEE($t,1$) admite la interpolación, es decir la estimación de valores en los momentos en los que no fueron registrados debido a la periodicidad en intervalos más largos. Otra ventaja de plantear el modelo de esta forma es que es posible hacer predicciones para ambas periodicidades como así también estimaciones de los componentes no observables en los períodos más cortos.

Esta formulación del MEE, la cual permite utilizar las dos partes de la serie de tiempo, resulta en una forma del filtro de Kalman que puede ser utilizada para el pronóstico de observaciones futuras y faltantes. El filtro para la forma del modelo FEE($t,1$) se ejecuta de forma directa, el mismo saltea la actualización de las ecuaciones en los períodos intermedios donde no se cuenta con información, pero los considera en la estimación de la variancias.

Por ejemplo, sin pérdida de generalidad, para un modelo sólo con nivel resulta:

$$y_t = \mu_t + \varepsilon_t \quad (2.2)$$

$$\mu_t = \mu_{t-1} + \eta_t$$



con $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2$ y $\text{Var}(\eta_t) = \sigma_\eta^2$.

Suponiendo que no se cuenta con la observación y_k , la serie $y = y_{t_\tau}^v$, $\tau = 1, \dots, T$

donde $t_\tau = \begin{cases} \tau & \tau = 1, \dots, k-1 \\ \tau+1 & \tau = k, \dots, T \end{cases}$, se tiene el modelo para el momento $k+1$

$$\begin{aligned} y_{k+1} &= \mu_{k+1} + \varepsilon_{k+1} \\ \mu_{k+1} &= \mu_k + \eta_{k+1} \end{aligned} \quad (2.3)$$

como no se tiene información sobre μ_k se reemplaza por $\mu_{k-1} + \eta_k$, en consecuencia

$$\mu_{k+1} = \mu_{k-1} + \eta_k + \eta_{k+1}, \quad (2.4)$$

llamando $\bar{\eta}_{k+1} = \eta_k + \eta_{k+1}$, se deduce que

$$\text{Var}(\bar{\eta}_{k+1}) = 2\sigma_\eta^2. \quad (2.5)$$

Si se considera que la primera parte de la serie se registra en forma bianual y la segunda en forma trimestral, el filtro se aplica a $FEE(t,1)$ y usa la primera observación como valor inicial para un modelo que considera sólo nivel. Como los componentes estacionales trimestrales no son observados en los datos bianuales, en esta primera etapa, los elementos asociados a ésta no aparecen en el vector de estado. Una vez que se arriba a los datos trimestrales, se expande el vector de estado para incluir los tres elementos estacionales y se necesitarían 4 observaciones como valores iniciales, pero en este caso sólo se requieren 3 porque ya se utilizó la primera observación de la serie bianual como valor inicial del nivel.

La función de verosimilitud exacta se obtiene a través de la descomposición del error usando los errores de predicción.

Usando este método se consigue obtener las dos series en estudio en forma trimestral en el período 1974-2005. Las cuales como se dice anteriormente posee datos estimados en el primer y tercer trimestre en el período 1974-2002.

3 MODELOS MULTIVARIOS DE ESPACIO DE ESTADOS

En forma similar a los modelos de ecuaciones simultáneas aparentemente no correlacionadas (SUR), propuestos por Zellner en 1963, [seemingly unrelated regression equation



(SURE) model], Harvey (1989) propone un sistema como un modelo de ecuaciones de series de tiempo aparentemente no correlacionadas (SUTSE), [seemingly unrelated time series equations a SUTSE model].

Sea \mathbf{y}_t un vector ($N \times 1$) de observaciones, (en este caso $N=2$), el cual depende de componentes no observables que también son vectores. La relación entre las distintas series se puede observar a través de los disturbios de los distintos componentes. En un modelo con factores comunes, algunas o todas las matrices de covariancias, serán de rango reducido.

La especificación del modelo en el caso de las tasas de desocupación donde existe nivel aleatorio y estacionalidad determinística es del tipo:

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon), \\ \boldsymbol{\mu}_t &= \boldsymbol{\mu}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\eta), \\ \gamma_t &= -\sum_{j=1}^3 \gamma_{t-j}, \end{aligned} \tag{3.1}$$

\mathbf{y}_t , $\boldsymbol{\mu}_t$ y $\boldsymbol{\gamma}_t$ son vectores 2×1 ; $\boldsymbol{\varepsilon}_t$ y $\boldsymbol{\eta}_t$ son vectores 2×1 de disturbios normales multivariados, mutuamente no correlacionados en todos los períodos; $\boldsymbol{\Sigma}_\varepsilon$ y $\boldsymbol{\Sigma}_\eta$ son matrices de covariancias 2×2 , las diagonales de estas matrices representan las variancias de los hiperparámetros.

3.1 MEEM PARA LAS SERIES TASAS DE DESOCUPACIÓN DE GR Y GBA

El modelo MEEM para las series trimestrales tasa de desocupación de GR y GBA tiene la siguiente representación:

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon), \quad t = 74.1 \text{ al } 04-4, \\ \boldsymbol{\mu}_t &= \boldsymbol{\mu}_{t-1} + \mathbf{CN}_{95.2} + \mathbf{CN}_{97.2} + \mathbf{CN}_{01.4} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim NID(\mathbf{0}, \boldsymbol{\Sigma}_\eta), \\ \gamma_t &= -\sum_{j=1}^3 \gamma_{t-j}, \end{aligned} \tag{3.1.1}$$

donde **CN** representa el vector de cambio de nivel.

Ambas series presentan nivel aleatorio, estacionalidad fija y cambios de nivel en el segundo trimestre de 1995 y 1997 y en el cuarto trimestre de 2001. Se toman estos tres cambios de nivel, porque son los que presentan para ambas series simultáneamente y además presenta el mejor ajuste.



4 RESULTADOS DE LA ESTIMACIÓN DEL MEE BIVARIADO PARA LAS SERIES TASA DE DESOCUPACIÓN DE GBA Y GR

A continuación se presentan las estimaciones correspondientes al modelo bivariado que considera ambas tasas de desempleo simultáneamente. En todos los casos se trabajó con el programa STAMP 3.0. En la Tabla 4.1 se muestran las estimaciones del nivel al final del estado, los coeficientes estacionales y los correspondientes a los cambios de nivel (CN).

Tabla 4.1: Estimación del modelo para GBA y GR

Coefic.		GBA			GR		
		Estimación	t	Prob.asoc.	Estimación	t	Prob.asoc.
Nivel		13.8	106.3	0.0000	15.4	34.1	0.0000
Estacionalidad	γ_1	0.647			0.669		
	γ_2	0.310	4.0	0.0001	0.345	2.7	0.0074
	γ_3	-0.167	-2.2	0.0301	0.023	0.2	0.8525
	γ_4	-0.790	-10.3	0.0000	-1.037	-8.2	0.0000
CN _{2,95}		6.0	8-1	0.0000	7.0	6.4	0.0000
CN _{2,97}		-2.6	-3.5	0.0006	-3.6	-3.3	0.0015
CN _{4,01}		3.2	4.3	0.0000	3.7	3.4	0.0009

El nivel al final del período, de la serie de tiempo tasa de desocupación de GR es levemente superior a la de GBA.

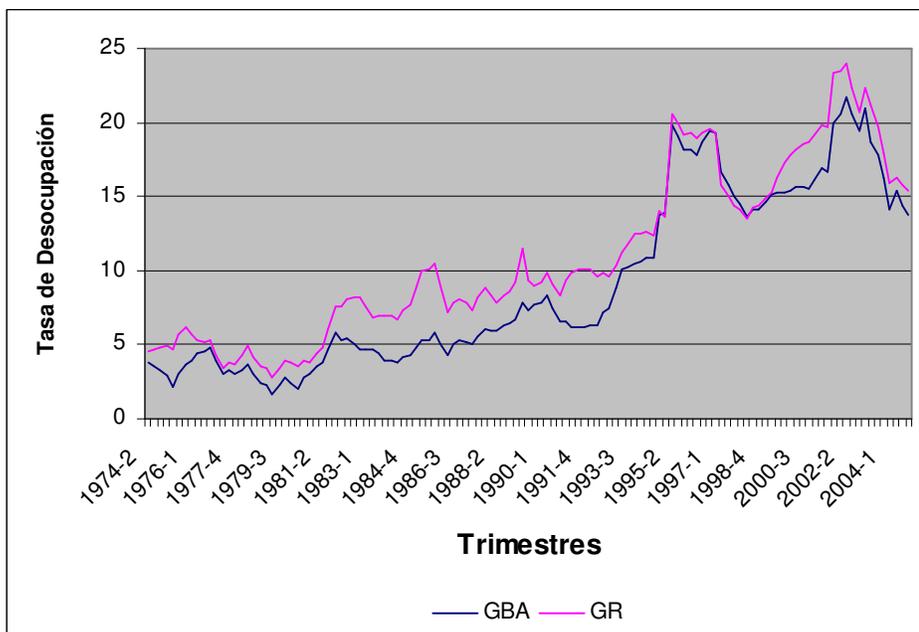
El comportamiento estacional es similar en ambos aglomerados, según el cual es de esperar que la mayor desocupación se presente en el primer trimestre luego vaya descendiendo hasta presentar el punto más bajo en el cuarto trimestre.

Los cambios de nivel, tanto positivos como negativos, siempre son más pronunciados para GR que para GBA. En el segundo trimestre de 1995 se presenta el mayor cambio positivo (7 y 6 puntos respectivamente), en el segundo trimestre de 1997 se da un cambio negativo y en el cuarto trimestre de 2001 nuevamente un cambio positivo.

En el gráfico 4.1 se muestra la tendencia para ambas series.



Gráfico 4.1: Tendencia de las series tasa de desocupación de GBA y GR



En el gráfico 4.2 se presentan los coeficientes estacionales para ambos aglomerados.

Gráfico 4.2: Coeficientes estacionales de las tasas de desocupación de GBA y GR

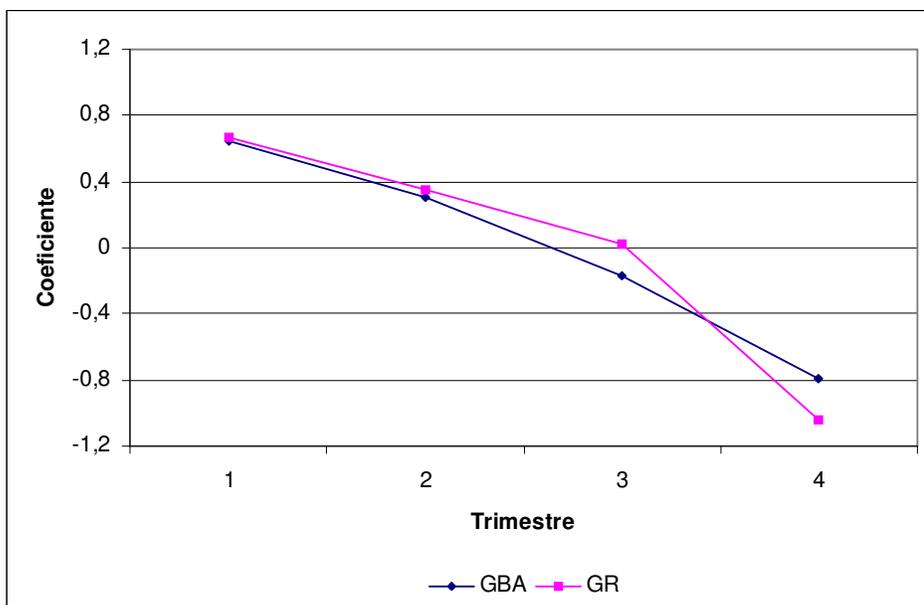




Tabla 4.2: Pronósticos para el año 2005 de la tasa de desocupación GBA y GR

	GBA				GR			
	Pronos.	Inter. Conf.	V. real	% error	Pronos.	Inter. Conf.	V. real	% error
1.05	14.45	13.7-15.2	14.5	0.3	16.1	14.9-17.2	14.0	15.0
2.05	14.11	13.1-15.1	13.4	5.3	15.7	14.3-17.2	12.4	26.6
3.05	13.63	12.4-14.9			15.4	13.7-17.1		
4.05	13.01	11.6-14.5			14.3	12.4-16.2		

En la tabla 4.2 se presentan los pronósticos para el año 2005, para los dos primeros trimestres se cuenta con los valores reales por lo que se calcula el error porcentual en valor absoluto. Estos errores son mucho más chicos para GBA (0.3 y 5.3 para el primer y segundo trimestre respectivamente) que los de GR (15.0 y 26.6). Esto podría estar indicando que los valores de la tasa de desocupación de GR en el primer y segundo trimestre del 2005, son menores que lo esperado, haciendo suponer un nuevo cambio de nivel con signo negativo.

En un análisis previo univariado de la serie de GR (Blaconá, M.T.; Bussi, J., 2005), se determinó la presencia de un cambio de nivel en el tercer trimestre de 2003, pero como este cambio no se presenta en la serie GBA, no se incluyó en el modelo multivariado. Esto podría explicar la sobreestimación de las tasas presentes en los pronósticos para la tasa de desocupación de GR.

5 CONSIDERACIONES FINALES

En este trabajo se realiza un análisis de la series de tiempo tasa de desocupación correspondientes a dos aglomerados urbanos de la República Argentina: Gran Buenos Aires y Gran Rosario. Estas series se presentan con dos periodicidades, de 1974 al 2002 se en forma bianual, mientras que en los años 2003 al 2005 las series son trimestrales.

Para estudiar el comportamiento de la tendencia, la estacionalidad y los puntos de quiebre en las series completas, se emplean los MEE combinando distintas periodicidades, para cada serie en forma univariada, con el fin de completar los datos faltantes, correspondientes al primer y tercer trimestre del período 1974-2002.

Una vez que se completan ambas series, logrando tener información trimestral desde el año 1974, se procede a utilizar los MEE multivariados analizando las series conjuntamente.

Entre las características a resaltar producto de este análisis se puede mencionar:

- i) el nivel de la ambas series es aleatorio
- ii) la estacionalidad es determinística y significativa para ambas series, el comportamiento de los coeficientes es similar para ambos aglomerados, indicando un



- deceso de la desocupación a medida que se avanza en los trimestres del año;
- iii) se presentan tres cambios de nivel comunes a ambas series, en los segundos trimestres de 1995 y 1997 y en el cuarto trimestre de 2001
 - iv) los pronósticos para el primer y segundo trimestre de 2005 para GBA presentan bajos errores de pronóstico, mientras que no ocurre lo mismo para GR, donde los pronóstico sobreestiman la tasa de desocupación real observada para el aglomerado.

De acuerdo a lo expresado en los puntos anteriores, el modelo multivariado presenta resultados muy alentadores para GBA, y superiores al modelo univariado (Blaconá y otros, 2005). En el caso de Rosario, esto no ocurre, lo que podría deberse a que esta serie presenta un cambio de nivel en el tercer trimestre de 2003 (Blaconá, M.T.; Bussi, J., 2005) que no se presenta en GBA, y por lo tanto no fue incluido en el modelo multivariado. Esto estaría mostrando que existe un deceso de la tasa de desocupación en GR que no puede ser explicado por el modelo multivariado, reflejándose este hecho en los pronósticos.

El análisis de las series en forma simultánea completando la información faltante permite hacer un análisis más detallado y completo de las series en la nueva periodicidad con la que se registra la información de la EPH.

Es de interés seguir estudiando distintos enfoques para la inclusión de los cambios de nivel en el modelo multivariado, como así también poder determinar si existe una relación de equilibrio a largo plazo entre las tasas de desocupación considerando los períodos trimestrales.

6 REFERENCIAS BIBLIOGRÁFICAS

- _____. Blaconá, María Teresa, Bussi, Javier. (2004) "Uso del concepto de cointegración para la modelación de la tasa de desocupación", *Actas de la XXXIX Reunión Anual de la Asociación Argentina de Economía Política*, CD. Buenos Aires, 10 -12 noviembre de 2004. Publicado página web: www.aaep.org.ar/anales/publicaciones.
- _____. Blaconá, María Teresa, Bussi, Javier, Ventroni, Nora. (2004) "Comparación de bondad de ajuste y pronósticos de modelos para tasa de desocupación. CLATSE IV, Universidad de Concepción, Chile.
- _____. (2003) "Análisis de la tasa de desocupación de Gran Rosario y Gran Buenos Aires a través de modelos univariados de series de tiempo- (1º 1974 – 2º 2002)", Octavas Jornadas de Investigación, Facultad de Ciencias Económicas y Estadística, UNR. Página Web: www.fcecon.unr.edu.ar/scyt/jor/jor2003.
- _____. Harvey, A.C. (1989). "Forecasting Structural Time Series and the Kalman Filter". Cambridge, U.K. Cambridge, University Press.
- _____. INDEC (2003).^(a). *La nueva Encuesta Permanente de Hogares de Argentina. 2003*.
- _____. INDEC (2003).^(b). *Encuesta Permanente de Hogares (EPH). Cambios metodológicos*.