

PROGRESOS EN ECONOMETRÍA

Editora:

Mariana Marchionni

Autores:

Hildegart Ahumada

Juan Mario Jorrat

Alfredo Martín Navarro

Walter Sosa Escudero

Serie Progresos en Economía



Asociación Argentina
de Economía Política

Marchionni, Mariana (Editora)

Progresos en econometría - 1a ed. - Buenos Aires :

Temas Grupo Editorial: Asociación Argentina de Economía Política, 2005.

150 p. ; 23x16 cm.

ISBN 950-9445-15-0

1. Econometría. I. Título

CDD 330.015 195

Fecha de catalogación: 23/09/2005

© A. A. E. P. - Asociación Argentina de Economía Política

Córdoba 637 piso 4 C1054AAF Ciudad de Buenos Aires

© Temas Grupo Editorial SRL, 2005

Bernardo de Irigoyen 972 piso 9

1072 - Buenos Aires - Argentina

Tel. 4307-4531 y rot.

E-mail: info@editorialtemas.com

www.editorialtemas.com

Derechos reservados en idioma español

Director editorial: Jorge Scarfi

Coordinación general: Julieta Codugnello

Producción editorial: Cecilia Leveratto

Diagramación: Daniel Schapces

Diseño de tapa: Diego Barros

Prohibida la reproducción total o parcial por
cualquier medio sin permiso escrito de la Editorial

PRÓLOGO

La Asociación Argentina de Economía Política tiene como misión “Promover el análisis económico en el país con miras al adelanto de la ciencia”.

Las Reuniones Anuales, de las que este año se cumplirán ya cuarenta ediciones, han sido siempre nuestro principal medio de contribución científica. Han creado un ámbito estable para la presentación de nuestros trabajos, para la evaluación crítica e independiente de la calidad de nuestra producción, y para la difusión de nuestros resultados innovadores.

Las Reuniones Anuales nos han permitido, también, poner a nuestros asociados, en particular a los jóvenes, en contacto con académicos de primer orden internacional.

A partir de noviembre de 2004 hemos tomado la iniciativa de reforzar los aspectos de difusión y formación académica incorporando paneles de Progresos en Economía. Esperamos sean un vehículo eficaz para hacer conocer a toda nuestra comunidad científica los más recientes y destacados avances en cada una de las especialidades de nuestra ciencia.

Este libro es el resultado del primer panel sobre Progresos en Econometría. Agradecemos a los expositores y autores Hildegart Ahumada, Juan Mario Jorrat, Alfredo Martín Navarro y Walter Sosa Escudero, y a Mariana Marchionni en su rol de coordinadora, su esfuerzo y dedicación para llegar a concretar esta edición. Es nuestra aspiración reproducir también como libros nuestras futuras mesas de Progresos en Economía, de modo de ampliar significativamente el acceso de profesores y estudiantes a estas síntesis y reflexiones de destacados investigadores argentinos.

*Consejo Directivo 2005
Asociación Argentina de Economía Política*

BREVE HISTORIA DE LA AAEP

La AAEP fue fundada por el Dr. Juan E. Alemann, el Dr. Roberto Alemann, el Dr. Julio Broide, el Dr. Benjamín Cornejo, el Dr. Aldo Ferrer, el Ing. Francisco García Olano, el Dr. Juan J. Guaresti (h), el Dr. Carlos C. Helbling, el Dr. Carlos Moyano Llerena, el Dr. Julio H. G. Olivera, el Dr. Federico Pinedo, el Dr. Oreste Popescu, el Dr. Ovidio Schiopetto y el Dr. Francisco Valsecchi.

El origen de la AAEP se remonta a sendas invitaciones cursadas por los Dres. Adolph Jöhr y Louis Baudin, a mediados de la década de los cincuenta, a los Dres. Oreste Popescu y Julio H. G. Olivera.* Jöhr y Baudin, por entonces pertenecientes a la International Economic Association, sugirieron constituir una asociación representativa de los investigadores en «economía política». La convergencia de iniciativas de los Dres. Popescu y Olivera cristalizó, el 26 de setiembre de 1957, con la decisión de crear la AAEP. El Dr. Olivera llevó adelante la ejecución de la fundación, la que tuvo lugar formalmente el 18 de noviembre del mismo año.

La historia de la Asociación puede dividirse en tres fases. Durante la primera etapa (1957-1965) la actividad se tradujo en encuentros periódicos para la discusión de temas específicos. En 1958 se realizó la primera reunión de análisis económico. Durante este período la AAEP constituyó varios «centros regionales». La segunda etapa (1966-1972) se caracterizó por la incorporación a la AAEP de representantes de centros e institutos de investigación. A partir de entonces, las reuniones de centros de investigación económica se realizaron en el marco de la AAEP. Se inició en 1968 la rotación de la sede y de las autoridades ejecutivas entre los principales centros. En 1972 tuvo lugar la última reunión de la AAEP organizada sobre la base de trabajos de los centros e institutos de investigación. Desde 1973 hasta el presente la AAEP se encuentra en su tercera etapa, con su sede permanente en la ciudad de Buenos Aires. La AAEP es una institución formada por investigadores y académicos en economía y que interactúa en forma directa con los mismos. Pese a ello, el espíritu de una amplia representación institucional y regional ha quedado impreso en la práctica de las Reuniones Anuales.

Desde su fundación, la AAEP fue presidida por Julio H. G. Olivera (1957/68), Benjamín Cornejo (1968/70), Víctor J. Elías (1970/72 y 1978/80), Miguel E. Martínez (1972/74), Horacio Núñez Miñana (1974/76), Aldo A. Arnaldo (1976/78), Rolf R. Mantel (1980/82), Mario L. Szychowski (1982/84), Ana M. Martirena Mantel (1984/86), Luisa Montuschi (1986/88), Alfredo M. Navarro (1988/90), Rinaldo Colomé (1990/92), Juan C. De Pablo (1992/94), Eusebio C. Del Rey (1994/96), Enrique A. Bour (1996/98), José A. Delfino (1998/00), Hildegart Ahumada (2000/02), José Luis Arrufat (2002/04) y Omar O. Chisari (2004/2006).

QUE ES LA ASOCIACION ARGENTINA DE ECONOMIA POLITICA

Los recursos de la AAEP provienen de las cuotas sociales y de los aportes de entidades adherentes. Actualmente la AAEP cuenta con más de 400 socios activos y 17 entidades adherentes: Banco Central de la República Argentina, Centro de Estudios de Estado y Sociedad (Cedes), Comisión Económica para América Latina y el Caribe (CEPAL), DePablo Consult, Estudio Broda & Asoc., Fundación de Investigaciones Económicas Latinoamericanas (FIEL), Fundación Capital, Internacional Monetary Fund (IMF), Instituto Torcuato Di Tella (ITDT), Nacional Association For Business Economics (NABE), Orlando J. Ferreres y Asociados, Universidad Argentina de la Empresa (UADE), Universidad del CEMA (UCEMA), Universidad Nacional de Córdoba, Universidad Nacional de Cuyo, Universidad Nacional de La Plata y Universidad Nacional del Sur.

La principal actividad de la AAEP es la celebración de una Reunión Anual de discusión de trabajos realizados en el ámbito de la economía, tanto por socios como por no socios, la que tradicionalmente se celebra el mes de noviembre de cada año en distintas ciudades del país, en consulta y cooperación con universidades y centros de investigación económica de la Argentina. En dichas reuniones han participado economistas del país e invitados de otros países, incluyendo prestigiosos economistas extranjeros como Albert Berry, M. Bruno, V. Corbo, J. Drèze, S. Fisher, R. Guesnerie, A.C. Harberger, H. Houthakker, J.J. Laffont, A. Leijonhufvud, Sir J. Mirrlees (Premio Nobel), F. Modigliani (Premio Nobel), M. Nerlove, L. Pasinetti, S. Rosen, P. Spiller, J. Tobin (Premio

Nobel), W. Oates, V. Volsky E. Prescott (Premio Nobel), T.N. Srinivasan, Finn Kydland (Premio Nobel), Aloisio Araujo. Los trabajos son aceptados por una comisión de socios de reconocidos méritos científicos y académicos, designada por el Consejo Directivo.

Desde 1964, la AAEP ha llevado adelante 39 Reuniones Anuales en las cuales han sido presentados y discutidos, en un marco de total libertad académica, 1858 trabajos de investigación. Los trabajos presentados en cada Reunión Anual son editados en los Anales de la Asociación Argentina de Economía Política. A estos trabajos hay que sumar los trabajos presentados antes de 1964 y en reuniones científicas, no incluidos en Anales. A partir de 1997, los trabajos y comentarios están incluidos también en el web site de la AAEP en Internet. La AAEP no es responsable ni de las opiniones incluidas en los Anales, ni de su protección intelectual.

La AAEP también organiza, en forma no sistemática y en la medida de sus posibilidades financieras, reuniones científicas a lo largo del año, en colaboración con otros organismos. Es miembro de la International Economic Association y mantiene relaciones con instituciones similares de otros países.

El Consejo Directivo de la AAEP es el órgano de gobierno de la AAEP, y está compuesto de un presidente, dos secretarios y nueve vocales, provenientes de distintos organismos y regiones del país.

El Dr. Julio H. G. Olivera es Presidente Honorario de la AAEP. Los Profesores Albert Berry, Vittorio Corbo, Jacques Drèze, Roger Guesnerie, Arnold C. Harberger, Jean-Jacques Laffont (fallecido), Axel Leijonhufvud, James Mirrlees, Marc Nerlove, Wallace Oates, Alberto Petrecolla, Sherwin Rosen, Pablo Spiller, T.N. Srinivasan, Aloisio Araujo y Finn Kydland son Miembros Honorarios.

SEDE DE LA AAEP

Av. Córdoba 637 - 4º piso - (1054) Buenos Aires - Argentina

Tel. (5411) 4314-0246 Fax (5411) 4314-8648 E-mail aaep@fiel.org.ar

Web-Site de la AAEP en INTERNET: <http://www.aaep.org.ar>

* Para esta sección, véase J. H. G. Olivera, La Asociación Argentina de Economía Política: los Años Iniciales, Anales de la A.A.E.P., XXIIa. Reunión Anual, Universidad Nacional de Córdoba, 1987, vol. I. Se agradecen al Dr. Olivera los comentarios formulados sobre los orígenes de la AAEP.»

CONSEJO DIRECTIVO DE LA ASOCIACIÓN ARGENTINA DE ECONOMÍA POLÍTICA

	2004	2005
Presidente	José Luis Arrufat	Omar O. Chisari
Presidente Electo	Omar O. Chisari	
Secretaria	Nuria Susmel	Cynthia Moskovits
Secretario	Ángel E. Neder	Ángel E. Neder
Vocal	Jorge Baldrich	Jorge Baldrich
Vocal	Manuel Fernández López	Enrique Folcini
Vocal	Enrique Folcini	Juan Mario Jorrat
Vocal	Juan Mario Jorrat	Miguel Lliteras
Vocal	Miguel Lliteras	Mariana Marchionni
Vocal	Mariana Marchionni	Ernesto Schargrodsky
Vocal	Ernesto Schargrodsky	Juan Verstraete
Vocal	Jorge Streb	Lidia Rosignuolo
Vocal	Juan Verstraete	Germán Coloma
Secretario Supl.	José M. Bulacio	José M. Bulacio
Secretario Supl.	Ricardo Crespo	Miguel Angel Asencio
Vocal Suplente	Héctor Benegas Prado	Héctor Benegas Prado
Vocal Suplente	Miguel A. Broda	Miguel A. Broda
Vocal Suplente	Andrés Chambouleyrón	Andrés Chambouleyrón
Vocal Suplente	Walter Cont	Walter Cont
Vocal Suplente	Guillermo Escudé	José María Fanelli
Vocal Suplente	José María Fanelli	Daniel Heymann
Vocal Suplente	Daniel Heymann	Carlos Rojas
Vocal Suplente	Carlos Rojas	Ernesto Seselovsky
Vocal Suplente	Ernesto Seselovsky	Laura D'Amato

SOBRE LOS AUTORES

Hildegart Ahumada

Es Licenciada en Economía de la Universidad Nacional de La Plata y Master of Philosophy in Economics de la Universidad de Oxford. Actualmente desarrolla su labor docente en la Universidad Torcuato Di Tella como profesora de Estadística y Econometría. Fue miembro del Comité Académico de la Rama Latinoamericana de la Sociedad Econométrica (1991-2001) y presidente de la Asociación Argentina de Economía Política (2001-2002). Fue distinguida con una beca de la J.S. Guggenheim Memorial Foundation (New York) y con el Premio Fulvio Salvador Pagani 1994 de la Fundación Arcor. Sus publicaciones se encuentran en prestigiosas revistas locales e internacionales.

Juan Mario Jorrat

Es Contador Público Nacional de la Universidad Nacional de Tucumán y Master of Arts en Economía de la Ohio State University. Desarrolla su actividad docente como profesor titular de Comercio Internacional en la Universidad Nacional de Tucumán desde 1983. En 1994 el equipo de investigadores a su cargo estableció una primera lista de series coincidentes y líderes para Argentina y construyó los índices compuestos correspondientes. Desde 1995 dirige el programa de investigación Ciclos Económicos de Argentina, aprobado y financiado parcialmente por el Consejo de Investigaciones de la UNT, y que fue declarado de interés del Honorable Senado de la Nación Argentina en 1996.

Mariana Marchionni

Es Licenciada en Economía, Mg. en Economía y Dra. en Economía de la Universidad Nacional de La Plata, donde actualmente se desempeña como Profesora de Econometría e investigadora del Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS). Fue coordinadora del Panel de Progresos en Econometría en la XXXIX Reunión Anual de la Asociación Argentina de Eco-

nomía Política y editora de este volumen. Recibió el Premio Fulvio Salvador Pagani 2001 de la Fundación Arcor y becas del Programa Mecovi-Argentina (2001) y de la Academia Nacional de Ciencias Económicas (2004). Se especializa en Econometría aplicada a temas sociales y en Economía de la Demografía.

Alfredo Martín Navarro

Es Doctor en Ciencias Económicas graduado en la Universidad de Buenos Aires. Es miembro titular de la Academia Nacional de Ciencias Económicas. Fue Presidente de la Asociación Argentina de Economía Política entre 1988 y 1990. Ha ejercido la docencia en diversas universidades estatales y privadas y sus trabajos de investigación están relacionados con aplicaciones econométricas y aspectos metodológicos de la Economía.

Walter Sosa Escudero

Es licenciado en Economía de la Universidad de Buenos Aires y Ph.D. en Economía de la Universidad de Illinois (Urbana-Champaign). Es profesor de Econometría y director de la Maestría de Economía de la Universidad de San Andrés, y profesor e investigador visitante en el Departamento de Economía y en el CEDLAS de la Universidad Nacional de La Plata. Fue distinguido en 2003 con el Premio Bernardo Houssay al Investigador Joven y en 2001 con el Premio Fulvio Salvador Pagani de la Fundación Arcor. Su investigación se concentra fundamentalmente en Econometría con aplicaciones a Microeconomía y cuenta con numerosas publicaciones en destacadas revistas académicas nacionales e internacionales.

ÍNDICE

1. Introducción	1
<i>Mariana Marchionni</i>	
2. Reflexiones sobre el estado actual de la metodología de la Econometría	5
<i>Alfredo Martín Navarro</i>	
3. Pronóstico con modelos econométricos	27
<i>Hildegart Ahumada</i>	
4. Construcción de índices compuestos mensuales coincidente y líder de Argentina	43
<i>Juan Mario Jorrat</i>	
5. Perspectivas y avances recientes en regresión por cuantiles	101
<i>Walter Sosa Escudero</i>	

INTRODUCCIÓN

MARIANA MARCHIONNI

CEDLAS - UNIVERSIDAD NACIONAL DE LA PLATA

La Econometría ha recorrido un largo camino en un lapso de tiempo relativamente corto. Citando a Clive W. J. Granger en un pasaje distendido de su discurso al recibir el Premio Nobel en Economía en 2003:

“...The first Nobel Prizes in Economics were to econometricians, Ragnar Frisch and Jan Tinbergen, for which we are very proud. There are now eight of us with the Prize, representing 15 percent of Economics winners. However, in the current millennium, we represent about 44 percent of the winners, which I view as a healthy local trend.”

Como señala Pesaran (1987), el progreso de la Econometría se produjo en todos los frentes: se han hecho avances considerables en la compilación de datos económicos y en el desarrollo de conceptos, teorías y herramientas econométricas, y en prácticamente todos los campos de la Economía pueden encontrarse aplicaciones basadas en métodos econométricos. Y si bien la Econometría sufre de importantes limitaciones, en parte originadas en la incompletitud de las teorías Económicas, no puede desconocerse el papel fundamental que ha jugado en el progreso de la Economía como ciencia.

Los cuatro trabajos que se compilan en este libro fueron presentados en el Panel de Progresos en Econometría que me tocó coordinar durante la XXXIX Reunión Anual de la Asociación Argentina de Economía Política, el 11 de noviembre de 2004 en Buenos Aires. El objetivo de este panel fue el de reunir a un grupo de destacados econométricos argentinos para discutir los desarrollos recientes tendientes a superar algunas de las limita-

ciones que la Econometría ha ido encontrando a lo largo del camino.

Según Heckman (2001), otro econometrista que obtuvo el Premio Nobel en 2000, la Econometría es útil sólo si ayuda a los economistas a conducir e interpretar investigaciones empíricas sobre la base de datos económicos. Es posible que la mayoría de los economistas y, en particular, de los econometristas, estemos más o menos de acuerdo con esta afirmación. Seguramente haya menos acuerdo sobre qué pensamos acerca de cómo se deberían “conducir” e “interpretar” las investigaciones empíricas y sus resultados. Y creo que encontraríamos aún más diferencias en la forma en que finalmente se materializan en la práctica esa “conducción” e “interpretación”. En el primero de los trabajos, Alfredo M. Navarro reflexiona sobre la manera en que la Econometría se desempeña como nexo entre la Teoría Económica y los datos económicos, presentando y discutiendo apreciaciones propias y de otros autores sobre la metodología de la Econometría.

La naturaleza particular de los datos económicos justifica el desprendimiento de la Econometría del cuerpo central de la Estadística. De la misma forma, la diversidad de técnicas econométricas y de áreas de especialización para los econometristas responde, en gran medida, a las características particulares de los distintos tipos de datos económicos cuya clasificación básica distingue los datos de series de tiempo de los de corte transversal. En los siguientes dos trabajos se presentan y discuten métodos propios del análisis de series de tiempo. En el primero, Hildegart Ahumada se refiere a los pronósticos económicos basados en modelos econométricos y a los determinantes del desempeño de esos pronósticos, discutiendo el efecto de diferentes estrategias de modelización. En el segundo, Juan Mario Jorrot también aborda el tema del pronóstico, pero en este caso aplicado directamente a predecir los niveles de actividad económica agregada. Su metodología se basa en el análisis del comportamiento de índices de actividad construidos a partir de un conjunto de variables económicas con el objetivo de monitorear la actividad económica del país y predecir futuras recesiones o recuperaciones.

Finalmente, Walter Sosa Escudero dedica un espacio al tratamiento de una metodología que, aunque últimamente ha sido extendida al análisis de series temporales, se desarrolló en el ámbito de los datos de corte transversal: el modelo de regresión por cuantiles. Su interés consiste en modelar el efecto de un grupo de factores sobre toda la distribución condicional de la variable dependiente, en contraposición a los modelos de regresión tradicionales que buscan únicamente caracterizar el efecto sobre la media condicional de esa variable y que, en palabras del autor, “pueden verse como resúmenes, a veces relevantes y otras veces no, de los modelos de regresión por cuantiles”.

REFERENCIAS

- Granger, Clive W. J. 2004. Time Series Analysis, Cointegration, and Applications. *The American Economic Review* 94(3): 421-425.
- Heckman, James J. 2001. Econometrics and empirical economics. *Journal of Econometrics* 100: 3-5.
- Pesaran , M. Hashem. (1987). Econometrics, en Eatwell, J., Milgate, M. y Newman, P. (eds). *The New Palgrave: A Dictionary of Economics*, Macmillan, New York.

REFLEXIONES SOBRE EL ESTADO ACTUAL DE LA METODOLOGÍA DE LA ECONOMETRÍA

ALFREDO MARTÍN NAVARRO

ACADEMIA NACIONAL DE CIENCIAS ECONÓMICAS

1. INTRODUCCIÓN

Los economistas que nos formamos al comienzo de la segunda mitad del siglo pasado presenciamos un gran cambio en las características de nuestra ciencia. Lo mismo les ocurrió, según nos cuenta Skidelsky (1983), a Marshall y a Sidgwick a fines del siglo XIX, cuando entró en crisis el mundo victoriano, con respecto a los economistas clásicos, y luego a Keynes y sus contemporáneos, respecto de ellos. Es que cambia la economía, porque cambia el mundo que los economistas estudiamos, y esos cambios se deben a acontecimientos de carácter social, pero también a la evolución de las ideas y de las demás disciplinas científicas.

El cambio que se produce en los años setenta con respecto a la actitud de los economistas respecto a la Econometría está muy bien descrito por Klammer (2001)¹:

“When I went to the university in the early seventies and opted for the study of econometrics...I wanted to work on the big economic problems such as inequality and poverty, and expected that econometrics would produce the best scientific solutions... My hero

¹ En 1924 Keynes en su nota necrológica sobre Marshall, decía, refiriéndose al estudio de la economía: “*It is not ... a very easy subject compared with the higher branches of philosophy and pure science*”, y con respecto a la formación de los economistas: “*He must be mathematician, historian, statesman, philosopher, in some degree. He must understand symbols and speak in words. He must be purposeful and disinterested in a simultaneous*

was Jan Tinbergen, the Dutch econometrician, who dedicated his life to the cause of justice and just had earned the Nobel Prize. His fusion of idealism with hard science seemed ideal...

The frustration came when I found out that econometrics was all algebra and mathematical statistics and economic topics hardly came up in the discussions. The second frustration came when I found out that even econometricians did not really know what was going on in the world. They ran their regressions but did not settle any theoretical issues. The inconclusiveness of their research did not bother them as they were preoccupied with mathematical issues...I had already resigned myself to the impotence of the science of economics to be an objective guide to a better world. Jan Tinbergen the scientist had become for me Jan Tinbergen the idealist.”

Esta experiencia es bastante similar a lo que nos ha sucedido a muchos economistas de todo el mundo, que hemos tenido las mismas dudas. Algunos dejaron la Econometría de lado, como Klammer, mientras que otros la seguimos utilizando, pero cuestionando el valor científico de los resultados obtenidos. Luego en los ochenta se produjo una profunda transformación en la Econometría: por un lado, la crítica de Lucas, referida a la estabilidad de los coeficientes ante cambios en las expectativas de los agentes, que golpeó duramente a los modelos paradigmáticos de los sesenta, compuestos por centenares de ecuaciones, que tenían la magia de la tecnología ultramoderna y que parecían irrefutables desde el punto de vista científico y por el otro la

mood, as aloof and incorruptible as an artist, yet sometimes as near the earth as a politician. (Keynes, 1972 [1924]). Es evidente el cambio operado en el perfil del economista en los ochenta años que median entre la redacción de esta nota y nuestros días. Seguramente el cambio es mucho mayor que el operado en el siglo precedente, ya que el ideal de Keynes no difiere de demasiado, por ejemplo, del de Stuart Mill. El cambio operado desde los sesenta hasta nuestros días es mucho mayor. Basta mirar una revista especializada, como el JPE o AER y comparar la naturaleza de los artículos y el grado de complejidad de las técnicas empleadas para tomar conciencia del enorme cambio operado en un período relativamente breve.

crítica de Sims (1980), quien sostiene que la identificación de esos modelos, tal como la habían definido los econométricos de la Cowles Commission, estaba basada en restricciones arbitrarias que no tenían fundamento económico. El carácter exógeno de ciertas variables, especialmente las monetarias, queda en duda cuando se observa que existen influencias de éstas sobre las exógenas, por lo que su supuesta exogeneidad pasa a ser cuestionada.

Estas dos cuestiones fueron los fundamentos principales de la aparición de la metodología VAR, de la conocida como “general to specific”, como la define Hendry, y de una importante serie de desarrollos en el análisis de las series temporales, así como técnicas econométricas para tratar los datos organizados en forma de panel. La aparición de las computadoras personales, que estuvieron al alcance de todos en los primeros años de la década de los ochenta y de paquetes de software cada vez más perfeccionados provocaron un desarrollo impensado en la utilización de los métodos econométricos, aunque las dudas respecto a qué estamos haciendo cuando realizamos regresiones u otros métodos econométricos sigue frente a nosotros. Por eso en este trabajo voy a tratar de encontrar alguna respuesta parcial a esa cuestión. Los economistas debemos preguntarnos qué estamos haciendo exactamente cuando aplicamos técnicas econométricas, para poder evaluar adecuadamente el valor de esa tarea para explicar el mundo que nos rodea y para pronosticar el comportamiento futuro de las variables que por alguna razón nos interesa analizar.

Es conveniente destacar que hay dos formas de encarar la metodología, de acuerdo a Machlup (1963): una es la que denomina *exógena*, y es la que realizan especialistas en filosofía de la ciencia, que no son economistas, y otra *endógena*, que es la que hacemos los propios economistas. En este trabajo nos referiremos a temas *endógenos*, porque son los que corresponde que analicemos los economistas.

Recientemente se han planteado muchas cuestiones interesantes relacionadas con la *metodología econométrica endógena*, de las cuales he seleccionado, tal vez arbitrariamente, algunas de las que me parecen de mayor relevancia. En primer lugar, voy a considerar el pensamiento de McCloskey y sus críticas a la Econometría y a la forma en que la estamos aplicando. En segundo término me

voy a referir a una nueva visión de la práctica conocida como “*data mining*”. Luego voy a tratar algunas consideraciones que se ha realizado con motivo de cumplirse el cincuenta aniversario de la publicación de la obra de Friedman (1953), “*The Methodology of Positive Economics*” y a su vigencia actual. Posteriormente voy a referirme al método conocido como *calibración*, que si bien hace ya veinte años que fue aplicado por primera vez, ha cobrado actualidad luego de haber obtenido el premio Nobel sus autores y por último trataré el cambio de actitud de los econométricos ante la teoría económica. Para finalizar voy a exponer algunas conclusiones propias.

2. McCLOSKEY Y LA RETÓRICA

2.1. La retórica y la economía

El término retórica tiene dos significados: por una parte es “el arte de bien decir, de embellecer la expresión de los conceptos, de dar al lenguaje escrito o hablado eficacia bastante para deleitar, *persuadir* o conmover” y por otra parte también significa “el uso impropio o intempestivo de este arte”². Es decir que hay dos retóricas, una buena, que sirve para persuadir de algo con medios honorables y otra mala, que es la que se utiliza para convencer a otro de algo que no es cierto. Pues bien, McCloskey (1998) sostiene que los científicos en general y los economistas en particular cuando escribimos nuestros trabajos estamos utilizando la retórica en la primera de las acepciones, para *persuadir* a nuestros colegas de que lo que *decimos es cierto*. Pero al poner el eje de la discusión en que el objetivo de la labor del científico *es convencer* a los demás (y *los demás* no pueden ser otros que los colegas, dado que sólo ellos comprenden el lenguaje utilizado), cambia el objetivo central de la tarea de los científicos, que ya no es el descubrimiento de la *verdad*, sino *el convencer* a los colegas de que lo que estamos afirmando es cierto.

² Esto según el Diccionario de la Lengua Española de la Real Academia Española.

2.2. Los dos pecados de la economía contemporánea

Así se titula el último libro de McCloskey (2002), y los dos pecados a los que se refiere son, en primer término, la práctica de construir modelos que no pueden confrontarse con la realidad, y en segundo lugar, la utilización de “tests” estadísticos, como por ejemplo, el conocido como “t de Student”, para aceptar o rechazar una teoría. De ambos responsabiliza a Samuelson, por el primero porque sus modelos son exclusivamente matemáticos y no verificables empíricamente, y por el segundo, porque según dice fue el director de la tesis de Klein, que utiliza el “test t” en sus regresiones en forma que considera inapropiada. Por supuesto, esto es una injusticia, y ella seguramente lo sabe, pero lo presenta de esta manera para hacer entender su mensaje.

¿Por qué considera pecado hacer modelos sin comprobación empírica? Porque define la ciencia como un “inquiry” acerca del mundo, lo que implica teorizar y observar, formalizar y ver los hechos, ambas cosas. Si hacemos una sola estamos desvirtuando la labor del científico, y con este sistema llegamos a un nivel de abstracción de la realidad que hace que los modelos sean similares a problemas abstractos de ajedrez, que ni siquiera sirven para ganar partidas de ajedrez. Ni los físicos llegan a este nivel de abstracción. En síntesis, hemos ido demasiado lejos en nuestro vuelo teórico y debemos pisar tierra firme.

El problema quedaría solucionado si realizáramos análisis empíricos para confirmar nuestros resultados. Pero la forma en que interpretamos los resultados de las regresiones tampoco es la correcta, ya que antepone la “significatividad estadística” a la “*significatividad económica*”. ¿Y cuál es la diferencia? Supongamos que estamos estudiando la paridad del poder adquisitivo, por lo que estimando la ecuación

$$p_t = \alpha p_t^*/e_t + \varepsilon_t$$

(donde p representa los precios domésticos, p^* los externos y e el tipo de cambio nominal, α un coeficiente de valor desconocido, ε , un término aleatorio con las propiedades habituales y t , el tiempo).

Si deseamos verificar empíricamente la teoría de la paridad del poder adquisitivo debemos obtener en la regresión un coeficiente á de valor unitario. Si la muestra es muy grande, es posible que obtengamos, al aplicar el test t que el coeficiente es significativamente distinto de la unidad, no obstante debemos considerar *cuánto* se aparta de uno para saber si la teoría debe ser descartada o aceptada. Tal vez un valor (por ejemplo) de .93 sea suficiente-mente cercano a uno como para que demos la teoría por cierta. Otro ejemplo es el siguiente: supongamos, dice McCloskey (2000) que estamos analizando la relación entre el producto bruto de Estados Unidos y el rendimiento de los cultivos de maíz en el estado de Iowa y que hacemos una regresión entre ambas variables y obtenemos que el valor del test “t” del coeficiente de esta última variable resulta significativamente distinta de cero de acuerdo a los criterios habituales. Ese resultado puede indicar que la variable tiene alguna influencia³, aunque el sentido común nos dice que no es así.

Se podría argumentar que existe una división del trabajo entre quienes realizan tareas teóricas y quienes buscan su comprobación empírica. Pero se opone a esta solución porque los modelos matemáticos están construidos de tal manera que no son aptos para ser verificados empíricamente. Si miramos una revista de física, afirma, veremos que la gran mayoría de los artículos presentan un modelo teórico y a continuación el resultado de un experimento o algún tipo de comprobación empírica, cosa que no ocurre en un número muy importante de artículos económicos. Einstein, en un artículo publicado en 1950, decía⁴:

“It may well be true that this system of equations is reasonable from a logical standpoint. But it does not prove that it corresponds with nature... Experience alone can decide about the truth...”

³ Debe tenerse en cuenta que cuando evaluamos el valor del coeficiente con respecto a su desviación standard generalmente no tenemos en cuenta que el valor del coeficiente varía a lo largo de la muestra, y que en ciertas partes de ella puede no ser significativamente distinto del valor de referencia.

⁴ Ver Navarro (2003).

Smith (1999) va aún más lejos que McCloskey cuando analiza la aplicación de algunos de los avances recientes en Econometría: raíces unitarias, VAR, causalidad de Granger y cointegración. Sostiene que recoger y elaborar los datos con estas técnicas no resuelve de por sí los problemas bajo análisis, sino que deben en todos los casos realizarse interpretaciones económicas. Si vemos a un timonel mover el timón frenéticamente mientras el barco se desliza con normalidad pueden ocurrir dos cosas: el timonel es eficiente o los controles no funcionan. Aquí lo importante es identificar el modelo para saber qué está sucediendo, y esto sólo puede hacerse aplicando la teoría económica, para poder interpretar los datos. Es decir, es preciso aplicar criterios económicos, que sinteticen los resultados estadísticos. Durante algún tiempo algunos economistas pensaron que estas técnicas de por sí eran suficientes para verificar o falsificar las teorías económicas, pero la amplia confianza de un primer momento que nos llevó a sobreestimar su poder, fue diluyéndose paulatinamente a través de los últimos años.

2.3. Caminos alternativos

La pregunta que nos formulamos después de analizar las ideas de McCloskey es qué hacemos entonces. Si utilizamos las técnicas econométricas en la forma convencional podemos exigirnos una lectura inteligente de los tests, pero no prescindir de ellos. Solamente encontré en sus obras una reflexión sobre este tema, y sugiere utilizar las técnicas de VAR y de Leamer⁵, la primera mucho más orientada a dejar hablar a los datos, con modelos con un propósito diferente, orientados a mostrar la interrelación entre las variables y sacar información de la realidad para elaborar o modificar las teorías, pero sin adherirse tan fuertemente a las teorías como hace la Econometría tradicional. Sin embargo, si bien Sims (1996) coincide con McCloskey en algunos aspectos, como en

⁵ La propuesta de Leamer (1983), conocida como Extreme Bound Análisis, tiene como punto de contacto, a mi juicio, con las ideas de McCloskey el hecho de que cuestiona la significatividad estadística y es más exigente a la hora de considerar la inclusión o no de una variable en un modelo.

que las teorías no son “verdaderas” o “falsas”, y en que para “testearlas” no es suficiente una sola confrontación con los datos, así como en su desconfianza de la ortodoxia, de las jerarquías y de las prescripciones metodológicas, considera que su entusiasmo por la retórica tiene el riesgo de empujar la economía hacia la charlatanería. Analiza la situación de dos teorías, una sencilla y otra complicada, y se pregunta si el mayor atractivo de la primera por su sencillez no puede conducirnos a un camino equivocado, y con respecto a las normas estrictas necesarias en la ciencia, dice lo siguiente:

“The only kind of argument that modern science treats as legitimate concern is the match of theory to data generated for experiment or observation... The fact that econometrics is not physics does not mean that we should not aim to apply the same fundamental standards for what constitutes legitimate arguments...and is not legitimate protect theories from the data...I think many economists now see themselves as experts in persuasion as much as experts in substantive knowledge”

3. NUEVA ACTITUD ANTE LA PRÁCTICA DEL “*DATA MINING*”

La práctica conocida como “*data mining*” en forma peyorativa consiste en que, una vez realizada una regresión, se realicen todos los cambios necesarios hasta lograr determinada una especificación, que se considere apropiada por alguna razón, presentándose solamente los resultados de esta última regresión, a pesar de haberse realizado una cantidad de ellas⁶. Dichos cambios pueden ser tanto en las variables explicativas incluidas, como en la estructura de rezagos, en el método de estimación, en la forma funcional elegida y en la muestra utilizada (es decir, cambiando la periodicidad cuando se trabaja con series temporales o alterando el período muestral).

⁶ Darnell (1994) define la práctica de “*data mining*” en términos similares. Esta práctica, cuando se hace en forma deshonesto, nos lleva a resultados inapropiados, ya que aunque sea por casualidad, algunas variables pueden resultar significativas.

Esta es una cuestión que importa a todos los economistas, sean o no econométricos, dado que quienes leen un trabajo deben saber hasta qué punto pueden creer en las conclusiones obtenidas⁷. Mayer (2000) considera que lo importante es saber cómo se debe informar al lector, a quien se trata de convencer pero no de engañar. Se pregunta entonces cuáles de los resultados obtenidos deben ser comunicados al lector y cuáles no. No es un problema sencillo, y lo más razonable es aplicar el sentido común. La práctica conocida como “*data mining*” casi nunca es definida, pocas veces defendida, muy criticada, pero ampliamente utilizada. Según Mayer lo que importa no es si se realiza o no esta práctica, sino la forma en que es comunicada. Es aceptable que a veces algunos resultados no se expongan en el trabajo, por ejemplo si en una de las regresiones el coeficiente del precio de una curva de demanda tiene signo positivo, pero no es lícito hacerlo en caso de que se obtengan resultados diferentes de los esperados.

Si miramos esta práctica desde la forma tradicional de la Econometría, es censurable: una vez definido el modelo teórico debemos confrontarlo con la realidad, y uno de los caminos posibles para lograr tal fin es el uso de la técnica de la regresión múltiple. Las variables incorporadas deben ser solamente las preestablecidas por la teoría. No son los datos los que establecen las características de la ecuación a estimar, sino la teoría diseñada con anterioridad, y la práctica de correr regresiones hasta que una arroje un satisfactorio coeficiente de determinación, coeficientes significativos estadísticamente y un DW cercano a dos, sin referencia al camino recorrido es un procedimiento que sin mayor análisis no puede recibir otra cosa que reprobación.

Pero si bien es cierto que todos la hemos criticado, siempre se ha practicado, a veces en forma explícita y otras en forma no explícita. Pagan y Veall (2000) cuentan que en una de sus películas, Woody Allen visita a su siquiatra y le dice que su hermano cree ser una gallina. El médico le sugiere que lo traiga a su

⁷ Mayer (2000) cita una encuesta cuyos resultados dicen que el 56% de los economistas desconfía un poco de los resultados econométricos, un 27% desconfía bastante y un 2% rechaza totalmente los resultados de los trabajos econométricos.

clínica para tratarlo, pero se niega porque manifiesta que su familia “necesita los huevos que pone” (*I can't, we need the eggs*). Esto, dicen, es algo similar a lo que ocurre con el uso del “*data mining*”, ya que todos sabemos que es incorrecto su uso, pero lo necesitamos para sobrevivir. Hoover y Perez (2000) sostienen que hay tres actitudes ante la práctica del “*data mining*”: una de rechazo, otra que sigue los lineamientos de Leamer (1983) y la considera inevitable y sólo importan aquellos resultados que sobreviven a múltiples especificaciones y por último la que considera que es esencial, y que debe hacerse en forma inteligente, como se hace, por ejemplo en el método denominado LSE (London School of Economics), donde se parte de un modelo general que se va reduciendo paulatinamente en base a criterios estadísticos.

Un ejemplo de la aplicación de métodos que se puede considerarse bastante cercanos al “*data mining*” es el debate acerca de la forma de determinar las variables que determinan el crecimiento económico. Levine Y Renelt (1992), continuando trabajos anteriores de Barro (1991 y otros, estudian una función del siguiente tipo:

$$Y = \beta I + \beta_m M + \beta_z Z + u$$

(donde Y es la tasa de crecimiento del PBI, **I**, un vector compuesto por tres variables que siempre están presentes, M la variable de interés, y **Z**, un vector de variables que se van cambiando, β_k (para $k = i, m, z$), vectores de coeficientes, y u un término de error con las propiedades habituales. Aplican el método de Leamer, que permite analizar qué sucede con la significatividad y el signo de cada uno de los coeficientes al producirse cambios en los otros regresores, método que se conoce como “*Extreme Bound Análisis*”⁸. Posteriormente Sala-i-Martin (1997) aplica una metodología similar, para lo que, según dice el título del trabajo, corre dos millones de regresiones, partiendo de 58 posibles regresores de los que selecciona 22 como significativos para explicar la tasa

⁸ Ver Leamer (1983).

de crecimiento. Más recientemente, Hendry y Krolzig (2004) realizan la misma tarea⁹, pero con una sola regresión, utilizando su nuevo software, PcGets¹⁰. El común denominador de estos trabajos es que ante las dificultades de la teoría de sugerir un modelo “verdadero”, buscan las variables que resultan explicativas utilizando procedimientos *automáticos*¹¹. No importan en nuestro caso las características de las técnicas que emplean, sino el hecho de que están aceptando algo que se parece mucho al procedimiento que hemos definido como “*data mining*”.

Para terminar este punto quiero hacer dos reflexiones. La primera es que este cambio en la valoración del *data mining*, que surge de los trabajos comentados, refleja un relajamiento de la convicción de que la teoría debe ser la guía exclusiva de la tarea del economista y que la comprobación o falsificación de esas teorías por vía de la Econometría u otras técnicas cuantitativas deben estar férreamente subordinadas a ella. La corriente actual, reflejada en los métodos de la LSE y del VAR, es más proclive a “*dejar hablar a los datos*”¹².

En segundo lugar creo que los medios técnicos con que se cuenta en la actualidad permiten que todo trabajo que utilice técnicas cuantitativas esté acompañado por los datos utilizados en las estimaciones, los que deberían estar disponibles en algún lugar accesible a través de Internet. De esta forma todos podríamos repetir las estimaciones y recorrer el camino que previamente ha transitado el autor. La actitud parece ser cada vez más tolerante, pero a condición de que todas las cartas estén sobre la mesa, es decir que si hacemos “*data mining*”, debemos hacerlo en forma explícita y transmitir toda la información disponible.

⁹ Agradezco a la Lic. Lorena Garegnani sus sugerencias respecto a este punto.

¹⁰ Para una descripción del software conocido como PcGets, ver Hendry y Krolzig (2005).

¹¹ Un análisis reciente y muy detallado sobre este tema puede verse en Phillips (2005).

¹² Ver Dow (2002).

4. LA METODOLOGÍA DE FRIEDMAN DESPUÉS DE CINCUENTA AÑOS

Hace cincuenta años Friedman (1953) desarrollaba, en su ya clásico trabajo “*The Methodology of Positive Economics*”, una idea sencilla, pero de una gran fuerza, que motivó discusiones entre los economistas por medio siglo: *las teorías valen por su poder predictivo*, y por lo tanto las teorías no deben ser evaluadas por sus supuestos. Estos pueden ser falsos, pero lo que importa es que la teoría sirva para predecir adecuadamente. La teoría de la gravedad de Galileo puede basarse en supuestos falsos: no existe el vacío perfecto, la tierra no es plana, etc., pero esas son circunstancias no relevantes. Lo que realmente interesa es que la teoría predice adecuadamente¹³.

Denomina “economía positiva” a aquella que procura interpretar el funcionamiento del sistema económico, a diferencia de la “economía normativa” y del “arte”, que procura la obtención de determinados resultados¹⁴.

“Viewed as a body of substantive hypotheses, theory is to be judged by its predictive power for the class of phenomena which it is intended to explain. Only factual evidence can show whether it is right or wrong, or better, tentatively accepted as valid or rejected ...the only relevant test of the validity of a hypothesis is the comparison of its prediction with experience...”

¿Por qué es importante para nosotros la teoría de Friedman? Porque sugiere el uso de técnicas econométricas y estadísticas para comprobar las teorías. Recientemente existe un renovado interés en su forma de analizar la

¹³ Sin embargo debe destacarse que supone que al confrontarse la teoría con sus predicciones se realiza un test implícito de la veracidad de los supuestos. Una cosa es que sean disparatados y otra que no se cumplan en determinadas circunstancias, como el vacío perfecto en el ejemplo referido.

¹⁴ Aquí sigue la clasificación que John Neville Keynes presenta en su “Scope and Methods of Political Economy” en 1891. (Friedman, 1953).

metodología, que fue fuertemente criticada desde perspectivas distintas. Hayek se lamentaba de no haber escrito más contra el enfoque de Friedman que contra el de Keynes. Los filósofos de la ciencia consideran su enfoque una manifestación del instrumentalismo y del positivismo. Por otra parte su ideología extremadamente liberal y antikeynesiana, expresada en otras áreas, contribuyó a que muchos se unieran en una actitud crítica que a veces se expresó en forma muy ácida.

Sin embargo pareciera existir en la actualidad una reconsideración de su pensamiento, y no podemos dejar de volver a recordarla, así como algunas opiniones en torno a estas ideas que han aparecido recientemente. Es importante notar que en su trabajo no cita ningún metodólogo, excepto el trabajo del padre de J.M. Keynes, sino que todas las citas son referencias de trabajos de economistas. Pero pareciera adoptar una posición compatible con Popper, cuando afirma *“it is accepted if its prediction are not contradicted and great confidence is attached to it if it has survived many opportunities for contradiction”*¹⁵. Sin embargo también destaca que las opiniones y el “background” del economista tienen también importante influencia.

Analizada su posición desde la filosofía de la ciencia, el trabajo de Friedman ofrece una confusión bastante importante de enfoques alternativos. Maki (2003) concluye que todos los enfoques pueden ser encontrados conviviendo en su trabajo, razón por la cual hay tantos amigos y enemigos que han encontrado material para hacer las evaluaciones más dispares. Esto ocurre, tal vez, porque el trabajo no está mirando la filosofía de la ciencia sino la economía y el trabajo de los economistas, desde la propia economía.

Boland (2003) sostiene que el trabajo de Friedman, a pesar de usar la expresión “economía positiva” en el título no está inspirado en el positivismo, sino en el instrumentalismo, tal como ya lo había planteado en su trabajo de

¹⁵ Al menos la teoría seleccionada será aquella menos contradicha por su poder predictivo y no interesa que los presupuestos en que se basa la teoría sean o no ciertos.

1979 publicado en la JEL¹⁶, y atribuye también el rechazo que ha generado su trabajo a las ideas de Friedman en otros campos. Cuenta que en una reunión de economistas keynesianos realizada en 1983 en Cambridge, en conmemoración de los cien años del nacimiento de Keynes, presentó las ideas de Friedman sin decir a quien correspondían y obtuvo un enorme respaldo de los que allí estaban. Luego preguntó quien estaba de acuerdo con la metodología de Friedman y casi la totalidad manifestó su oposición. Por eso propone un nuevo análisis del referido trabajo de Friedman desprovisto de contenido político, aunque destaca que la forma de trabajar en los departamentos de economía de las principales universidades, orientada a la formulación de modelos matemáticos hace que los partidarios del enfoque de Friedman sean minoritarios.

Según Reder (2003) su teoría estaba destinada a defender la teoría neoclásica de los precios, que estaba siendo atacada por la teoría de la competencia imperfecta. Friedman decía que los modelos nunca describen adecuadamente los hechos, pero demostrar que la existencia de mercados funcionaba en forma bastante similar a la competencia perfecta, confiere a la teoría neoclásica mejor poder predictivo. Reder sostiene que la metodología de Friedman se refiere básicamente a pronósticos de tipo cuantitativo, y que el desarrollo explosivo de la Econometría, de la computación y de la disponibilidad de datos abren nuevas perspectivas para esta metodología, las que permiten inclusive cuestionar teorías como el principio de eficiencia en los mercados financieros, que hacen al corazón de la propia teoría neoclásica. Esta última cuestión está relacionada con la famosa expresión de Samuelson sobre lo que denomina el “F-Twist”. Esto viene de la reunión anual de la American Economic Association de 1962, donde se dedica un panel a los problemas metodológicos de la economía. Allí Ernest Nagel (1963) rescata el trabajo de Friedman:

¹⁶ Ver Boland (1979). En este trabajo cambia la posición que había mantenido en trabajos anteriores y sostiene que lo que Friedman propone no tiene nada que ver con el positivismo (a pesar de que la palabra “positiva” aparece en el título, sino que debe interpretarse como una manifestación del instrumentalismo, denostado tanto por los positivistas como por Popper.

“Professor Friedman’s essay... is marked by an ambiguity that perhaps reflects unresolved tensions in his views on the status of economic theory. Is he defending the legitimacy of unrealistic theoretical assumptions because he thinks theories are at best only useful instruments, valuable for predicting observable events but not to be viewed as genuine statements whose truth or falsity may be significant investigated? But if this is the way he conceives theories (and much in his arguments suggest that it is), the distinction between realistic and unrealistic theoretical assumptions is at best irrelevant...”

Por otra parte, Samuelson describe lo que denomina el “F-Twist”, que consiste en la elaboración de una formulación de la metodología concordante con la finalidad de dar fuerza a sus propias ideas. Señala, como ejemplo, que Friedman procura de esta forma presentar como verdadero el modelo de competencia perfecta, y dejar de esta forma de lado la teoría de la competencia monopólica. Si lo que Reder sostiene es cierto, entonces esa crítica quedaría a un lado, y Boland tiene razón cuando dice que debemos desvincular la idea que sostiene Friedman con el resto de su pensamiento y de sus ideas controvertidas en otros campos¹⁷. Después de transcurridos cincuenta años desde su aparición, el trabajo metodológico de Friedman vuelve a renacer como tema de discusión y promete interesantes cambios de opinión respecto a las ideas expuestas en él en el futuro.

5. CALIBRACIÓN

Si bien las técnicas conocidas como *calibración* no son nuevas, ya que el trabajo original de Kydland y Prescott (1982) fue publicado hace 23 años y no se trata de una técnica econométrica en su alcance restringido, es conveniente

¹⁷ Debe tenerse en cuenta que en algunos casos una teoría con excelente poder explicativo puede no pronosticar correctamente, como es el caso de la teoría de la evolución, que puede explicar el proceso evolutivo pero no las direcciones futuras. Lo mismo ocurre con la geología, que puede explicar los terremotos pero no predecirlos. Sin embargo, es posible que un desarrollo posterior permita también predecir adecuadamente los fenómenos a los se refiere.

que nos refiramos a ella por la actualidad que ha cobrado al ser la metodología utilizada por sus autores para “*testear*” la teoría del ciclo real (RBC), que es una de las dos contribuciones que los hicieron acreedores del Premio Nobel de Economía del año pasado. Vale la pena entonces que hagamos algunas reflexiones sobre este procedimiento.

Como es sabido, consiste en tomar los coeficientes de otros estudios o análisis cuantitativos, sin estimarlos mediante las técnicas econométricas habituales, lo que se justifica por la complejidad de los modelos sofisticados de equilibrio general con que se expresa la teoría del RBC.

Sin embargo sería útil comparar los resultados obtenidos mediante las técnicas de calibración con otras formas de estimación alternativas, mediante el uso de estadísticos como el RMSE¹⁸ (*Root Mean Square Error*). Seguramente vamos a ver en el futuro cómo otras técnicas alternativas se aplican a este problema, especialmente las relacionadas con modelos VAR y también veremos cómo la calibración es exportada a otras áreas. Seguramente el debate en torno a esta técnica va a cobrar un nuevo impulso en el futuro.

6. TEORÍA ECONÓMICA Y ECONOMETRÍA

Como hemos sostenido anteriormente¹⁹, cuando comienza a consolidarse la Econometría es a partir del trabajo de Haavelmo (1944), y a partir de ese momento y durante los próximos treinta años, es decir hasta 1974 aproximadamente, reina un paradigma consistente en la primacía de la teoría económica y la subordinación de la Econometría a sus dictados. Los modelos son abstracciones de la realidad, generalmente suponiendo equilibrio en el largo plazo y parsimoniosos (compuestos por pocas variables) y tienen adicionado un término aleatorio. Luego de recogidos los datos estadísticos restaba realizar las estimaciones econométricas, que sólo podían confirmar la

¹⁸ Esta es la opinión de Fair (1992) y de Hansen y Heckman (1996). En este trabajo los autores citan, refiriéndose a los antecedentes de la calibración, un trabajo de Mantel (1974).

¹⁹ Ver Navarro (1997).

teoría. Una vez realizadas se realizaban sólo algunos tests, fundamentalmente para verificar que no hubiera “patologías”, tales como la autocorrelación, la heterocedasticidad o la colinealidad. En caso de descubrirse alguna de ellas se aplicaba el método de los mínimos cuadrados generalizados u otras técnicas correctivas para depurar el modelo de la supuesta “*enfermedad*”.

La crisis de los años setenta conduce a probar otras metodologías, como señalamos en la Introducción, lo que lleva a la aparición de los nuevos métodos precitados (los diseñados por Leamer, Sims y Sargan y Hendry).

El paulatino abandono de la metodología original, nacida básicamente en el seno de la Cowles Commission, para desplazarse hacia la de Sargan, Hendry y Spanos, conocida como LSE (London School of Economics) es analizado por Cook (2003 y 2005) y por Davis (2005). El primero presenta el cambio que se produce desde la metodología descripta más arriba hacia el enfoque de la LSE, como un cambio de paradigma en el sentido de Kuhn (1970), cuya más distintiva característica metodológica es la introducción del concepto de Proceso Generador de Datos (DGP), que sienta las bases para el método de la LSE y que consiste en un mecanismo que motiva los datos bajo análisis y que es “complicado, interdependiente, dinámico, estocástico, multidimensional y no lineal, por lo que debe ser modelado teniendo en cuenta esas características y por lo tanto la teoría económica no puede darnos un modelo “verdadero” para ser estimado. Por lo tanto la modelización econométrica debe ser vista como una serie de reducciones del DGP, con una variedad de tests, tratando de no perder nada de la información existente en los datos. Las demás herramientas incorporadas por enfoque LSE tienen su fundamento en el análisis de los datos. El mecanismo de corrección de errores (ECM), la comparación con otros modelos alternativos (encompassing) y la actitud ante la autocorrelación de los errores es consecuencia de ese cambio de enfoque, o de un cambio de paradigma, como dice Cook en el citado trabajo. Davis, alternativamente presenta este cambio desde la perspectiva de la concepción semántica de las teorías. Pero más allá de la pertinencia del modelo de Kuhn para el análisis del cambio operado es de rescatar que lo que distingue el cambio es la actitud de “*dejar hablar a los datos*” y quitar a la teoría el papel casi tiránico que tenía en el enfoque original, al que denomina “*textbook approach*”.

7. ALGUNAS REFLEXIONES FINALES

Durante los últimos años se han publicado infinidad de libros y artículos sobre la relación entre metodología y la Econometría. Pero, con las limitaciones de la subjetividad, si tuviera que elegir el más representativo de la década pasada, me inclinaría por “*The Limits of Econometrics*”, de Darnell y Evans (1990)²⁰. De los aparecidos recientemente creo que vale la pena analizar cuidadosamente la obra de C. Favero (2001), “*Applied Macroeconometrics*”. De la lectura de ambos es fácil apreciar que el eje de la discusión ha cambiado: mientras en el primero se discuten los fundamentos filosóficos de la Econometría, en el segundo se analiza la relación entre metodología y Econometría desde una perspectiva más aplicada, analizando, en ambos casos las características de los distintos métodos que están siendo usados con mayor frecuencia: los modelos inspirados en la labor de la Cowles Commission, en la metodología de la London School of Economics (LSE), los modelos de Vectores Autorregresivos y los métodos de calibración relacionados con la teoría del ciclo real.

También quiero reiterar que estas cuestiones deben ser tenidas presentes no sólo por los que practicamos Econometría. Quienes no la practican seguramente leen muy frecuentemente trabajos que utilizan estas técnicas, y es preciso que sepan cuáles son los alcances y las limitaciones de las conclusiones que esos trabajos les están transmitiendo. Por eso creo que este tema debe interesar a todos los economistas, de todas las ramas y de todas las especialidades.

¿Cuál es el sentido en que está evolucionando la ciencia económica y cuál es el futuro de la Econometría como herramienta para proponer y verificar las teorías? En mi opinión va a tener un rol cada vez más relevante, sobre todo a medida en que se van perfeccionando las técnicas utilizadas y vamos disponiendo de más y mejores datos estadísticos. Dow (2002) diferencia dos

²⁰ Ver Navarro (1997), donde se puede encontrar reflexiones sobre esta cuestión y un análisis de la situación a esa fecha.

formas de trabajar en Economía: a la primera la llama el método Cartesiano/Euclidiano, que parte de axiomas evidentes y deduce de ellos sus conclusiones que no lo son. La otra es la que llama “babilónica”, que tiene como base la realidad tal como la observamos, sin atarse a esos principios axiomáticos, más propios de la matemática, que no se preocupa de la realidad. Hemos vivido un período en el que ha reinado el primer método, y predice para los próximos diez años un giro hacia la segunda forma de análisis, es decir que prevé una orientación menos matemática y más preocupada por el análisis de la realidad.

Por último, debemos también preguntarnos acerca de las metodologías más adecuadas para analizar la economía de nuestro país, donde los cambios en los regímenes han sido tan abruptos y la presencia de “outliers” tan frecuente, que hacen difícil trasladar sin elaboración los modelos econométricos elaborados para países desarrollados, que son mucho más estables y donde los cambios son mucho más suaves.

REFERENCIAS

- Boland L. (1979) “A critic of Friedman’s critics”. *Journal of Economic Literature*. 17:503-22.
- Boland L. (2003) “Methodological criticism vs. ideology and hypocrisy”. *Journal of Economic Methodology*. 10:4. 521-526.
- Cook, S. (2003) “A Kuhnian perspective of econometric methodology”. *Journal of Economic Methodology*. 10:1. 59-78.
- Cook, S. (2005) “On the semantic approach to econometric methodology”. *Journal of Economic Methodology*. 12:1. 117-124.
- Darnell, A. (1994) “*A Dictionary of Econometrics*”. Aldershot: Edward Elgar Pub.
- Darnell, A. y Evans, L. (1990) “*The Limits of Econometrics*”. Aldershot: Edward Elgar Pub.
- Davis G. (2005). “Clarifyng the puzzle between the Textbook and LSE approaches to econometrics: A comment on Cook’s Kuhnian perspective

- on econometric modelling”. *Journal of Economic Methodology*. 12:1. 93-116.
- Dow, S. (2002) “*Economic methodology: an inquiry*”. Oxford: Oxford University Press.
- Fair, R. (1994). “*Testing Macroeconomic Models*”. Cambridge: Harvard University Press.
- Friedman, M. (1953) “*The Methodology of Positive Economics*”, en *Essays in Positive Economics*, Chicago: The University of Chicago Press, 3-46.
- Hansen, L. y Heckman, J. (1996) “*Journal of Economic Perspectives*”. 10:1. 87-104.
- Hendry, D. y Krolzig, H. (2004) “We Ran One Regresión”. Department of Economics. Oxford University.
- Hendry, D. y Krolzig, H. (2005) “The Properties of Automatic PcGets modelling”. *Economic Journal*. Vol. 115 No. 502. C32-C61.
- Hoover, K. y Perez S. (2000) “Three attitudes toward data mining”. *Journal of Economic Methodology* 7:2 195-210.
- Keynes, J. M. (1972) [1924]. “Alfred Marshall”, en “*Essays on biography*”. Cambridge: Cambridge University Press. 161-231.
- Klamer, A. (2001) “Making sense of economists: from falsification to rhetoric and beyond”. *Journal of Economic Methodology*. 8:1, 69-75.
- Kuhn, T. S. (1970) “*The Structure of Scientific Revolutions*”. 2da. Edición. Chicago: Chicago University Press.
- Kydland, F. y Prescott, E. (1982). “Time to Build and Aggregate Fluctuations”. *Econometrica*. 50:6. 1345-1370.
- Leamer, E. (1983) “Let’s Take the Con Out of Econometrics”. *American Economic Review* 73:31-43.
- Levine, R. y Renelt, D. (1992) “A Sensitivity Analysis of Cross-Country Growth Regressions”. *American Economic Review*. 82-4, 942-963.
- Lucas R. (1976) “Econometric Policy Evaluation: A Critique” *Journal of Monetary Economics*. 2:19-46.
- Machlup, F. (1963) *Problems of Methodology: Introduction*. American Economic Review. Papers and Proceedings. Vol. LII. Number 2. 231-237.

- Maki, U. (2003) "The methodology of positive economics (1953) does not give us the methodology of positive economics". *Journal of Economic Methodology*. 10:4. 495-505.
- Mantel, R. (1974). "On the Characterization of Aggregate Excess Demand". *Journal of Economic Theory* 7, 348-55.
- McCloskey, D. (2000) "*How to be Human, though an Economist*". Ann Arbor: University of Michigan Press.
- McCloskey, D. (2002) "*The Secret Sins of Economics*". Chicago: Prickly Paradigm Press.
- Nagel, E. (1963) Assumptions in Economic Theory. *American Economic Review. Papers and Proceedings*. Vol. LII. Number 2. 211-219.
- Navarro A. (1997) "*Reflexiones sobre la relación entre Economía, Econometría y Epistemología*". Anales de la Academia Nacional de Ciencias Económicas.
- Navarro, A. (2003) "El pensamiento de Mark Blaug en cuanto a la comprobación empírica de las teorías económicas". En "*Temas de Macroeconomía*". Instituto de Economía Aplicada Dr. Vicente Vázquez-Presedo. Academia Nacional de Ciencias Económicas.
- Phillips, P. (2005) "Automated Inference and the Future of Econometrics". *Econometric Theory*. 21, 2005.
- Reder, M. (2003) "Remarks on the methodology of positive economics". *Journal of Economic Methodology*. 10:4, 527-530.
- Sala-i-Martin, X. (1997) "I Just Ran Two Million Regression". *American Economic Review. Papers and Proceedings*. Vol. 87-2. 178-183.
- Samuelson, P. (1963). "Problems of Methodology: Discussion". *American Economic Review. Papers and Proceedings*. Vol. LII. Number 2. 231-237.
- Sims, C. (1980) "Econometric and Reality". *Econometrica*. 48:1 3-48.
- Sims, C. (1996). "Macroeconomics and Methodology". *Journal of Economic Perspectives*. 10:1 105-120.
- Smith, R. (1999) "Unit roots and all that: the impact of time series methods in macroeconomics". *Journal of Methodology Economics*. 6:2 238-258.
- Skidelsky, R. (1983) "*John Maynard Keynes. Hopes Betrayed 1883-1920*". New York: Penguin Group.

PRONÓSTICO CON MODELOS ECONOMÉTRICOS

HILDEGART A. AHUMADA
UNIVERSIDAD TORCUATO DI TELLA

*“When weather forecasts go awry,
meteorologists get a new supercomputer
When economist mis-forecast,
we get our budgets cut”*

(Hendry, 2001)

Tradicionalmente se identifican tres objetivos de la modelación económica: descripción de relaciones económicas, pronóstico y evaluación de políticas. Recientemente el segundo objetivo, el uso de modelos econométricos para pronóstico¹, ha sido profundamente estudiado (especialmente en Clements and Hendry, 1998, 2001) alcanzando conclusiones por un lado sorprendentes pero que a la vez explican la “performance” de los pronósticos económicos. Asimismo, estos resultados dan validez formal a algunas prácticas usuales en economía aplicada. Aunque la falla predictiva puede extenderse a otras ciencias (el ejemplo más cotidiano es el de la meteorología), en el caso de la economía ha sido muchas veces excesivamente penalizado por lo cual aparece como más necesario tratar de entender sus causas y las diferentes alternativas disponibles para su corrección.

Esta nota resume en la sección primera las principales cuestiones referidas a pronóstico con modelos econométricos, en la sección segunda presenta formalmente los resultados más destacados y en la sección tercera ejemplifica

¹ A lo largo de esta nota no distinguiremos “pronóstico” de “predicción”, una discusión de estos conceptos puede verse en la literatura indicada .

los mismos tomando como base la relación entre importaciones y producto de Argentina.²

1. UN PRIMER ANÁLISIS

La primera cuestión que conviene analizar es ¿de qué depende el “éxito” de los pronósticos económicos? En sentido amplio un pronóstico depende de tres aspectos:

- La *existencia de regularidades* en el pasado
- Que dichas regularidades sean *captadas por el método de pronóstico*
- Que ellas sean *informativas sobre el futuro*

Pero en el caso de los pronósticos económicos debe tenerse en cuenta que las economías son inherentemente evolutivas y caracterizadas como *no estacionarias*. Específicamente, las series económicas pueden ser no estacionarias porque están sujetas a:

- i) pequeños y recurrentes shocks que se acumulan como en el caso de series integradas de primer orden, $I(1)$, y/o
- ii) intermitentes, probablemente grandes y no anticipados shocks originados en cambios estructurales y de régimen

La econometría evolucionó con respecto a i) a través del estudio de las relaciones de *cointegración* (relaciones de largo plazo entre variables integradas) y la modelación en términos de *EC*, “Error Correction” o “Equilibrium Correction Models” (re-parametrizando las relaciones entre variables integradas en forma de diferencias y desvíos de las relaciones de cointegración).³

² La sección segunda puede omitirse sin pérdida de continuidad.

³ Debe preferirse la denominación “Equilibrium correction” a “Error correction” ya que este tipo de modelo si bien puede funcionar apropiadamente cuando se estima dentro de la

A raíz de ii) la econometría ha tendido a repensar el objetivo de pronóstico. Históricamente los pronósticos con modelos econométricos (por ejemplo en la derivación de los ampliamente conocidos estadísticos de Chow, 1960) han supuesto no solo que dichos modelos son buenas representaciones de la realidad sino que “la realidad” (o más precisamente los “parámetros del modelo”) es (son) relativamente invariable (s)⁴. Este segundo supuesto es muy difícil de validar empíricamente y el punto central es que, en general, los resultados sobre pronósticos basados en un mundo de parámetros constantes se modifican sustancialmente. A continuación se presentan algunos de estos nuevos resultados, tal vez los más inquietantes.

Cuando se levanta el supuesto de parámetros constantes puede darse que

- un modelo que incluya variables causales (con sentido económico) no es necesariamente mejor en cuanto a pronóstico que uno que incluya *variables no causales*.
- *un modelo mal especificado* puede pronosticar mejor que uno que esté más próximo al proceso generador de los datos (PGD) .

Intuitivamente estos resultados se explican porque para pronóstico no interesa tanto que el modelo “ajuste bien” dentro de la muestra sino que tenga una *rápida adaptación al cambio*, una vez que el shock no anticipado tuvo lugar de forma de evitar errores sistemáticos de pronósticos.

En este contexto encontramos que, en lo referente a pronóstico, los modelos econométricos son dominados por modelos extremadamente simples de series temporales incluso solo en diferencias o doble diferencias. La explicación de esta mejor performance debe buscarse en el tipo de componente de los modelos econométricos que más afectan a los pronósticos económicos. Los componentes de un modelo econométrico pueden agruparse en:

muestra, tiende a “equivocarse” (no corregir errores) cuando se lo utiliza para pronóstico y ocurren cambios estructurales en dicho período de pronóstico.

⁴ Ver Anexo 1

- el efecto de las variables aleatorias observables
- errores no observables
- términos determinísticos (constante, tendencia, estacionales)

De ellos, el comportamiento de los componentes “determinísticos” resulta ser el principal responsable de las fallas en pronóstico (aunque, probablemente, interactuando con la falta de una adecuada especificación del modelo).

Por lo tanto, se obtienen mejores resultados en cuanto a pronóstico con un modelo en diferencias (porque “eliminan” las constantes o medias). También explican la utilidad de la corrección de ordenada en los modelos econométricos (porque “ajustan” las medias). A continuación se presentan formalmente estos resultados y luego se ejemplifican para el caso de las importaciones de Argentina.

2. PRESENTACIÓN DE ALGUNOS RESULTADOS

El efecto de variables causales sobre pronósticos puede apreciarse para el siguiente PGD. Supongamos que x_t es una variable genuinamente “causal” de y_t tal que

$$y_t = \varphi x_{t-1} + \varepsilon_{1t} \quad (1a)$$

$$x_t = x_{t-1} + \varepsilon_{2t} \quad (1b)$$

Entonces, esperaríamos que todo pronóstico que incluya a “ x ” sea superior al que no lo incluya. Pero este tipo de razonamiento sería válido si el PGD de y_t permaneciese constante. Si no lo es, por ejemplo, porque hay un cambio como el siguiente a partir del período T ,

$$y_{T+i} = \varphi^* x_{T+i-1} + \varepsilon_{1T+i} \quad (2a)$$

$$x_{T+i} = x_{T+i-1} + \varepsilon_{2T+i} \quad (2b)$$

los errores de pronóstico después del “break” con este modelo (condicional a conocer x_{T+i-1}) tendrán un valor esperado,

$$E [y_{T+i} - y_{T+i}^f \mid x_{T+i-1}] = (\varphi^* - \varphi) x_{T+i-1} \quad (3)$$

Los errores tenderán a ser persistentes y grandes al ser x_t integrada de orden 1. En cambio, una simple regla no causal tal como la de repetir el valor anterior (una simple extrapolación, e)

$$y_{T+i}^e = y_{T+i-1} \quad (4)$$

puede ser superior en cuanto a pronóstico como puede verse a partir de T+1

$$\begin{aligned} E [y_{T+i} - y_{T+i}^f \mid x_{T+i-1}] &= E [\Delta y_{T+i} \mid x_{T+i-1}] \\ &= (\varphi^* - \varphi) x_T \quad \text{si } i = 1 \\ &= 0 \quad \text{si } i > 1 \end{aligned} \quad (5)$$

En consecuencia, un pronóstico con un modelo no causal pueden ser mejor que uno que incluya variables causales.

Para apreciar cómo opera la corrección de ordenada o del error de pronóstico usando el error anterior⁵ veamos el siguiente simple caso de un pronóstico a través de un modelo autoregresivo

$$y_{T+1}^f = b y_{T-1} \quad (6)$$

el error de pronóstico (f) es (suponiendo que β es el parámetro del PGD),

$$f_{T+1} = y_{T+1} - y_{T+1}^f = (\beta - b) y_T + \varepsilon_{T+1} \quad (7)$$

mientras que si se corrige el pronóstico con el error anterior,

$$y_{T+1}^{fc} = b y_T + f_T \quad (8)$$

⁵ Este efecto es similar al realizado a través una variable dummy. Recordemos que si la variable dummy tiene un 1 para un solo período, el residuo es nulo para dicha observación.

siendo el resultante error de pronóstico,

$$\begin{aligned} f_{T+1}^c &= y_{T+1} - y_{T+1}^{fc} = (\beta - b) y_T + \varepsilon_{T+1} - f_T \\ f_{T+1}^c &= f_{T+1} - f_T = \Delta f_{T+1} \end{aligned} \quad (9)$$

El error de pronóstico corregido es la diferencia de pronósticos consecutivos. Su beneficio ($f_{T+1}^c < f_{T+1}$) se origina cuando los “breaks” dan lugar a errores de pronósticos positivamente correlacionados, lo cual ocurre frecuentemente.

Finalmente, para ver analíticamente las principales causas de fallas de pronóstico se puede considerar el siguiente sistema de n variables que ya ha sido reparametrizado en términos de variables $I(0)$, transformando las variables originalmente $I(1)$ a través de relaciones de cointegración y primeras diferencias.

$$y_t = \delta + \pi y_{t-1} + \varepsilon_t \quad \varepsilon_t \sim IN_n(0, \Omega) \quad t = 1, \dots, T \quad (10)$$

$$E[y_t] = [I - \pi]^{-1} \delta = \gamma \quad \text{ó} \quad \delta = [I - \pi] \gamma \quad (11)$$

Entonces,

$$(y_t - \gamma) = \pi (y_{t-1} - \gamma) + \varepsilon_t \quad (12)$$

Y si ocurre un cambio estructural a partir de $T+1$ (pero que mantiene las mismas propiedades de integración y cointegración)

$$(y_{T+1} - \gamma^*) = \pi^* (y_T - \gamma^*) + \varepsilon_{T+1} \quad (13)$$

por lo que los errores de pronóstico serán

$$\begin{aligned} f_{T+1} &= y_{T+1} - y_{T+1}^f \\ f_{T+1} &= \delta^* + \pi^* y_T + \varepsilon_{T+1} - \delta^e + \pi^e y_T \end{aligned} \quad (14)$$

donde “e” indica los estimadores de los respectivos parámetros. Los valores esperados condicionales serán,

$$\begin{aligned} E[f_{T+1} | y_T] &= \delta^* + \pi^* y_T + E[\varepsilon_{T+1}] - E[\delta^e] + E[\pi^e] y_T \\ E[f_{T+1} | y_T] &\approx (\delta^* - \delta^p) + (\pi^* - \pi) y_T \end{aligned} \quad (15)$$

donde

$$\begin{aligned} E[\pi^e] &= \pi^p \\ E[\delta^e] &= \delta^p = [I - \pi^p] \gamma \end{aligned}$$

y p denota al respectivo parámetro poblacional que no es necesariamente π ni δ , los parámetros del PGD. Por lo tanto sus estimadores pueden ser inconsistentes. Sin embargo, $\gamma^p = \gamma$ ya que casi todos los métodos que estiman la media obtienen como valor esperado al poblacional γ .

Resulta así el pronóstico no condicional

$$\begin{aligned} E[f_{T+1}] &\approx [(I - \pi^*) \gamma^* - (I - \pi^p) \gamma] + (\pi^* - \pi^p) \gamma \\ E[f_{T+1}] &\approx (I - \pi^*) [\gamma^* - \gamma] \end{aligned} \quad (16)$$

Por lo que el valor esperado del error de pronóstico será nulo cuando

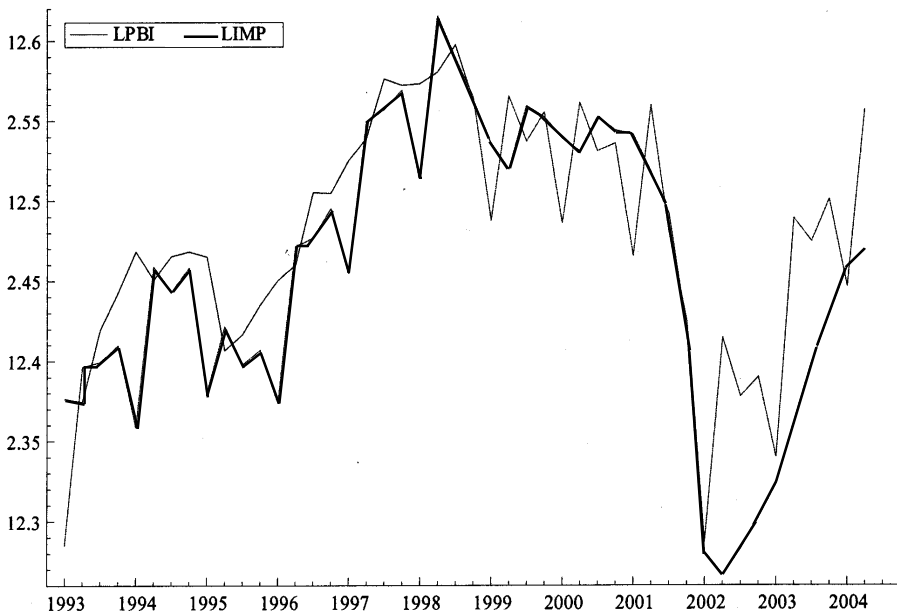
- i) un proceso tenga media nula ($\delta = \gamma = 0$)
- ii) la media de largo plazo sea constante ($\gamma^* = \gamma$)

El caso ii) puede darse si los cambios de π y δ se compensan de forma tal que γ no cambie. Debe notarse que un cambio en γ no será muy notado en la medida en que π se acerque a la identidad (un modelo en diferencias de las desviaciones con respecto a la media). Sin embargo, el punto central de esta formulación es la importancia de los cambios determinísticos, la media de largo plazo, γ , con respecto a los cambios en dinámica (π) y a problemas de especificación que deriven en parámetros inconsistentes.

3. EJEMPLOS CON EL CASO DE LAS IMPORTACIONES DE ARGENTINA

El siguiente gráfico muestra la relación entre importaciones y PBI ajustada por medias y rangos que se usará para ilustrar los resultados anteriores. Como puede verse a partir de 2002 la relación muestra un “quiebre” que hace sospechar que los pronósticos basados en dicha relación sobre la base de datos hasta 2001 experimentarían errores sistemáticos. Vale la pena enfatizar que la elección de este caso responde al objetivo de exponer en forma sencilla los resultados encontrados formalmente y no de sugerir un modelo econométrico para las importaciones de Argentina, dando por descontado que presenta problemas de especificación tal como el de "variables omitidas".

Gráfico 1. Importaciones y PBI (series originales)



Tales sospechas son confirmadas en la Tabla 1⁶ donde se presentan los pronósticos efectuados con un modelo de "Equilibrium correction" (*EqC*) para las importaciones estimado con la muestra 93(3)-2001(4) y que incluye los desvíos de la relación de largo plazo con el PBI ($LIMP = 24.8 + 2.8 LPBI$). En todo el período 2002 (1)-2004 (2) hay errores de pronósticos sistemáticos y en valor absoluto mayores a 4 veces su SE.

Tabla 1. EqC. estimado con datos 93(3)-2001(4)

Analysis of 1-step forecasts EqC					
Date	Actual	Forecast	Y-Yhat	Forecast SE	t-value
2002 1	9.59400	9.83483	-0.240828	0.0365336	-6.59196
2002 2	9.54883	9.83131	-0.282485	0.0705898	-4.00178
2002 3	9.60124	10.0484	-0.447199	0.0825816	-5.41523
2002 4	9.66524	10.0429	-0.377615	0.0788820	-4.78709
2003 1	9.73359	10.1205	-0.386881	0.0740701	-5.22317
2003 2	9.85795	10.2537	-0.395756	0.0707354	-5.59487
2003 3	9.98087	10.3828	-0.401981	0.0751609	-5.34826
2003 4	10.0821	10.3760	-0.293949	0.0652890	-4.50227
2004 1	10.1770	10.4290	-0.251994	0.0599102	-4.20620
2004 2	10.2117	10.4477	-0.236007	0.0487511	-4.84106

Con fines de comparación se realizaron los pronósticos para el mismo período con un modelo "naive" equivalente al de primeras diferencias, un modelo univariante de series temporales AR(2). Como puede observarse en la Tabla 2, después del salto en 2002 (1) los errores tienden a disminuir tanto en forma absoluta como en relación con su SE.

⁶ En las siguientes tablas se presentan en las columnas sucesivamente las fechas de los pronósticos, el valor observado de la variable y el pronosticado, el error de pronóstico, el error típico (SE) de dicho error de pronóstico y el cociente de cada error relativo a su SE expresado como "t" (ver Hendry and Doornik, 1996).

Tabla 2. Modelo naive: AR(2) en niveles

Analysis of 1-step forecasts Dif /AR2

Date	Actual	Forecast	Y-Yhat	Forecast SE	t-value
2002 1	9.59400	9.91462	-0.320616	0.0781873	-4.10061
2002 2	9.54883	9.31213	0.236701	0.110383	2.14436
2002 3	9.60124	9.68633	-0.0850916	0.0762294	-1.11626
2002 4	9.66524	9.62862	0.0366211	0.0770136	0.475515
2003 1	9.73359	9.78660	-0.0530117	0.0758947	-0.698490
2003 2	9.85795	9.79902	0.0589313	0.0740629	0.795692
2003 3	9.98087	10.0832	-0.102345	0.0729037	-1.40384
2003 4	10.0821	10.0201	0.0620407	0.0678037	.915005
2004 1	10.1770	10.1884	-0.0113600	0.0684382	-0.165988
2004 2	10.2117	10.2173	-0.00567140	0.0669084	-0.0847636

Finalmente también se calcularon los pronósticos para el período 2003(1)-2004(2) (después del cambio estructural) utilizando el mismo modelo de “Equilibrium Correction” inicial pero realizando una corrección de ordenada a través de una dummy con valores no nulos a partir de 2002 (1). Los resultados se presentan en la Tabla 3 donde se puede apreciar la mejor performance resultante, con relación a los de la Tabla 1 y para los tres primeros pronósticos de 2003 con respecto a la Tabla 2.⁷

En consecuencia un simple modelo univariante pueden ser mejor en cuanto a capacidad predictiva de las importaciones que un modelo mejor especificado que incluya una variable indudablemente causal como es el PBI para el presente caso, si se ha producido algún tipo de cambio estructural que afecte a la relación entre tales variables. Solo si el modelo se “enmienda” en la ordenada a través

⁷ Nótese que, posiblemente, para los últimos tres períodos del pronóstico (2003(4)- 2004 (2)) una re-estimación del coeficiente de la dummy sería necesaria.

Tabla 3. EqC de Tabla 1 con corrección de ordenada
(dummy de cambio de ordenada en 2002)

Analysis of 1-step forecasts con dummy crisis					
Date	Actual	Forecast	Y-Yhat	Forecast SE	t-value
2003 1	9.73359	9.73091	0.00268131	0.0372786	0.071926
2003 2	9.85795	9.88006	-0.0221127	0.0386498	-0.572129
2003 3	9.98087	9.97889	0.001973	0.0392141	0.050318
2003 4	10.0821	10.0073	0.0748260	0.0372130	2.01075
2004 1	10.1770	10.0659	0.111078	0.0390240	2.84640
2004 2	10.2117	10.1123	0.0993226	0.0414411	2.39672

una variable dummy (una práctica bastante frecuente en economía aplicada) el modelo econométrico “causal” aparece como competitivo en cuanto a pronósticos.

Con el objetivo de poder apreciar la importancia de los cambios determinísticos con respecto a otras posibles fuentes de fallas de pronóstico se realizaron como ejercicio modificaciones en la relación entre importaciones y PBI (pm, millones de pesos de 1993) a partir de un modelo de *EqC* que supone que las importaciones son, en el largo plazo, 0.10 del PBI y que el coeficiente de ajuste (de corto plazo) es 0.5 en el trimestre, es decir

$$LP \quad IMPA = 0.10 PBI$$

$$CP \quad \Delta LIMPA = 0.5 [LIMPA_{t-1} - \ln(0.10) - LPBI_{t-1}]$$

donde IMPA indica importaciones “artificiales” y el modelo de *EqC* está expresado en logaritmos (L). La serie resultante reparametrizada en niveles se denota como LIMPA.

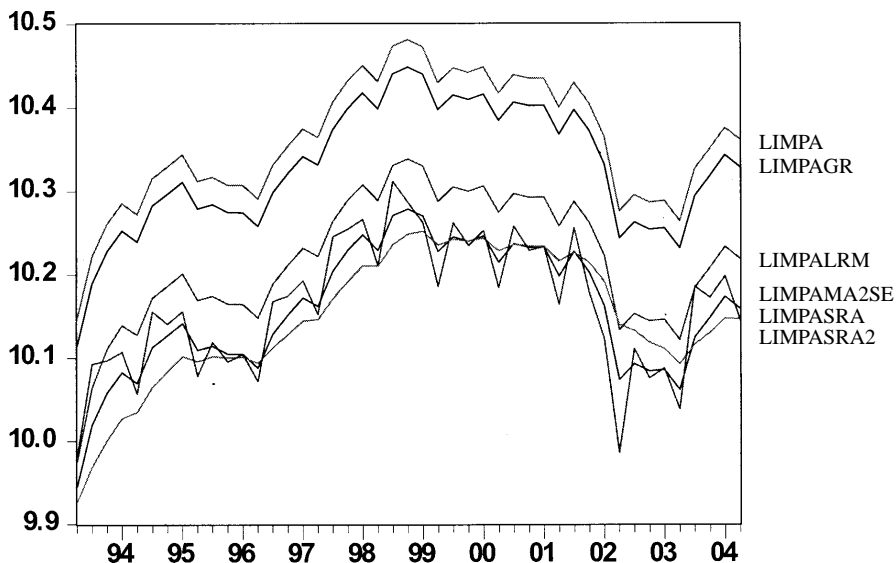
A partir del mismo se generaron otras series artificiales de importaciones cambiando⁸:

- la tasa de crecimiento de 0 a 3% (LIMPAGR)
- el coeficiente de ajuste a 0.25 (LIMPASRA) y 0.99 (LIMPASRA2)
- la media de largo plazo a 0.15 (LIMPALRM)

Estas series artificiales en conjunto se presentan en el gráfico 2. Asimismo, con fines de comparación se incluyó el límite superior del intervalo de confianza al 95 % de la serie base del *EqC* (el SE obtenido de la estimación con series originales, usado para la Tabla 1).

Como puede observarse en el gráfico 2 el cambio en la relación de largo

Gráfico 2. Series de importaciones artificiales



⁸ Si se cambiara la elasticidad de las importaciones al PBI en este tipo de modelo también se cambiaría la media de largo plazo

plazo entre importaciones y producto del 0.10 al 0.15 es el único caso que está fuera del límite superior del intervalo. Nótese que este cambio implica una variación en la media de largo plazo en la formulación de EqC en logaritmos. Este sería el cambio determinístico dominante. El otro cambio determinístico que lleva la tasas de crecimiento trimestral autónoma de las importaciones de 0% a 3% (uno de los máximos crecimientos experimentados por la serie original de importaciones desde 1993) es el segundo en importancia pero muy lejos está del límite del intervalo. Lo más sorprendente son los cambios en la dinámica expresados en este caso por variaciones en el coeficiente de ajuste al equilibrio de largo plazo. Si el ajuste se reduce a la mitad, a solo una cuarta parte de las desviaciones del largo plazo o el ajuste es casi dentro del mismo trimestre no afecta mayormente el comportamiento de la serie.

4. REFLEXIONES FINALES

La pregunta final es qué implican estos resultados de pronósticos para la econometría. Para responderla hay que tener en cuenta al menos tres aspectos.

- **Necesidad de separar objetivos:**

Un método “naive” para pronóstico no nos puede dar la respuesta del efecto de un instrumento sobre un objetivo (evaluación de políticas). Un modelo econométrico (tal vez “fortalecido” por cambios en la media), sí. Incluso puede utilizarse para “corregir” un pronóstico de un modelo “naive” con la respuesta al cambio de política que el cambio estructural induce. (Hendry and Mizon, 2001)

- **No validez para la crítica de Lucas a los modelos econométricos**

Esto se debe a que son los cambios en los componentes determinísticos (no los de los efectos de los instrumentos de política) los principales responsables de los malos pronósticos. Mas aún la dominancia en pronóstico de modelos no

causales a causales cuestiona al mismo principio de “Expectativas racionales” que se basa en relaciones causales (Hendry, 2002)

• **Estimulación a la investigación teórica y empírica en lo relativo a “Co-breaking”**

A medida que las series económicas son más largas, “algunos” shocks no anticipados pueden volverse anticipados. Así como las combinaciones lineales de variables integradas resultan estacionarias cuando hay cointegración, ciertos “breaks” pueden ser removidos a través de combinaciones lineales de variables sujetas a los mismos shocks. Esa es la idea detrás de “Co-breaking” (Clements and Hendry, 2001)

Finalmente vale la pena recordar que en el objetivo de pronóstico, sea o no a través de modelos econométricos, está inmerso el concepto de *futuro* que, por definición, es desconocido y como Singer ha expresado:

“Because of the things we don’t know we don’t know, the future is largely unpredictable” (M. Singer, 1997)

ANEXO. RELACIÓN ENTRE CAPACIDAD PREDICTIVA Y CAMBIOS EN PARÁMETROS

Para recordar la formulación de los tests convencionales de pronóstico supongamos que, para el período de estimación el PGD

$$y_t = \beta x_t + \varepsilon_t \quad \varepsilon_t \approx IID(0, \sigma_\varepsilon) \quad t = 1, \dots, T \quad (A1)$$

mientras que para el período de pronóstico el PGD es $t = T + 1, \dots$

$$y_t = \beta^* x_t + \varepsilon_t \quad \varepsilon_t \approx IID(0, \sigma_\varepsilon^*) \quad t = T + 1, \dots \quad (A2)$$

por lo que los errores de los pronósticos condicionales tendrán como valor esperado

$$\begin{aligned} E[y_{T+i} - y_{T+i}^f | x_{T+i}] &= E[\beta^* x_{T+i} + \varepsilon_{T+i} - \beta x_{T+i}] \\ &= (\beta^* - \beta) x_{T+i} \end{aligned} \quad (A3)$$

dado los supuestos usuales (no hay variables omitidas, regresores endógenos, etc.) en los modelos econométricos que aseguran $E(\varepsilon) = 0$.

Se puede notar que sólo cuando no haya cambio de parámetros ($\beta^* = \beta$) los valores esperados de los errores de pronóstico serán nulos. Esta justamente es la hipótesis nula en los tests de pronósticos más usuales que implican evaluar que los parámetros del modelo son constantes o equivalentemente que los errores de pronóstico tienen esperanza nula (Chow, 1960).

REFERENCIAS

- Clements M.P. and Hendry D.F. (1998), *Forecasting Economic Time Series*, Cambridge Univ. Press.
- Clements M.P. and Hendry D.F. (2001), *Forecasting Non-stationary Economic Time Series*, MIT Press.
- Chow G.(1960), Tests of Equality between sets of coefficient in two linear regressions, *Econometrica*, 64, pp. 5.
- Hendry D.F and Doornik J. (1996), *Empirical Econometric Modelling using PcGive for Windows*, International Thomson Business Press.
- Hendry D.F. and Ericsson N. (2001), *Understanding Economic Forecasts*, MIT Press.
- Hendry D.F. and Mizon G. (2000), “Reformulating Empirical Macroeconometric Modelling”, *Oxford Review of Economic Policy*, vol. 16, 4, pp.138-159.
- Hendry D.F. (2002), “Forecast Failure, Expectation Formation and the Lucas Critique”, *Nuffield College Economics Working Papers*, w8.
- Hendry D.F. (2001), “How Economists Forecast”, in Hendry D.F. and Ericsson N. (eds), *Understanding Economic Forecasts*, MIT Press, chapter 1.
- Singer M. (1997), Thoughts of a nonmillenarian. *Bulletin of the American Academy of Arts and Sciences*, 51, pp 36-51.

CONSTRUCCIÓN DE ÍNDICES COMPUESTOS MENSUALES COINCIDENTE Y LÍDER DE ARGENTINA

JUAN MARIO JORRAT¹
UNIVERSIDAD NACIONAL DE TUCUMÁN

1. RESUMEN

En este capítulo se presentan algunas metodologías para el estudio de los ciclos económicos. Se clasifican las series de Argentina en coincidentes y líderes y se discuten algunos resultados empíricos sobre cada grupo.

Posteriormente se aborda la metodología de construcción de índices compuestos de cada grupo de series. Al mencionar el ajuste por estacionalidad se propone un ajuste adicional por irregulares extremos que aprovecha la información del método X11.

Se presentan los resultados obtenidos para Argentina, el Índice Compuesto Coincidente (ICCO) y el Índice Compuesto Líder (ICLD). Se analiza el ICCO en el contexto de la cronología del ciclo económico argentino, se presentan medidas de las duraciones medianas, amplitud y velocidad de cada una de las fases del ciclo argentino.

Al estudiar el ICLD se verifica que la metodología de construcción del índice compuesto líder agrega eficiencia a los resultados observados en la predicción.

Se propone una manera de analizar las variaciones anuales de estos indicadores en forma de tener una idea de su nivel de significación estadística, respondiendo a la pregunta: ¿Es una variación significativa o no?

E-mail: Jorrat@herrera.unt.edu.ar

¹ Se agradece la colaboración de los Licenciados María J. Granado y Lucas Sal Paz en el proceso de las series.

Finalmente se presentan ejemplos del análisis individual de sendas series, Importaciones Totales (serie coincidente) y la Base Monetaria (serie líder).

2. INTRODUCCIÓN

El estudio de los ciclos económicos se inició poco después de la Gran Depresión² en el National Bureau of Economic Research (NBER) de Estados Unidos de Norteamérica. El Secretario del Tesoro, Henry Morgenthau, solicitó al NBER el desarrollo de un sistema de indicadores económicos confiables, que ayudara al gobierno a detectar con anticipación el inicio de las recesiones, y que permitiera, en lo posible, tomar las medidas económicas correctivas para evitar catástrofes similares a la descrita como Gran Depresión³.

El proyecto fue encomendado a Wesley C. Mitchell y Arthur F. Burns del NBER, quienes presentaron en 1938 la lista de las series propuestas, clasificadas como indicadores líderes, coincidentes o rezagados, según su sincronía con el ciclo económico. Este sistema de clasificación es utilizado aún por el NBER de EE.UU. y por otros países, aunque la lista de series utilizadas ha cambiado a través del tiempo y entre países.

La metodología propuesta por Burns y Mitchell era intensiva en la cantidad de series económicas que representaban la producción de bienes y servicios de un gran número de sectores.⁴ Para cada uno de los tres grupos de indicadores, coincidente, líder y rezagado, se construían índices que medían el porcentaje de series que crecían menos aquellas que decrecían. Estos índices se denominaron *Índices de Difusión por Fase*. Así, se definía el fin de una expansión e inicio de una recesión (pico) cuando, por lo menos, la mitad de las series coincidentes estaban contrayéndose durante un tiempo mínimo

² La Gran Depresión de EE.UU. iniciada en 1929, duró 43 meses, el producto interno bruto real cayó 33%, la producción industrial se contrajo 57%, el empleo cayó 32% y el desempleó se disparó al 25%. G. H. Moore (1985), pag. 21.

³ C. W. J. Granger (1980), cap. 7.

⁴ El número de series en las que se basó la selección en EE.UU. fue de 801 series mensuales.

establecido; es decir, cuando el índice de difusión por fase coincidente se hacía negativo y permanecía menor a cero. El fin de la contracción e inicio de la recuperación (valle), ocurría cuando por lo menos la mitad de las series coincidentes estaban aumentando o, de otra manera, cuando el índice de difusión por fase coincidente se hacía positivo. Esta metodología, basada en *Índices de Difusión*, está todavía en uso y representa una alternativa del tipo “no-paramétrica” a la metodología de los índices compuestos, que se presenta más adelante.

Así, Burns y Mitchell, establecieron las bases empíricas para los estudios de los ciclos económicos (clásicos) y de los ciclos de crecimiento (ciclos de desviaciones de la tendencia). Sin embargo, los índices de difusión no tomaban en cuenta la magnitud de las oscilaciones: no reflejaban ni la profundidad de las recesiones ni el vigor de las recuperaciones. Dando respuesta a estas inquietudes, se desarrolló el método de los *Índices Compuestos* para combinar los movimientos de series económicas tan heterogéneas entre sí que no podían ser agregadas en un número índice. Esta metodología fue introducida por Geoffrey Moore y Julius Shiskin⁵. En particular Moore, modificó la primera lista de indicadores líderes de la economía norteamericana, desarrolló el método de índices compuestos aplicados a varios países, y estudió la relación entre los ciclos de EE.UU. y de los demás países. Su centro de actividades se trasladó posteriormente del NBER al Center for International Business Cycle Research (CIBCR) de la Universidad de Columbia, Nueva York. El CIBCR fue el encargado de construir los indicadores para EE.UU. y los países europeos, en colaboración con la Organization for Economic Cooperation and Development (OECD).

El análisis de los ciclos económicos se generalizó rápidamente, primero a los países más industrializados del mundo y paulatinamente a los países en desarrollo. Las naciones que actualmente publican regularmente los indicadores cíclicos son: En América: Estados Unidos de Norteamérica, Canadá y, luego del *efecto tequila*, México; en Europa: Alemania, Austria, España, Francia,

⁵ G. H. Moore (1958) y J. Shiskin (1961)

Gran Bretaña, Italia, Suecia y Suiza; en el Sudeste Asiático: Australia, China, Corea del Sur, Japón, India, Nueva Zelanda y Taiwán; y otros países: Sudáfrica y Jordania.⁶

Al presente, dentro del NBER, es el Business Cycle Dating Committee el que establece la cronología del ciclo económico de EE.UU. Para ello, define recesión como “*una caída significativa en la actividad económica que afecta a varios sectores, dura por lo menos algunos meses y es generalmente visible en el Producto Interno Bruto real, ingreso personal disponible real menos pagos de transferencias, empleo no agrícola, producción industrial y ventas de productos industriales y de comercios mayoristas y minoristas.*”⁷ El comité se reúne cuando existen datos definitivos de las series mencionadas y evidencia incuestionable que se produjo un punto de giro en la economía norteamericana, para fijar el mes en que dicho cambio se produjo. De esta manera, el NBER estableció la cronología del ciclo de referencia de EE.UU. desde junio de 1857.

En la actualidad, la institución que lleva los indicadores económicos y construye los índices compuestos para EE.UU. de Norteamérica, es The Conference Board.

3. CICLO ECONÓMICO: CONCEPTOS BÁSICOS

Los *ciclos económicos* son fluctuaciones del nivel de actividad económica general que ocurren en todas las naciones que organizan la producción basándose en empresas.⁸ El ciclo económico consiste en periodos de aumentos del nivel de actividad económica, *expansiones*, que ocurren aproximadamente al mismo tiempo en *muchos sectores económicos*; seguidos por etapas de caídas del nivel general de actividad, *recesiones*, que se *funden con la fase de expansión*

⁶ K. R. Phillips, L. Vargas, y V. Zarnowitz (1996). Ver también: <http://www.conference-board.org/economics/bci/>; y <http://www.businesscycle.com/>.

⁷ Business Cycle Dating Committee. Comunicado del 21 de octubre del 2001. NBER: Washington.

⁸ Burns y Mitchell (1946), p. 3

del ciclo siguiente. La sucesión de fases diferentes: expansiones, recesiones, expansiones, etc. se designa *ciclo económico*. En la visión de Burns y Mitchell, la economía está, en todo momento, ya sea en recesión o en expansión.⁹

Los máximos relativos del nivel de actividad económica se denominan *picos*; mientras que los puntos de actividad relativa mínima se designan *valles*. Ambos puntos, picos y valles, se llaman puntos de giro de la economía y constituyen el ciclo económico o ciclo de referencia de la economía.

El período entre un valle (excluido ese mes) y un pico, que se caracteriza por el aumento de la actividad económica, se llama *expansión (o recuperación)*; mientras que el lapso entre un pico (excluido ese mes) y un valle, donde cae el nivel de actividad, es una *recesión (o contracción)*.

Las *expansiones* son periodos de prosperidad caracterizados por aumentos en producción, ingreso, consumo, inversión, ventas, empleo, etc.; y por disminuciones de la tasa de desempleo, del número de empresas en quiebra, etc. Mientras que las *recesiones* se caracterizan por caídas de la producción, el consumo, la inversión, el ingreso, las ventas, el empleo, mientras que aumentan la tasa de desempleo, las quiebras, etc.

Conforme con la tradición del NBER, expresada en la cita precedente, no se establece duración mínima para cada fase del ciclo económico ni límite inferior para la amplitud. Sólo se hace mención de “durante algunos meses” y “ser visible en el PIB”. Por eso seguiremos las pautas de Moore¹⁰ que representan el sentir del CIBCR de la Universidad de Columbia, EE.UU. Para que una fluctuación en la actividad económica general sea considerada ciclo económico se deben cumplir las siguientes condiciones:

1. Duración mínima del ciclo completo de 15 meses (ambas fases sucesivas).
2. Longitud de cada fase, expansión o recesión, de por lo menos 5 meses.
3. Alternancia de las sucesivas fases cíclicas: a un pico siga un valle y viceversa.

⁹ Burns y Mitchell (1946), varios capítulos.

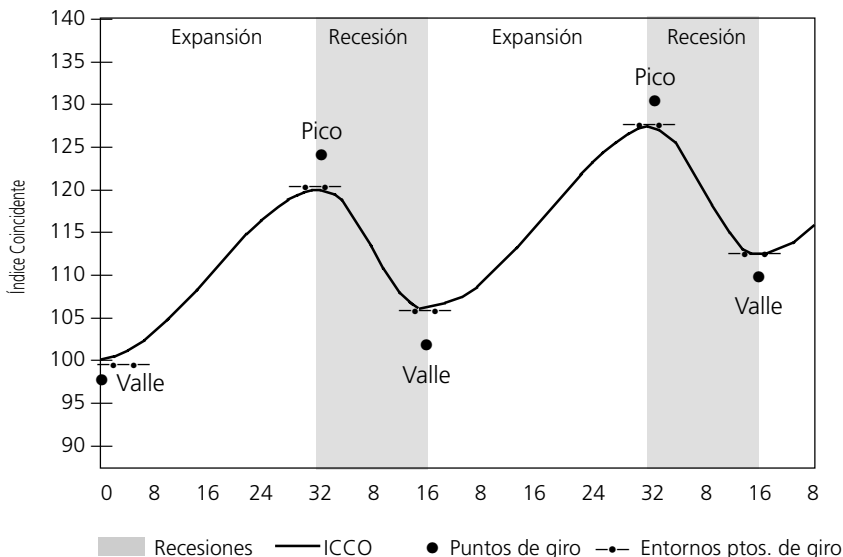
¹⁰ Moore y Moore (1985).

4. Una amplitud significativa dentro del contexto de cada periodo, reflejado en un cambio en el PIB.

Como forma de aclarar estos conceptos se presenta en el Gráfico 1 un esquema estilizado del Índice Compuesto Coincidente de Argentina (ICCO) que representa el nivel de actividad económica. En él se marcan los sucesivos puntos de giro, picos y valles, que determinan las recesiones, indicadas como zonas grises, y las expansiones que corresponden a las regiones claras.

La duración en meses de las expansiones, las recesiones y del ciclo completo, se corresponden aproximadamente con las duraciones medianas del ciclo económico de Argentina. Asimismo, la tendencia que se observa en la serie del ICCO y la amplitud del movimiento cíclico son semejantes a la tendencia mensual del Producto Interno Bruto (PIB) y a su oscilación en el periodo 1970-2003. Más adelante se verán estos temas en detalle.

Gráfico 1. Índice coincidente. Análisis de Ciclo Económico.



Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat - UNT

4. CICLO DE CRECIMIENTO: CONCEPTOS BÁSICOS

Las economías modernas se caracterizan más bien por fluctuaciones del nivel de actividad económica alrededor de su tendencia de largo plazo, antes que por subas y caídas del nivel absoluto de actividad económica. Estas fluctuaciones por arriba o por debajo de la tendencia se designan *ciclo de crecimiento* económico.

En los *ciclos de crecimiento* el término de *aceleración* del crecimiento se refiere a períodos donde el nivel de actividad económica se encuentra *por arriba* de su tendencia de largo plazo. Mientras que se llama *desaceleración* del crecimiento a los lapsos donde el nivel general de actividad está *por debajo* de su tendencia. En los ciclos de crecimiento se estudian los *desvíos* de la actividad económica con respecto a su tendencia de largo plazo; mientras que, en los ciclos económicos, las fluctuaciones se definen en los *niveles* de la actividad económica.

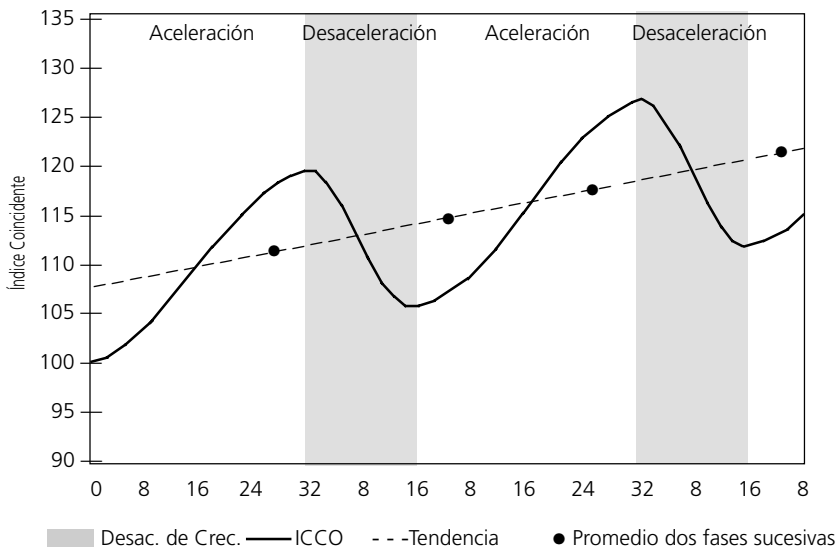
A efectos prácticos, se define como *tendencia de largo plazo* una interpolación suavizada entre los promedios de los valores de la serie correspondientes a dos fases sucesivas del ciclo de crecimiento, estimado en una primera etapa de forma preliminar¹¹. Esta es la metodología seguida por el programa Turning Point Determination (TPD) del NBER que se usa para fechar los ciclos económicos y de crecimiento.

Los puntos de máximo desvío del nivel de actividad económica *por arriba* de su tendencia son los *picos de crecimiento*. En estos picos, el ciclo de crecimiento tiende a preceder a la recesión del ciclo económico: primero se produce una desaceleración de la economía y luego, eventualmente, sigue una recesión, con caída en el nivel de actividad económica.

Los puntos de mayor desvío del nivel de actividad *por debajo* de su tendencia son los *valles de crecimiento*. En ellos, el ciclo de crecimiento coincide o se rezaga ligeramente respecto al ciclo económico.

¹¹ Pueden utilizarse otros métodos para separar la tendencia del ciclo como los filtros de Hodrick-Prescott o de Kalman.

Gráfico 2. Índice coincidente. Análisis de Ciclo de Crecimiento: Índice coincidente y Tendencia.



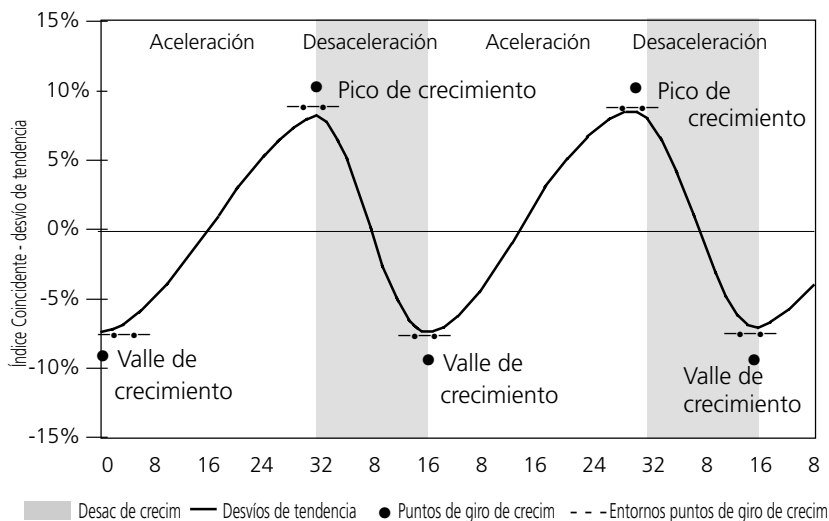
Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat - UNT

Por ello, los ciclos de crecimiento son más simétricos en la duración de sus fases, tienen una menor duración promedio y son más frecuentes que los ciclos económicos.

A modo de ejemplo se presenta en el Gráfico 2 un esquema estilizado del ciclo de crecimiento del ICCO de Argentina. En él se presenta el ICCO y su tendencia de largo plazo. La tendencia de largo plazo se estima interpolando entre los puntos medios de dos fases sucesivas del ciclo de crecimiento preliminar.

La definición de los puntos de giro del ciclo de crecimiento y el establecimiento de las fechas de aceleraciones y desaceleraciones se realiza analizando de “manera clásica” (ciclos absolutos) los desvíos de la tendencia. Esto se presenta esquemáticamente en el Gráfico 3.

Gráfico 3. Índice coincidente. Análisis de Ciclo de Crecimiento: Desvíos de Tendencia.



Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat - UNT

5. CLASIFICACIÓN DE LOS INDICADORES ECONÓMICOS

Dentro de las medidas del nivel de actividad económica ocupa un lugar destacado el Producto Interno Bruto (PIB) por su amplia cobertura y por el cuidado que se tiene para medirlo, pero su periodicidad trimestral y la demora típica en publicarse de 4 meses, disminuyen su utilidad como medida del nivel de actividad económica¹². No obstante, cualquier referencia a ciclo económico debe tener en cuenta el comportamiento del PIB.

Ya vimos que el NBER toma en cuenta para fechar los picos y valles de la economía norteamericana además del PIB otras series mensuales como: el ingreso personal real, la producción industrial y las ventas mayoristas y minoristas.

¹² En Argentina los datos del primer, segundo tercer y cuarto trimestres se publican en junio, septiembre, diciembre y marzo, respectivamente.

Los distintos indicadores económicos que se disponen se clasifican conforme a su comportamiento con el ciclo económico en series coincidentes, series líderes y series rezagadas.

Las *series coincidentes* son aquellas que se mueven en aproximada sincronía entre ellas y con el PIB, y van a constituir el nivel de actividad económica mensual, expresado en el Índice Compuesto Coincidente (ICCO). Los *indicadores líderes* son series que consistentemente se anticipan por algunos meses a los puntos de giro en el nivel de actividad. Mientras que las *series rezagadas* son las que presentan sus puntos de giro con demoras respecto a los del ciclo económico.

La clasificación de las series en coincidentes, líderes y rezagadas propuesta por Moore¹³, demostró la permanencia de los resultados de esta sistematización a lo largo del tiempo en un mismo país y, a través de diferentes países, aún cuando la clasificación original de las series económicas se hizo teniendo en cuenta su desempeño con el ciclo económico en Estados Unidos de Norteamérica. En la Tabla 1 se presentan las series originalmente presentadas por Moore (A) y las actuales según The Conference board (B), clasificadas por sector económico y por su correspondencia cíclica.

Es importante señalar que de las diez series líderes usadas actualmente en EE.UU., siete fueron ya propuestas por Moore en 1950. Un indicador que capte información sobre las primeras decisiones o condiciones relacionadas con la inversión, el empleo o las expectativas de los agentes económicos, tenderá a liderar el nivel de actividad del sector.

Los puntos de giro de una serie particular pueden coincidir con los del ciclo económico, pueden ocurrir algunos meses antes o poco después; en estos tres casos se dice que existe correspondencia entre los puntos de giro específicos y los del ciclo económico. Para cada correspondencia se calcula la diferencia, en meses, entre las fechas del punto de giro de la serie (ciclo específico) y del punto de giro del ciclo de referencia. Algunos de estos rezagos serán negativos, cero o positivos, según la serie se adelante, coincida o se demore en ese punto

¹³ Moore, G. (1950 y 1983).

del ciclo económico. La mediana, un promedio robusto, indica el comportamiento promedio de la serie bajo estudio respecto al ciclo de referencia. El rezago mediano mide el promedio, en meses, en que el indicador se demora (positivo), se adelanta (negativo) o coincide (cero) con los puntos de giro del ciclo económico o de crecimiento, respectivamente. Indica así la sincronía de la serie respecto al ciclo de referencia: coincidente, líder o rezagada.

Así, se define como coincidente aquella serie que en promedio (mediana) presenta una demora en sus puntos de giro entre ± 2 meses, es decir, hasta dos meses antes (-) o después (+), en promedio del correspondiente giro en el ciclo económico.

Las series que se anticipan más de dos meses (mediana menor a -2 meses) en promedio, se clasifican como líderes; mientras que se definen rezagadas las que presentan promedios mayores a dos meses (mediana mayor a +2 meses).

El valor de dos meses es indicativo, discutiremos más al respecto cuando se presenten los resultados de los índices compuestos para Argentina.

Cuando la serie tiene algún punto de giro que no se refleja en el ciclo económico se está ante una falsa señal, puntos de giro extras, y una falta de correspondencia temporal entre los ciclos específicos y de referencia. En otro caso, el ciclo específico puede no contener punto de giro que indique un cambio en el ciclo económico, se presenta entonces un faltante o ausencia de señal. Ambos casos, ciclos extras o faltantes, se consideran una falta de correspondencia temporal disminuye la calidad del indicador. Por ello, para las series incluidas en los estudios de ciclo económico, se exige un alto grado de correspondencia o aciertos entre los puntos de giro de la serie particular y el ciclo económico.

La correspondencia de una serie se mide como el porcentaje del número de puntos de giro específico que pueden claramente relacionarse con el ciclo de referencia (éxitos), dividido en la suma del total de puntos de giro específicos de la serie, más el número de puntos del ciclo de referencia faltantes en el indicador, es decir, se incluye en el denominador los puntos del ciclo de referencia no informados por el ciclo específico. De otra manera, la

Tabla 1. Clasificación cruzada de los Indicadores Económicos de EE.UU. por Actividad Económica y Correspondencia Cíclica. Periodo 1950-2005

Actividad Económica	Correspondencia Cíclica		
	Series Líderes	Series Coincidentes	Series Rezagadas
Empleo y desempleo	(1) Horas promedio semanal por obrero industrial. (A), (B) (2) Nuevos pedidos de seguro de desempleo (invertido). (A) y (B)	(1) Empleo no agrícola. (A), (B) (2) Desempleo (invertido). (A)	(1) Duración media del desempleo. (A), (B)
Producción, ingreso consumo y comercio	(3) Nuevos pedidos de insumos industriales y bienes de consumo. (A), (B) (4) Índice de confianza del consumidor. (B)	(3) Producto Bruto Interno. (A), (B) (4) Producción industrial. (A), (B) (5) Ingreso personal. (A), (B). (6) Ventas industriales y comerciales. (A), (B)	
Inversión fija	(5) Formación de empresas. (A) (6) Nuevos pedidos industriales de plantas y equipos, excluido defensa. (A), (B) (7) Permisos de construcción de nuevas viviendas. (A), (B)		(2) Gasto en inversión en plantas y equipos. (A)
Inventarios e inversión en inventarios	(8) Cambios en los inventarios comerciales. (A) (9) Demoras en las entregas de insumos industriales. (B)		(3) Inventarios comerciales. (A) (4) Razón inventario a ventas en industria y comercio. (B)
Precios, costos y beneficios	(10) Índice de precios de insumos industriales. (A) (11) Índice de precios de acciones. (A), (B) (12) Razón precio a costo unitario laboral, no agrícola. (A)		(5) Cambio en la productividad por hora en la industria (invertido). (A) (6) Costo laboral por unidad de producto en la industria. (B) (7) Índice de precios de servicios al consumidor. (B)
Dinero y crédito	(13) Oferta monetaria. (A), (B) (14) Cambio en crédito en cuotas para consumo. (A) (15) Diferencia entre tasas de interés de largo y corto plazo. (B)		(8) Préstamos industriales y comerciales. (A), (B) (9) Tasa de interés activa a empresas de primera línea. (A), (B) (10) Razón préstamos al consumidor a ingreso personal. (B)

Fuentes: (A): Moore, Geoffrey H. (1959). «Statistical indicators revivals and recessions». Occasional paper 31. National Bureau of Economic Research- New York, NY. (B): The conference board. Business cycle indicators - General information. <http://www.conference-board.org/economicscbci/general.cfm>

correspondencia mide el porcentaje de éxitos del indicador; difiere del 100% por la aparición de señales falsas (puntos de giro extras en el indicador) o por la ausencia de señales (puntos de giro del ciclo económico que faltan en el aserie).

Asimismo, se pueden definir los rezagos (+) ó adelantos (-) medianos y porcentajes de correspondencia tomando como referencia el ciclo económico o el ciclo de crecimiento. En general, se observa que los mismos indicadores andan bien para ambos ciclos, teniendo muchos de ellos mejor desempeño en el de crecimiento.¹⁴

En la Tabla 2 se presenta la lista de series coincidentes determinadas para Argentina¹⁵, además, para cada indicador se presenta su relación con ambos ciclos de referencia, el económico y el de crecimiento, mediante dos valores, el rezago mediano y la correspondencia.

El comportamiento del grupo de indicadores coincidentes es muy adecuado, en el ciclo económico el rezago mediano es de 0,5; menos de un mes, y la correspondencia del 75%. En el análisis de crecimiento los resultados mejoran: mediana de 0,0 meses y aciertos del 83%.

Las dos primeras series, el PIB y el IPI, ya fueron mencionadas al considerar los datos que se evalúan como coincidentes en EE.UU. de Norteamérica y, los resultados de Argentina corroboran su clasificación; el sector de la construcción (ISAC) tiene una conducta pro-cíclica que es deseable aprovechar para determinar el nivel de actividad; las importaciones totales (IMPD) incluyen compras de productos finales para el consumidor, materias primas e insumos para la industria y bienes de capital y manifiesta una conducta pro-cíclica con una alta correspondencia. La recaudación nacional total se incluye por su buen comportamiento en el ciclo de crecimiento (correspondencia 83%) y como reconocimiento a la importancia del estado nacional en la economía.

¹⁴ Moore & Moore (1985).

¹⁵ Para iniciar la clasificación de los indicadores de Argentina consideramos solamente el PIB y la producción industrial, en razón de la disponibilidad de datos. En el Apéndice 1 se describe brevemente el concepto, método y fuente de cada indicador.

Tabla 2. Series coincidentes de Argentina, adelantos y rezagos medianos en meses y correspondencia en los ciclos económico y de crecimiento. Periodo 1970-2005

Nº	Serie Coincidente	Frecuencia e Inicio	Ciclo Económico		Ciclo de Crecimiento	
			Adelantos (-) ó Rezagos (+) Medianos (1)	Correspondencia (2)	Adelantos (-) ó Rezagos (+) Medianos (1)	Correspondencia (2)
1	Producto Bruto Interno a Precios de Mercado (PIB)	Trimestral 1956.1	0.0	100%	0.0	100%
2	Índice de Producción Industrial (IPI)	Mensual 1970.01	0.5	88%	0.0	90%
3	Índice de Actividad del Sector Construcciones (ISAC)	Mensual 1966.01	-1.0	75%	-0.5	67%
4	Importaciones Reales Totales (IMPD)	Mensual 1970.01	0.5	88%	1.0	90%
5	Recaudación Nacional Total (RECTOT)	Mensual 1970.01	0.0	63%	-1.0	83%
6	Índice de Salarios Reales Totales pagados en el Sector Industrial (ISRTE)	Trimestral 1976.1	-3.0	60%	0.0	62%
7	Remuneración Real Total de Asalariados en el SIJP (RRT)	Mensual 1994.07	4.0	100%	2.0	50%
8	Número de Asalariados en el SIJP (NPT)	Mensual 1994.07	4.0	50%	1.0	83%
9	Ventas Minoristas (VMTOT)	Mensual 1995.01	5.0	60%	3.0	100%
Medianas, Grupo Series Coincidentes			0.5	75%	0.0	83%

Notas: (1) Mediana en meses adelantos (negativos) o rezagos (positivos)

(2) Número de puntos de giro de cada serie que corresponden al ciclo de referencia respectivo, sobre el total de puntos de giro de la serie más los puntos faltantes.

Fuente: Programa Ciclos Económicos de Argentina, Director: Juan M. Jorraj, UNT.

Especial consideración merece la inclusión de los salarios reales totales pagados por la industria (ISRTE), a pesar de que es una serie líder en el ciclo económico (rezago mediano -3,0). En Argentina no disponemos de serie mensual prolongada que corresponda al concepto amplio de ingreso personal real,

además, el sector industrial representó en el pasado una fracción importante del empleo total no agrícola, situación que está recuperando actualmente. Asimismo, al considerar como ciclo de referencia el de crecimiento, el ISRTE se comporta como coincidente. Por todas las razones expuestas, ISRTE se incluyó en el grupo de indicadores coincidente argentinos.

Las siguientes tres series: ingreso real total de asalariados (RRT) y número de asalariados (NPT), ambos datos del Sistema Integrado de Jubilaciones y Pensiones (SIJP), y las ventas minoristas reales (VMTOT), son series cortas y, por lo tanto, las medidas de rezago promedio y correspondencia pierden exactitud al calcularse sobre pocos ciclos completos.¹⁶ Sin embargo, se verificó que su exclusión no modificó significativamente los puntos de giro clásicos ni de crecimiento de la economía argentina, ya determinados.

En la Tabla 3 se listan los indicadores líderes de Argentina. También se aprecia que el grupo de series líderes se comporta según lo esperado, en el ciclo económico el adelanto mediano es de 5,5 meses y la correspondencia del 56%, este último valor es relativamente bajo pero mejora cuando la referencia es crecimiento, 72%. Como se observó en otros países, las series líderes brindan también información sobre las aceleraciones o desaceleraciones de crecimiento.

En la tabla mencionada, se nota que todas las series líderes anticipan el ciclo económico y, excepto las horas promedio por obrero industrial (IHPE), también anticipan el ciclo de crecimiento.

En particular, seis de las nueve series arrojan liderazgos aproximados de más de seis meses, estas son: Índice General de la Bolsa de Comercio de Buenos Aires (ICBA), Base Monetaria (BM), Número de Bancarrotas (NBI), Precio Relativo de Manufacturas (VSPRMAN), y Productividad Media del Trabajo en la Industria (VSPMLE).

Las dos últimas series entran con una transformación particular: la variación anual suavizada de seis meses que se calcula como el cambio entre el

¹⁶ Desde fines de 1994 se presentó sólo un ciclo completo en Argentina.

Tabla 3. Series líderes en Argentina, adelantos y rezagos medianos en meses y correspondencia en los ciclos económico y de crecimiento. Periodo 1970-2005

Nº	Serie Líder	Sigla, Frecuencia e Inicio	Ciclo Económico		Ciclo de Crecimiento	
			Adelantos (-) o Rezagos (+) Medianos (1)	Correspon-dencia (2)	Adelantos (-) o Rezagos (+) Medianos (1)	Correspon-dencia (2)
1	Índice General de Valor de la Bolsa de Comercio de Buenos Aires (IBCBA)	Mensual 1970.01	-7.0	65%	-3.0	91%
2	Razón Índices de las Bolsas de Buenos Aires y de Nueva York (RBANY)	Mensual 1980.01	-2.0	56%	-2.0	81%
3	Base Monetaria (BM)	Mensual 1970.01	-5.5	80%	-3.0	59%
4	Número de Bancarrotas (NBR)	Mensual 1969.01	-7.0	53%	-4.5	58%
5	Precio Relativo de Manufacturas. Variación anual suavizada (VSPRMAN)	Mensual 1975.01	-8.0	43%	-5.5	52%
6	Recaudación Real del IVA (RECIVA)	Mensual 1975.01	-2.0	50%	-4.0	68%
7	Índice de Horas Mensuales Promedio por Obrero Industrial (IHPE)	Trimestral 1970.1	-4.5	61%	0.0	91%
8	Índice de Productividad Media del Trabajo en la Industria. Var. anual suavizada (VSPMLE)	Trimestral 1971.1	-7.5	64%	-3.0	72%
9	Sup. Cubierta Autorizada p/ Construcciones Privadas Nuevas y Ampliaciones (SUPCBA)	Mensual 1991.07	-1.5	50%	-2.0	100%
Medianas, Grupo Serie Líderes			-5.5	56%	-3.0	72%

Notas: (1) Mediana en meses adelantos (negativos) o rezagos (positivos)

(2) Número de puntos de giro de cada serie que corresponden al ciclo de referencia respectivo, sobre el total de puntos de giro de la serie más los puntos faltantes.

Fuente: Programa Ciclos Económicos de Argentina, Director: Juan M. Jorrat, UNT.

valor de un mes y el promedio de los doce meses anteriores, expresado como tasa compuesta anual. Para ello se usa el exponente 12/6,5, dónde 6,5 es el intervalo, en meses, entre el mes actual y el promedio de los doce meses anteriores. En fórmula:

$$VAS(x_t) = \left\{ \left[\frac{x_t}{\left(\frac{1}{12} \sum_{j=1}^{12} x_{t-j} \right)} \right]^{12/6.5} - 1 \right\} \cdot 100$$

Esta transformación suaviza la serie y enfatiza el movimiento cíclico, comparada con la tradicional tasa de cambio inter-anual.

6. METODOLOGÍA DE LOS ÍNDICES COMPUESTOS

Aún cuando los grupos de series coincidentes y líderes que se presentaron en la sección anterior tienen comportamientos adecuados, no podemos evitar que en un mes particular algunas nos den información contradictoria. ¿Eso indica un cambio en el ciclo? ¿Estamos ya en presencia de un punto de giro? La metodología de los índices compuestos intenta responder a estas preguntas, entre otras.

Al construir sendos índices compuestos para cada grupo de indicadores, coincidentes y líderes, se avanza en un sistema para la medición, monitoreo y predicción de los ciclos económicos que se realiza mediante la elaboración mensual de los *Índices Compuestos: Coincidente y Líder*¹⁷.

Cada *Índice Compuesto* resume en un único número mensual la conducta de un conjunto de series. Ese conjunto de series se caracteriza por un comportamiento homogéneo respecto del ciclo económico. Todas ellas son aproximadamente sincrónicas (coincidentes), o sistemáticamente se adelantan (líderes), o se demoran (rezagadas); respecto al nivel general de actividad.

¹⁷ En un futuro cercano se completará la trilogía con el *Índice Compuesto Rezagado* para Argentina.

El *Índice Compuesto Coincidente (ICCO)* se construye con el conjunto de series coincidentes que representan el nivel general de actividad económica o estado actual del ciclo económico. El *Índice Compuesto Líder (ICLD)* se forma con el grupo de indicadores líderes que anticipan confiablemente por varios meses a los picos y valles del Índice Coincidente. El Índice Compuesto Líder revela así la dirección futura de la economía.

Como queremos excluir el movimiento recurrente con periodo intra-anual, todas las series se ajustan previamente por estacionalidad.

Luego del ajuste estacional, la variación mensual del índice compuesto se calcula como el promedio de las tasas de cambio logarítmicas mensuales estandarizadas de las series que componen el índice.

El proceso de construcción del índice compuesto se explicita mejor como un proceso por etapas. Pasamos a considerar con mayor detalle cada una de ellas.

6.1. Ajuste por estacionalidad y propuesta de corrección por irregulares extremos

Para el ajuste estacional de las series se seleccionó el programa X-12-SEATS del Census Bureau de EE.UU.¹⁸. Para esta elección se tuvo en cuenta su disponibilidad, sus actualizaciones periódicas, su relativamente fácil utilización, la abundante bibliografía y manuales existentes, la muy buena disposición del *staff* técnico del Census Bureau para evacuar consultas, y la experiencia ya acumulada¹⁹.

La familia de programas X11/X12-ARIMA/X12-SEATS domina el ajuste estacional a nivel mundial, son usados en más del 90% de los datos ajustados por estacionalidad publicados en el mundo. La oficina de estadística de la Unión Europea, Eurostat, seleccionó oficialmente dos programas, el X12-ARIMA del Census Bureau de EE. UU y el TRAMO/SEATS del Banco de España, para el ajuste estacional de todas las series de la Unión Europea.

¹⁸ U. S. Census Bureau. *X-12-ARIMA Reference Manual*, versión 0.3 Beta. May, 2005.

¹⁹ Jorrat (1996-b), Jorrat *et al* (1999, 2002).

El programa X12-SEATS consiste en dos módulos. El primero es el módulo de Modelados RegARIMA (o SEATS). En él se estima un modelo ARIMA estacional (SARIMA), de la forma $(p\ d\ q)\ (P\ D\ Q)_{12}$ para la serie bajo estudio o para una transformación (logarítmica) de la misma, cuando es posible hacerlo. En la expresión anterior p es el orden del modelo autoregresivo regular $AR(p)$; d es el número de diferencias no estacionales; q es el orden del modelo de promedios móviles regular $MA(q)$; P es el orden del modelo autoregresivo estacional $AR(P)$; D es el número de diferencias estacionales; y Q es el orden del promedio móvil estacional $MA(Q)$.

El modelo estimado se usa para extrapolar los valores finales de la serie a fin de extender el número de observaciones (predicción); estimar valores anteriores al primer valor observado (retroproyección), cuando las observaciones son pocas; y detectar y estimar directamente diferentes efectos calendarios en la serie para preajustarla por ellos (días hábiles, pascua y valores extremos). Todo esto se realiza para mejorar posteriormente la estimación de los factores estacionales.

En el segundo módulo, el Ajuste Estacional X-11, la serie mensual o trimestral, ya procesada por el Modelado RegARIMA (extendida y filtrada de los efectos calendarios), se descompone en tendencia-ciclo (T), estacionalidad (S) e irregular (I); según varias descomposiciones alternativas posibles, las más usadas para los datos de Argentina son:

$$\begin{array}{ll} \text{Descomposición multiplicativa} & Y_t = T_t \cdot S_t \cdot I_t \\ \text{Descomposición aditiva} & Y_t = T_t + S_t + I_t \end{array} \quad (1)$$

Para estimar los diferentes componentes de la serie, el X-11 procede iterativamente. El algoritmo básico²⁰ realiza tres estimaciones, sucesivamente más refinadas, de los distintos componentes y en especial del término irregular, para detectar y corregir valores extremos. A su vez, en cada estimación se ejecutan tres iteraciones. Las últimas tres iteraciones de la tercera etapa se denominan B, C y D²¹, letras con la que se designan también las diferentes salidas.

²⁰ Para una descripción más detallada del algoritmo básico consultar Jorrat et al. (2002).

²¹ La etapa A designa las salidas, tablas, cuadros y gráficos del primer módulo, el RegARIMA y otros ajustes previos.

Además se proveen, como etapas E, F, y G, varias medidas de la bondad y calidad del ajuste estacional que interesan para el análisis del ciclo económico²².

El método X-11 incorpora en cada una de sus etapas una elaborada herramienta para la detección y corrección de valores extremos del componente irregular. En particular, se realiza una estimación robusta de su desvío estándar móvil. Esta estimación del error típico se hace en dos pasos. En el primero se estima $s_t^{(I)}$ como el desvío móvil de sesenta meses que se asigna a todas las observaciones del año central; para los dos primeros y los dos últimos dos años se repite el valor estimado más próximo. Posteriormente, se excluyen aquellas observaciones con irregulares muy distantes, es decir, se eliminan los I_t tales que: $|I_t - \bar{I}| > 2.5 \cdot s_t^{(I)}$. Luego, en un segundo paso, se estima de nuevo el desvío estándar móvil $s_t^{(II)}$.

A cada observación del componente irregular se adjudica el peso, W_t , inversamente proporcional a la distancia del irregular respecto a su media, en términos del desvío estándar, según la expresión (2).

$$\begin{aligned}
 W_t &= 1 && \text{para } |I_t - \bar{I}| \leq 1.5s_t^{(II)} \\
 W_t &= \frac{2.5s_t^{(II)} - |I_t - \bar{I}|}{2.5s_t^{(II)} - 1.5s_t^{(II)}} && \text{para } 1.5s_t^{(II)} < |I_t - \bar{I}| < 2.5s_t^{(II)} \\
 W_t &= 0 && \text{para } 2.5s_t^{(II)} \leq |I_t - \bar{I}|
 \end{aligned} \tag{2}$$

Nótese que las observaciones correspondientes al mismo año tienen el mismo valor de $s_t^{(II)}$ y que \bar{I} toma el valor 1 para la descomposición multiplicativa; y 0 para el caso aditivo.

La gráfica de los pesos (W_t) como función de la distancia de su media se presenta en el Gráfico 4. Los W_t finales se informan en la salida denominada tabla C17 (no reportada aquí)²³ del Ajuste Estacional X-11.

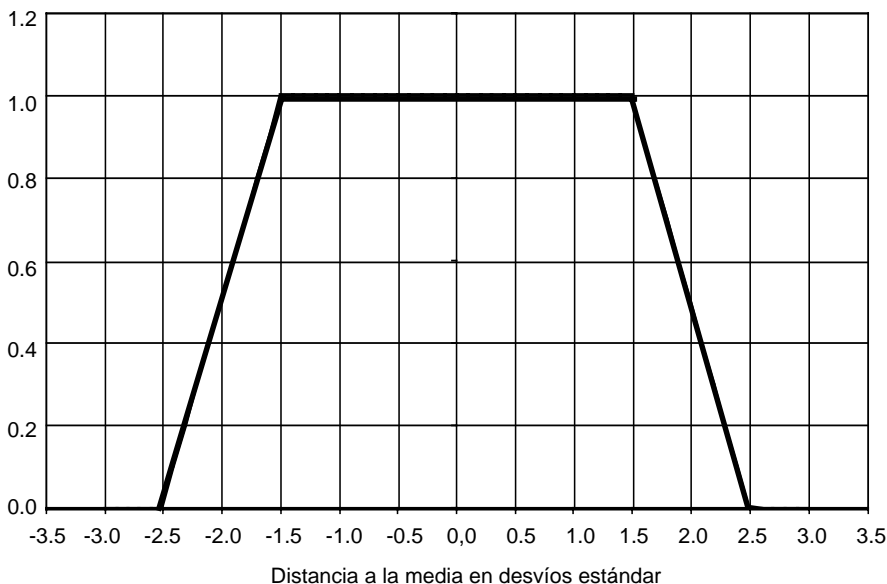
²² Jorrat (2003).

²³ El nombre de la tabla, como ser C17, coincide con la extensión del archivo que contiene los datos.

Otros importantes resultados del X-11 son: La tabla D12^{23a} que contiene la estimación final de la tendencia-ciclo: T_t , La última estimación de los factores para el ajuste estacional, S_t , se presenta en la tabla D16* (para descomposiciones multiplicativa y aditiva). La tabla D11* contiene la serie ajustada por estacionalidad (A_t) definida como:

$$\begin{aligned}
 A_t &= Y_t / S_t = T_t \cdot I_t && \text{para la descomposición multiplicativa} \\
 A_t &= Y_t - S_t = T_t + I_t && \text{para la descomposición aditiva}
 \end{aligned}
 \tag{3}$$

Gráfico 4. Valores de los pesos W_t



Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat - UNT

^{23a} Estas tablas no se presentan en este trabajo.

La tabla E2 comprende la serie ajustada por estacionalidad (A_t ó D11) modificada para las observaciones en las que el componente irregular recibe peso cero ($W_t = 0$), en cuyo caso, el valor de A_t (tabla D11*) es reemplazado por la estimación de la tendencia-ciclo: T_t (tabla D12*). Recordemos que el peso es cero cuando el componente irregular dista de su media, más de dos veces y medias el desvío estándar móvil, como lo indica la expresión .

Tenemos así tres series filtradas alternativas para usar en la construcción del índice compuesto, estos son: 1)La serie ajustada por estacionalidad (tabla D11*); 2)La tendencia-ciclo (tabla D12 *); y 3)La serie ajustada por estacionalidad y modificada por extremos con peso cero (tabla E2*).

La serie ajustada por estacionalidad (tabla D11*) tiene una importante variabilidad asociada al componente irregular, como se aprecia en su definición en la expresión . La tendencia-ciclo (tabla D12*) es una serie demasiado suavizada, muy aplanada en los picos y valles por lo que resulta difícil establecer las fechas de los puntos de giro con precisión. Una alternativa algo mejor es usar la tabla E2*.

La opinión del instituto pionero en el estudio de los ciclos económicos el NBER, es usar las series ajustadas sólo por estacionalidad para la determinación de los ciclos específicos de cada serie, reconociendo que “*para el caso de datos con muchos movimientos irregulares, tales como series para firmas individuales, el análisis de datos suavizados puede ser conveniente*”²⁴. A pesar de este principio general, se reconoce la necesidad de usar datos más suavizados cuando nos movemos a situaciones o países donde las series tienen mayor componente irregular, como es el caso de Argentina.

Las oficinas de estadísticas nacionales están interesadas en proveer estimaciones más suaves que el mero ajuste estacional, como una respuesta a la inestabilidad creciente observada en los indicadores a nivel mundial, para poder determinar los puntos de giro de las series económicas. Así se

²⁴ Bry y Boschan (1971), pag. 79.

ha propuesto²⁵ utilizar la tendencia-ciclo (tabla D12) estimada a partir de ajustar (y proyectar) a la serie de la tabla E2 un modelo SARIMA de la forma $(011)(001)_{12}$ u otro similar. Esta sugerencia brinda estimaciones de T_t con menores revisiones al final del periodo de observaciones, pero no resuelve el problema del aplanado de T_t en los puntos de giro de la serie ajustada.

La alternativa que se propone en este trabajo, extiende el concepto de la tabla E2 a los casos en que el peso está entre cero y uno ($0 < W_t < 1$) y es menos complicada desde el punto de vista computacional. Consiste en aprovechar la metodología robusta del X-11 para detectar y corregir los valores irregulares extremos, definiendo la serie ajustada por estacionalidad y filtrada por irregulares como F_t de la siguiente forma:

$$F_t = W_t \cdot A_t + (1 - W_t) \cdot T_t \quad (4)$$

De manera que cuando $W_t = 0$, la serie filtrada es $F_t = T_t$ como en el caso de la tabla E2; cuando $0 < W_t < 1$, la serie propuesta F_t es un promedio entre A_t y T_t ; y cuando $W_t = 1$, la serie filtrada coincide con la ajustada por estacionalidad, $F_t = A_t$.

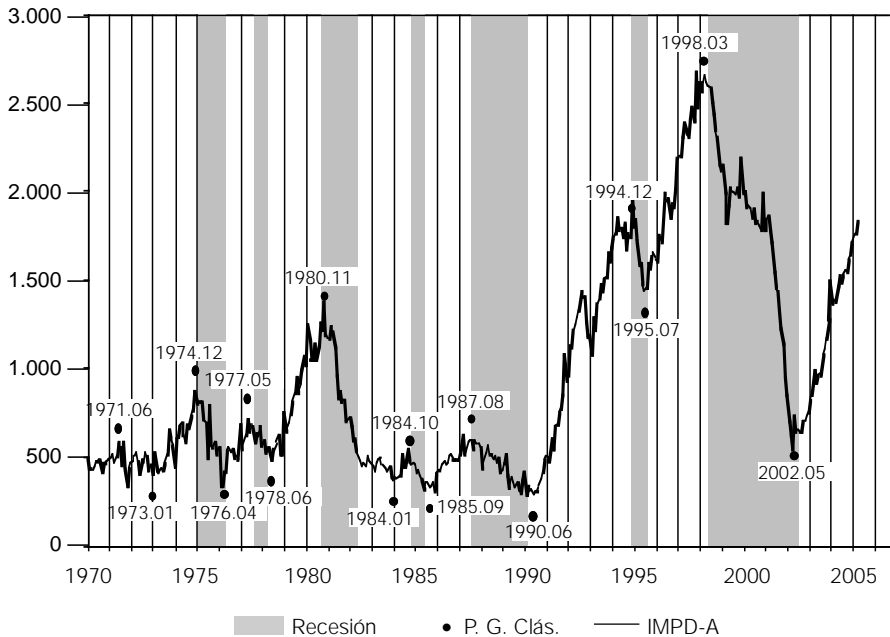
Para ilustrar esta propuesta, presentamos a continuación tres gráficos referidos a la serie de Importaciones Totales, con los puntos de giro definidos conforme a la metodología del NBER²⁶ en la serie F_t .

En el Gráfico 5 se presenta la serie ajustada por estacionalidad, D11, se aprecia el ruido del componente irregular que dificulta la determinación de los puntos de giro.

²⁵ Dagum y Capitanio (1998).

²⁶ Bry y Boschan (1971).

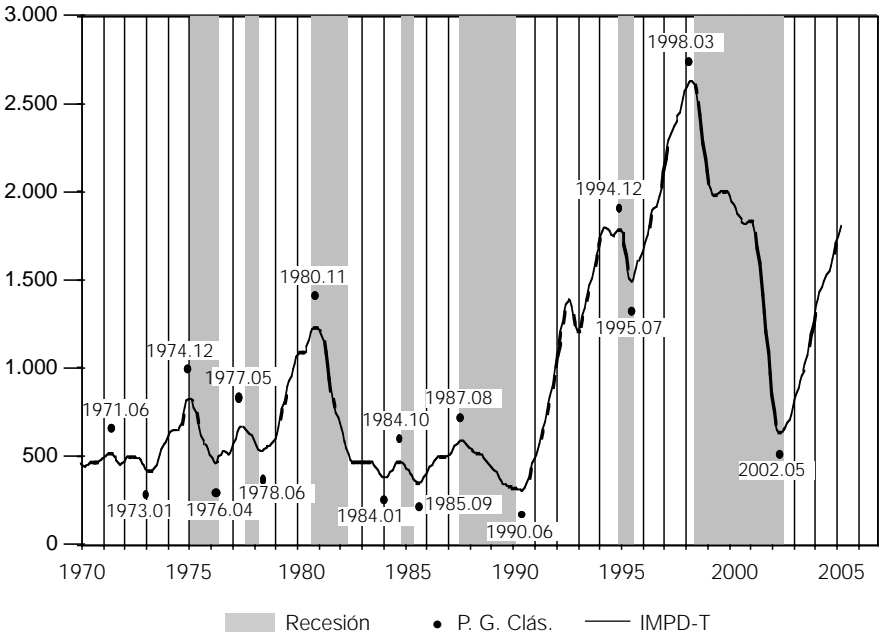
Gráfico 5. Importaciones Totales de Argentina
Millones de dólares de 1993. Serie D11.



Fuente: INDEC y Programa Ciclos Económicos de Argentina; director: Juan M. Jorrot - UNT

En el Gráfico 6 se presenta la serie tendencia final, D12, se observa la demasiada suavidad de la serie que también dificulta la determinación de los puntos de giro.

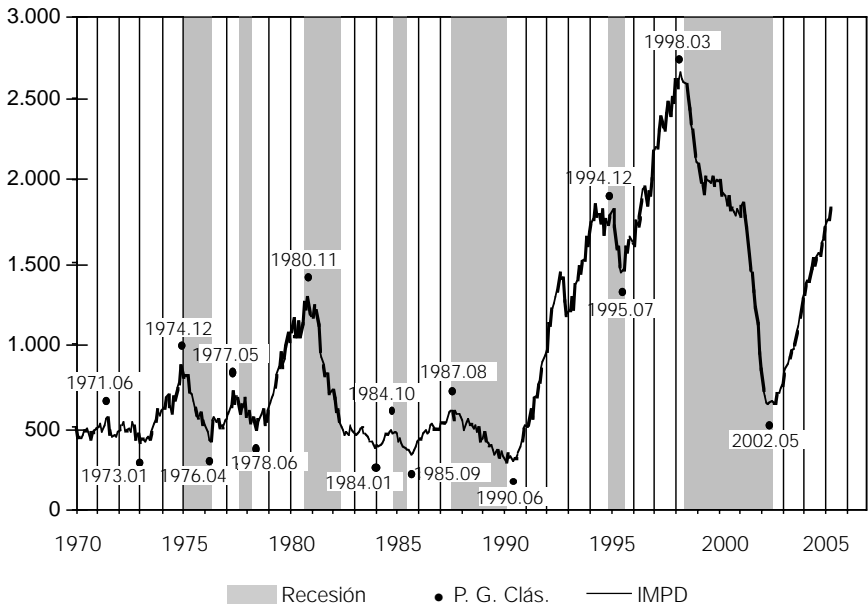
Gráfico 6. Importaciones Totales de Argentina
Millones de Dólares de 1993. Serie D12.



Fuente: INDEC y Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat - UNT

En el Gráfico 7 se presenta la serie propuesta F_t , donde es fácil la determinación de los puntos de giro.

Gráfico 7: Importaciones Totales de Argentina
Millones de Dólares de 1993. Serie F.



Fuente: INDEC y Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat - UNT

6.2. Cómputo de la media y desvío estándar de las tasas de cambio logarítmicas mensuales de cada serie²⁷

Una vez que todas las series se ajustaron por estacionalidad e irregulares de la manera propuesta en el punto anterior, se procede a expresar todas las observaciones como tasas de cambio mensuales logarítmicas. Sea x_{jt} el valor de la serie j en el mes t , entonces:

²⁷ Basado en algoritmo de Moore y otros, Moore y Moore (1985), y Moore (1983, 1990).

$$\hat{x}_{jt} = \ln \left(\frac{x_{jt}}{x_{j(t-1)}} \right) \quad (5)$$

Para las series que ya están definidas como tasas, se toma su diferencia y la expresión queda:

$$\hat{x}_{jt} = \left(x_{jt} - x_{j(t-1)} \right) \quad (6)$$

Las series trimestrales se *expanden*: el valor de cada trimestre se repite en los meses del trimestre correspondiente. Su tasa de cambio trimestral se mensualiza (se divide en tres) y su valor se asigna desde el tercer mes de observación de la serie en adelante, de manera que a cada mes le corresponda la variación adecuada.

Se define un periodo de tiempo donde se estimarán la tendencia (media) y el desvío estándar de cada serie, ya que estos valores son diferentes para cada variable económica y diversos conforme al periodo que se utilice para su cálculo. Este periodo, común para la mayoría de las variables, se define desde el inicio del índice compuesto o el de cada serie (si es posterior) y hasta el final de algún año (completo) previo al momento actual. Conviene elegir el último límite dos o tres años antes del último dato disponible, de manera de reducir las revisiones de tendencia (media) y el desvío estándar en los años finales, ya sean por cambios en los datos o cambios en el ajuste estacional²⁸; asimismo, se debe evitar considerar un año donde ocurre un punto de giro de la economía nacional.

Para el caso de Argentina, la estimación de las tendencias y desvíos estándares se hace desde 1970.01 (enero de 1970) hasta 2003.12 (diciembre de 2003), excepto para las variables que empiezan con posterioridad a 1970.01

²⁸ Para estimar los factores estacionales finales se usa en la mayoría de las series un promedio móvil compuesto 3x5 (siete términos). En este caso, aún sin revisiones de los datos anteriores, las estimaciones dentro de los tres últimos años cambian por la incorporación de nuevos datos a la serie.

que se considera su inicio respectivo. Así las expresiones de la tendencia m_j y desvío estándar s_j de la serie j en el intervalo de tiempo $a_j < t < b$ son:

$$m_j = \frac{1}{(b-a_j)} \sum_{t=a_j}^b \hat{x}_{jt} = \ln \left(\frac{x_{jb}}{x_{ja_j}} \right) / (b-a_j) \quad (7)$$

$$s_j = \sqrt{\frac{1}{(b-a_j-1)} \sum_{t=a_j}^b (\hat{x}_{jt} - m_j)^2} \quad (8)$$

Nótese que, para las series ya expresadas como tasas, el numerador de la expresión debe interpretarse como $(x_{jb} - x_{ja_j})$.

Para las series trimestrales se deben considerar los meses centrales del primer y último trimestre, respectivamente, al definir $a_j < t < b$.

6.3. Primera variación mensual logarítmica del índice compuesto, sin ajuste por amplitud ni por tendencia

La tasa de cambio mensual logarítmica del índice compuesto, sin ajuste por tendencia ni por amplitud, se define como el promedio de las variaciones mensuales estandarizadas de las series que forman el índice. Para las variables ya expresadas como razones, se consideran las diferencias mensuales estandarizadas. Así, para n series componentes tenemos:

$$\hat{c}_t^{(1)} = \frac{1}{n} \sum_{j=1}^n \left(\frac{\hat{x}_{jt} - m_j}{s_j} \right) \quad (9)$$

La expresión tiene media cero y varianza que dependerá de las covarianzas temporales entre las series componentes ya que no son variables mutuamente independientes por la presencia del co-movimiento cíclico.²⁹

²⁹ Recordemos que la suma de variables *no independientes* tienen como varianza:

$$\text{var} \left(\sum_j x_j \right) = \sum_j \sigma_j^2 + \sum_{i \neq j} \sigma_i \sigma_j \rho_{ij}$$

6.4. Segunda variación mensual logarítmica del índice compuesto, con ajuste por amplitud pero sin corrección por tendencia

Como se desea que el índice compuesto tenga la amplitud cíclica del Producto Interno Bruto (PIB), se ajusta el desvío estándar de (9) al del PIB:

$$\hat{C}_t^{(2)} = \hat{C}_t^{(1)} \cdot \frac{s_{PIB}^{(1)}}{s_{C_t}^{(1)}} \quad (10)$$

donde $s_{C_t}^{(j)}$ es el desvío estándar del la tasa de cambio definida en (9) y $s_{PIB}^{(1)}$ es el error típico del PIB.

Así la expresión (10) define la variación mensual logarítmica del índice compuesto sin tendencia. Podemos construir el índice compuesto sin tendencia a partir de la expresión (10) recordando que las tasas son logarítmicas y definiendo como 100 el valor inicial:

$$C_t^{(2)} = C_{t-1}^{(2)} \cdot \exp \left\{ \hat{C}_t^{(2)} \right\} \quad (11)$$

Este índice compuesto -sin tendencia- puede interpretarse como desvíos alrededor de la tendencia en el enfoque del ciclo de crecimiento. Para el caso de Argentina, con una tasa de crecimiento del PIB muy baja en el periodo 1970-2003, el fechado de las recesiones en este índice compuesto –sin tendencia- no difieren de las establecidas con el índice compuesto siguiente.

6.5. Tercera variación mensual logarítmica del índice compuesto, con ajustes por amplitud y por tendencia (final)

El paso final consiste en ajustar a la ecuación , la tendencia del PIB, para reconstruir el Índice Compuesto con tendencia y amplitud cíclica del PIB. Así la tasa de cambio mensual logarítmica del índice compuesto queda definida como:

$$\hat{C}_t^{(3)} = \hat{C}_t^{(2)} + m_{PIB} \quad (12)$$

y el correspondiente índice compuesto final es:

$$C_t^{(3)} = C_{t-1}^{(3)} \cdot \exp \left\{ \hat{c}_t^{(3)} \right\} \quad (13)$$

El índice compuesto se computa mientras se tengan observaciones para la mitad de las series que lo forman, por lo menos.

Analizando las expresiones, (9), (10) y (12), podemos reformular la ecuación (12) de manera que queden explícitas las *contribuciones* de cada serie a la variación mensual del índice compuesto, excepto por la división por el número de series, de la manera siguiente:

$$\hat{c}_t^{(3)} = \frac{1}{n} \left[\sum_{j=1}^n \left(\frac{\hat{x}_{jt} - m_j}{s_j} \cdot \frac{s_{PIB}}{s_{c_t^{(1)}}} + m_{PIB} \right) \right] \quad (14)$$

De esta manera, los índices compuestos de indicadores coincidentes y líderes, resumen información de un grupo de series, tratadas de manera que la variación de cada una puede compararse con las demás y, sobre todo, pueden promediarse de manera que ninguna de las series incluidas imponga la conducta al índice por ser más variable.

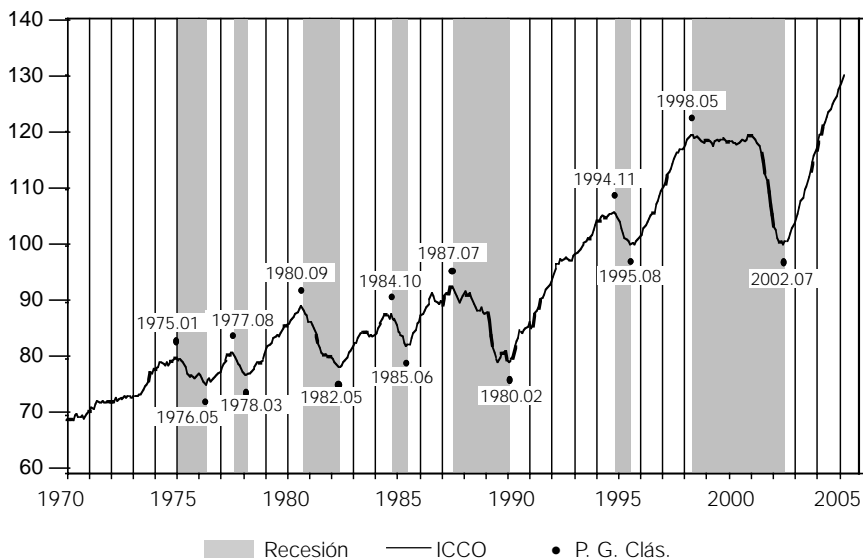
7. ÍNDICE COMPUESTO COINCIDENTE Y CRONOLOGÍA DEL CICLO ECONÓMICO DE ARGENTINA

Con las series coincidentes ya descritas y la metodología previa se confecciona el Índice Compuesto Coincidente de Argentina (ICCO) que mide mensualmente el nivel de actividad económica. Esos resultados se muestran en el Gráfico 8 conjuntamente con los picos y valles determinados con el programa TPD del NBER de acuerdo a las condiciones sobre duraciones mínimas y alternancia mencionadas previamente. Las recesiones, siguiendo la tradición se marcan como zonas grises.

¿Cómo podemos saber que lo que estamos midiendo es realmente actividad económica y no alguna correlación espuria al tener los datos del PBI dentro del índice compuesto? La mejor manera de contestar esta pregunta es hacer el experimento de construir un índice compuesto excluyendo al PBI,

trimestralizarlo³⁰ y correlacionarlo, en niveles y en tasas de cambio con el PBI. Los datos correspondientes se muestran en la Tabla 4. Con esos resultados podemos estar tranquilos. El ICCO mide actividad económica de manera mensual.

Gráfico 8. Índice coincidente de Argentina y el ciclo económico.
Base 1993 = 100. Datos mensuales. Período Enero 1970 - Marzo 2005



Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat - UNT

Tabla 4. Coeficientes de correlación del Índice coincidente (sin PIB) Trimestralizado con el PIB de Argentina. Período 1970-2005

Variable	Coefficiente de Correlación	Número de Trimestres
En niveles	0.98	140
Tasas de cambio Trimestrales	0.80	139

Fuente: Programa Ciclos Económicos de Argentina, director: Juan M. Jorrat, UNT

³⁰ Para trimestralizar directamente se tomó el promedio correspondiente a cada trimestre.

Volviendo a considerar el Gráfico 8, se detecta en él un crecimiento moderado entre 1970-1980, un estancamiento entre 1980-1990 caracterizado por dos recesiones con una gran amplitud, el rápido crecimiento desde 1990, la corta recesión de 1994-1995, la recesión 1998-2002 que resalta por su duración y por la caída en el nivel de actividad y la rápida recuperación desde mediados del 2002.

El ciclo económico de Argentina, las fechas de valles y picos y las duraciones de cada fase y del ciclo total definido de valle a valle o de pico a pico se exhiben en la Tabla 5. En los más de 34 años del estudio, en el periodo 1970-2005, se detectaron 7 expansiones y 6 recesiones. Las recesiones duraron entre 7 meses la mínima (1977-1978), y 50 meses la más larga (1998-2002), con una mediana de 16 meses.

Las 6 expansiones se extendieron desde 15 meses la más corta (1976-1977) hasta 57 meses, la más prolongada (1990-1994) con una duración promedio de 30 meses.

El ciclo económico completo, definido de pico a pico, se extendió de 31 meses (1975-1977) a 88 (1987-1994) el más extenso, con una duración mediana de 40 meses.

Las duraciones promedio de cada fase del ciclo completo y la mayor extensión de las recesiones son hechos estilizados que se detectaron en todos los países.

También se ha demostrado en el caso de los Estados Unidos de Norteamérica que la posibilidad de que se presente el fin de una expansión no depende de su duración. Por ello, no es suficiente contar con las duraciones de cada fase o del ciclo completo para hacer inferencias sobre cambios en la economía.

Además de la duración variable, también podemos apreciar en el Gráfico 8 que las magnitudes de las caídas y de las recuperaciones cambian de ciclo en ciclo. Para analizarlas se usan variaciones logarítmicas entre el inicio y fin de cada fase porque estas tasas son simétricas en el tiempo.

Asimismo, como la duración de las fases son diferentes, se anualizan las variaciones o amplitudes; esta magnitud puede interpretarse como la ve-

Tabla 5. Ciclo económico de Argentina: Fechas de referencia y duración en meses. Período 1970-2005

Fechas de Referencia				Recesión: Pico previo a Valle	Expansión: Valle previo a Pico	Ciclo Económico (1)	
Valles		Picos				Valle previo a Valle	Pico previo a Pico
Año	Mes	Año	Mes				
		1975	1				
1976	5			16			
		1977	8		15		31
1978	3			7		22	
		1980	9		30		37
1982	5			20		50	
		1984	10		29		49
1985	6			8		37	
		1987	7		25		33
1990	2			31		56	
		1994	11		57		88
1995	8			9		66	
		1998	5		33		42
2002	7			50		83	
Mediana (2)				16	30	53	40
Promedio (2)				20	32	52	47
Desv. Est. (2)				16	14	21	21
No. de fases				7	6	6	6

- Notas: (1) La duración total del ciclo se mide de dos maneras: pico a pico y valle a valle.
 (2) Los promedios y los desvíos estándares se redondearon al mes entero.
 (3) La fase incompleta no está incluida en los promedios.

Fuente: Programa Ciclos Económicos de Argentina, Director: Juan M. Jorraj, UNT.

locidad (anual) de la caída o de la recuperación. Estas medidas se presentan en la Tabla 6 para las expansiones y Tabla 7 para las recesiones, referidas al ICCO y al PIB.

En la Tabla 6 se muestra el comportamiento del ICCO y del PIB en la expansión iniciada a mediados del 2002 y que todavía está en curso.

En ambas tablas se observa que el aumento del PIB en las expansiones es aproximadamente el doble que su caída en las contracciones (13,9% versus -7,4%). Así la economía crece a lo largo del tiempo por dos moti-

vos, primero, las expansiones duran más que las recesiones (30 meses versus 16) y segundo, la recuperación del PIB en la fase ascendente es aproximadamente el doble que su disminución en las contracciones.

Particularmente, la recesión 1998-2002 fue la peor en caída del PIB (amplitud -21,8%) pero como tuvo la mayor duración, fue la penúltima peor en velocidad (-5,5% anual) aventajada sólo por la contracción 1975-1976.

Tabla 6. Expansiones del ciclo económico de Argentina. Fechas de referencia, duración, cambios total y anualizado del Producto Interno Bruto y del Índice Compuesto Coincidente, en cada fase. Período 1970-2005.

Valle		Pico		Duración (Meses)	Cambio Total: Amplitud (1)		Var. Anualizada: Velocidad (1)	
Año	Mes	Año	Mes		Producto Interno Bruto	Índice Compuesto Coincidente	Producto Interno Bruto	Índice Compuesto Coincidente
1976	5	1977	8	15	9.2	7.1	7.4	5.7
1978	3	1980	9	30	15.0	14.5	6.0	5.8
1982	5	1984	10	29	7.5	11.2	3.3	4.6
1985	6	1987	7	25	12.8	11.8	6.4	5.6
1990	2	1994	11	57	35.5	28.8	7.5	6.1
1995	8	1998	5	33	19.6	17.5	7.1	6.4
Expansión Incompleta (No incluida en promedios)								
2005	3			32	20.7	25.0	8.3	9.4
Mediana				30	13.9	13.1	6.8	5.8
Promedio				32	16.6	15.2	6.3	5.7
Desv. Est.				14	10.2	7.5	1.5	0.6
Número de Observaciones				6	6	6	6	6

Notas: (1) La variación se calcula como diferencia logarítmica en porcentajes para garantizar simetría en el tiempo.

Fuente: Programa Ciclos Económicos de Argentina, Director: Juan M. Jorrat, UNT.

Tabla 7. Recesiones del ciclo económico de Argentina. Fechas de referencia, duración, cambios total y anualizado del Producto Interno Bruto y del Índice Compuesto Coincidente, en cada fase. Período 1970-2005.

Pico		Valle		Duración (Meses)	Cambio Total: Amplitud (1)		Var. Anualizada: Velocidad (1)	
Año	Mes	Año	Mes		Producto Interno Bruto	Índice Compuesto Coincidente	Producto Interno Bruto	Índice Compuesto Coincidente
1975	1	1976	5	16	-2.8	-5.9	-2.3	-4.4
1977	8	1978	3	7	-7.4	-5.0	-14.8	-8.6
1980	9	1982	5	20	-11.7	-12.7	-7.8	-7.6
1984	10	1985	6	8	-8.5	-6.5	-17.1	-9.7
1987	7	1990	2	31	-17.7	-15.4	-7.1	-6.0
1994	11	1995	8	9	-6.1	-5.5	-8.1	-7.3
1998	5	2002	7	50	-21.8	-17.6	-5.5	-4.2
Mediana				16	-7.4	-5.9	-7.8	-7.3
Promedio				17	-8.6	-7.5	-8.6	-6.3
Desv. Est.				9	4.8	4.8	5.5	3.0
Número de Observaciones				7	7	7	7	7

Notas: (1) La variación se calcula como diferencia logarítmica en porcentajes para garantizar simetría en el tiempo.

Fuente: Programa Ciclos Económicos de Argentina, Director: Juan M. Jorrat, UNT.

8. ÍNDICE COMPUESTO LÍDER DE ARGENTINA

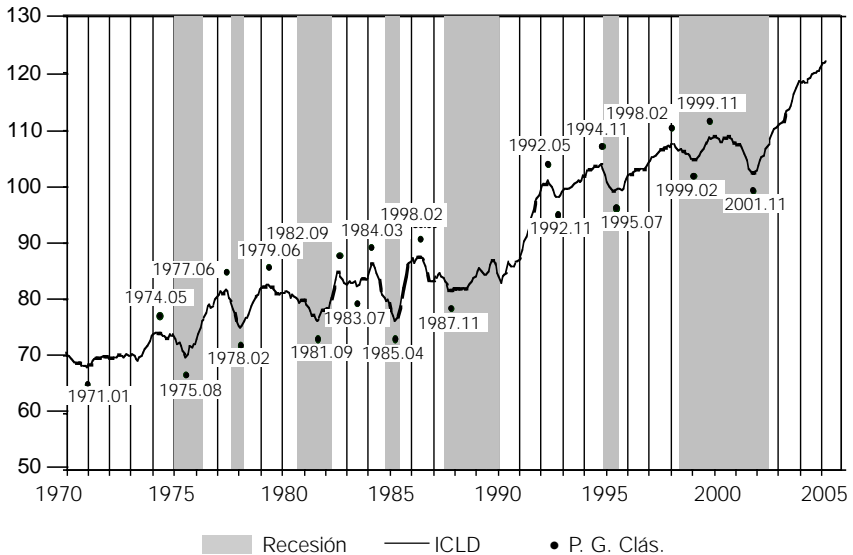
Con el grupo de series líderes se construye mensualmente el Índice Compuesto Líder de la Economía Argentina (ICLD), siguiendo la metodología ya descripta.

En esta sección se muestran los resultados obtenidos y se sugiere una manera de verificar la ventaja de usar indicadores compuestos en adición al análisis de los grupos de indicadores ya detallados.

En el Gráfico 9 se presenta el ICLD, el ciclo de referencia de Argentina, marcando en gris las recesiones del ciclo económico, y los puntos de giro del ciclo específico del ICLD.

Al tratarse de un indicador líder, se espera que sus giros den información anticipada sobre giros en la economía. Así es deseable que los picos del líder se anticipen al inicio de las recesiones, es decir, ocurran antes que el inicio de

Gráfico 9. Índice líder de Argentina y el ciclo económico.
Base 1993 = 100. Datos mensuales. Período Enero 1970 - Abril 2005



Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat - UNT

las zonas grises; mientras que los valles del ICLD, que predicen el fin de las recesiones (fin de las zonas grises) se presentan antes del fin de las zonas recesivas.

Se aprecia en el Gráfico 9 que la conducta del índice compuesto líder sigue el patrón descrito: Picos en zonas claras, antes del principio de las contracciones y valles en las zonas grises, previos al inicio de las recuperaciones. Sin embargo, se presentan algunas señales extras que conviene analizar puntualmente.

En la Tabla 8 se realiza la correspondencia de cada punto de giro del ciclo específico del ICLD con el ciclo económico de Argentina. Observemos en ella que todos los puntos de giro del ciclo económico, con excepción del pico 1994.11, son anticipados por el índice líder y en el caso de 1994.11 hay coincidencia. Recuérdese que la recesión de 1994-1995 fue denominada vulgarmente como “tequila” por la importancia que tuvo el shock externo de México.

La anticipación mediana en los valles es de 8 meses, en los picos es de 7 meses y considerando ambos puntos de giro es 7,5 meses. Es un resultado muy bueno, especialmente si observamos que siempre tenemos en el ICLD un dato más que en el índice coincidente.

La correspondencia en ambos puntos de giro es del 65%, aproximadamente 2 de cada 3 puntos de giro fue anticipado correctamente. Además, de los puntos de giro relacionados con el ciclo económico, el 93% son líderes.

Consideremos ahora las cuatro señales extras en el ICLD que no se reflejan en el ciclo económico. La primera: el valle en 1971.01 corresponde al fin de una desaceleración en índice coincidente; la segunda: contracción extra 1982.09-1983.07 se relaciona con la Guerra de Malvinas; la tercera: expansión extra 1992.05-1992.11 corresponde a una desaceleración de la actividad económica que surge en el análisis de crecimiento; y la cuarta: recuperación extra 1999.02-1999.11 corresponde a un amesetamiento del índice coincidente (expansión en la industria) que se traduce en una aceleración en crecimiento.³¹

Así concluimos que las señales extras del ICLD se explican por graves acontecimientos sociales, Guerra de Malvinas, o porque el índice líder anticipa aceleraciones y desaceleraciones que no llegan a ser expansiones y recesiones, respectivamente. Cabe señalar que no se presentan señales “faltantes” en el índice líder, es decir movimientos en la actividad económica que no fueron anticipados por el ICLD.

Es importante recordar los resultados de la Tabla 3 y compararlos con los de la presente. Los resultados se aprecian en la Tabla 9.

La Tabla 9 no debe sorprender al demostrar que la construcción del índice compuesto líder mejora ambas características deseables: la anticipación mediana en dos meses (de 5,5 a 7,5) y la correspondencia en 11 puntos porcentuales (de 56% a 67%). Es otro caso manifestación del resultado estadístico de que la variabilidad de la media es menor que la de la población y, por ende, la información del promedio es mayor que la de la muestra.

³¹ Recordemos que la economía, en la visión de Burns y Mirchell, no puede star “estancada”.

Tabla 8. Meses de adelantos (-) o rezagos (+) del Índice Compuesto Líder respecto al Ciclo económico de Argentina. Período 1970-2005

Ciclo Económico de Argentina: Fechas de Puntos de Giro				Índice Compuesto Líder: Ciclo Específico					
Valles		Picos		Valles		Picos		Meses de Adelantos (-) o Rezagos (+)	
Año	Mes	Año	Mes	Año	Mes	Año	Mes	Valles	Picos
				1971	1			sct	
		1975	1			1974	5		-8
1976	5			1975	8			-9	
		1977	8			1977	6		-2
1978	3			1978	2			-1	
		1980	9			1979	6		-15
1982	5			1981	9			-8	
						1982	9		sct
				1983	7			sct	
		1984	10			1984	3		-7
1985	6			1985	4			-2	
		1987	7			1986	6		-13
1990	2			1987	11			-27	
						1992	5		sct
				1992	11			sct	
		1994	11			1994	11		0
1995	8			1995	7			-1	
		1998	5			1998	2		-3
				1999	2			sct	
						1999	11		sct
2002	7			2001	11			-8	
Medianas (Meses)								-8.0	-7.0
Ambos puntos de giro								-7.5	
Promedios (Meses)								-8.0	-6.9
Ambos puntos de giro								-7.4	
Desv. Estánd. (Meses)								9.1	5.6
Ambos puntos de giro								7.3	
Correspondencia (%)								64%	70%
Ambos puntos de giro								67%	
Ptos. Giro Adelan. (%)								100%	86%
Ambos puntos de giro								93%	
Ptos. Giro Rezag. (%)								0%	14%
Ambos puntos de giro								7%	

Notas: sct: sin correspondencia temporal.

Espacio de tiempo nulo porque no existen datos o porque alguna serie tiene más puntos de giro que el ciclo de referencia.

Fuente: Programa Ciclos Económicos de Argentina, Director: Juan M. Jorrat, UNT.

Para no agotar al lector, los resultados del análisis cíclico de crecimiento se presentan para ambos índices en el Apéndice 2.

Tabla 9. Comparación del desempeño del Grupo de Indicadores Líderes y el Índice Compuesto Líder de Argentina. Período 1970-2003

Variable	Adelantos Medianos en Meses	Correspondencia
Grupo de Series Líderes	-5.5	56%
Índice Compuesto Líder	-7.5	67%

Fuente: Programa Ciclos Rconómicos de Argentina, Director: Juan M. Jorrat, UNT.

9. OTROS INSTRUMENTOS METODOLÓGICOS

Brevemente se presentan dos instrumentos adicionales a los índices compuestos, útiles para un mejor análisis de la coyuntura económica.

El primero de ellos son los *Índices de Difusión*. Los índices compuestos nos indican el valor y el crecimiento promedio del grupo de indicadores líderes o coincidentes para cada mes. Los *Índices de Difusión* se definen como el porcentaje de series que en los últimos seis meses contribuyeron positivamente al índice compuesto.

Para el caso de las series coincidentes, el Índice de Difusión de Series Coincidentes (IDSC), mide el porcentaje de series que crecen, es decir, que contribuyen positivamente al índice coincidente. Cuando el IDSC es menor al 50% señala un exiguo número de series que están creciendo.

De manera similar para el Índice de Difusión de Series Líderes (IDSL).

Puede observarse en el Apéndice 4 los gráficos de ambos índices de difusión. En la proximidades de un pico (inicio de recesión) ambos índices caen por debajo del 50%, mientras que alrededor de un valle (inicio de una recuperación), ambos índices suben por arriba del 50%.

Este análisis es volver, de alguna manera, a las ideas originales de Michel

y Burns, pero complementa la información cuantitativa de una forma “no-paramétrica”.

El segundo tema se refiere a proponer una manera de computar intervalos de confianza para las variaciones interanuales o suavizadas de seis meses, con algunas propiedades estadísticas. Las series filtradas por estacionalidad e irregulares extremos, conforme a la metodología propuesta previamente, y los índices compuesto tienen básicamente tres componentes: tendencia, ciclo e irregular. Suponiendo un modelo multiplicativo, que es el que más se ajusta a la s series bajo estudio, la serie se descompone como:

$$X_t = T_t \cdot C_t \cdot I_t \quad (15)$$

donde podemos suponer que el componente irregular se distribuye como log-normal con media uno desvío estándar σ_I

$$I_t \sim iid LN(1, \sigma_I) \quad (16)$$

Definiendo la tasa de cambio inter-anual como logarítmica:

$$X_t = \ln\left(\frac{X_t}{X_{t-12}}\right) = \mu + c_t + \ln\left(\frac{I_t}{I_{t-12}}\right) \quad (17)$$

donde μ es la tasa de crecimiento de largo plazo, c_t representa la variación cíclica y el último término es el componente aleatorio de la tasa de cambio inter-anual. Podemos esperar que los componentes irregulares separados 12 meses sean independientes, entonces, la varianza de la tasa de cambio se expresa como:

$$\sigma_X^2 = \sigma_c^2 + 2\sigma_{\ln I}^2 = \sigma_c^2 + 2 \cdot \ln(1 + \sigma_I^2) \approx 2 \cdot \sigma_I^2 \quad (18)$$

La última igualdad se basa en la relación entre las distribuciones normal y log-normal, y la aproximación final en que mientras σ_c^2 es despreciable, $\ln(1 + \sigma_I^2) \approx \sigma_I^2$. Así, un intervalo aproximado del 95% de confianza para la diferencia respecto a la tendencia se expresa:

$$X_t - \mu \pm 1.96 \cdot \sqrt{2} \cdot \sigma_t \quad (19)$$

Así la tasa crítica para anunciar las señales sucesivas de recesión queda:

$$\text{Tasa crítica} = \mu + 1.96 \cdot \sqrt{2} \cdot \sigma_t \quad (20)$$

Las variaciones anuales suavizadas de ambos índices coincidente y líder y sus respectivas tasas críticas se grafican en el Apéndice 3.

AGRADECIMIENTOS

Se agradece a la Asociación Argentina de Economía Política la gentil invitación cursada para participar en la mesa panel de Avances en Econometría que se llevó a cabo en la XXXIX Reunión Anual en la Universidad Católica Argentina en Buenos Aires.

También se agradece al Consejo de Investigaciones de la Universidad Nacional de Tucumán quien apoya financieramente este proyecto de investigación desde 1994.

Asimismo, se agradece la eficiente colaboración de los Licenciados María José Granado y Lucas Sal Paz

BIBLIOGRAFÍA DEL TEMA CICLOS ECONÓMICOS Y ANTECEDENTES EN ARGENTINA

1. Antecedentes metodológicos previos al enfoque de Índices Compuestos Cuantitativos

Arranz, Juan M. & Lidia R. Elías. (1983). "Estudio Preliminar del Ciclo de Referencia para Argentina". Universidad Nacional de Tucumán y *Serie Seminarios del Instituto Torcuato Di Tella*. Buenos Aires: Instituto Torcuato Di Tella.

- Arranz, Juan M. & Lidia R. Elías. (1984). “Ciclo de Referencia para la Economía Argentina, 1960-1982”. Universidad Nacional de Tucumán y *Serie Estudios Técnicos del Centro de Estudios Monetarios y Bancarios*, N° 60. Buenos Aires: Banco Central de la República Argentina.
- Arranz, Juan M. & Lidia R. Elías. (1985). “Medidas de los Ciclos Específicos y del Ciclo de Referencia”. *Instituto de Investigaciones Estadísticas*, Nota N° 41. Tucumán: Universidad Nacional de Tucumán.
- Cuccia, L. (1981). *Tendencias y Fluctuaciones en la actividad del Sector Agropecuario Argentino*. Comisión Económica para América Latina. Buenos Aires: CEPAL.
- Elías, Lidia R. (1982, editor). “Análisis Estadístico de los Ciclos Económicos Argentinos: Reunión de trabajo”. *Instituto de Investigaciones Estadísticas*, Nota N° 27. Tucumán: Universidad Nacional de Tucumán.
- Elías, Lidia R. (1983). “Programas del National Bureau of Economic Research”. *Instituto de Investigaciones Estadísticas*, Nota N° 31. Tucumán: Universidad Nacional de Tucumán.
- Elías, Lidia R.; Juan M. Arranz & Carlos G Rivas. (1989). “Análisis Cíclico de la Economía Argentina”. Universidad Nacional de Tucumán y en *Métodos Estadísticos para Análisis Cíclico y Estacional*, editado por Mentz, R. P.; E. De Alba; A. Espasa & P. A. Morettin. Panamá: Instituto Interamericano de Estadística.
- Elías, Lidia R. & Carlos G Rivas. (1986). “Índices Compuestos de Series Adelantadas, Coincidentes y Atrasadas: Pronósticos”. Universidad Nacional de Tucumán y *Serie Estudios Técnicos del Centro de Estudios Monetarios y Bancarios*, N° 83. Buenos Aires: Banco Central de la República Argentina.
- Elías, Victor J.; Raúl P. Mentz & Nora M. Jarma. (1987). “Bondad de las Predicciones de los Índices Agregados, Adelantados y Atrasados en el Ciclo Económico Argentino: Período 1985-1987”. Universidad Nacional de Tucumán y *XVI Coloquio Argentino de Estadística*. Córdoba, Argentina: Sociedad Argentina de Estadística.

Gorban, Edgardo M. (1979). “Los Ciclos Económicos en Argentina en el Período 1950-1978”. *Seminario*. Tucumán: Universidad Nacional de Tucumán.

Heymann, Daniel. (1980). “Las Fluctuaciones de la Industria Manufacturera Argentina”. *Comisión Económica para América Latina*. Buenos Aires: CEPAL.

2. Metodología Sobre Ciclos Económicos usando el enfoque de Índices Compuestos Cuantitativos

Boschan, Charlotte; & W. W. Ebanks. (1978). “The Phase-Average Trend: A New Way of Measuring Economic Growth”. *Proceedings of the American Statistical Association*.

Bry, Gerhard; & Charlotte Boschan. (1971). “Cyclical Analysis of Time Series: Selected Procedures and Computer Programs”. *Technical Paper National Bureau of Economic Research*, No. 20. New York: Columbia University Press.

Burns, Arthur F. & Wesley C Mitchell. (1946). *Measuring Business Cycles*. National Bureau of Economic Research. New York: Columbia University Press.

Diebold, Francis X. & Glenn D. Rudebusch. (1999). *Business Cycles: Duration, Dynamics, and Forecasting*. New Jersey: Princeton University Press.

Durbin, J.; & S. J. Koopman. (2001). *Time Series Analysis by State Space Methods*. London: Oxford University Press.

Friedman, Milton; & Anna Schwartz. (1982). *Monetary Trends in the United States and the United Kingdom: Their relation to Income, Prices, and Interest Rate, 1869-1975*. Chicago: University of Chicago Press.

Hodrick, R.J. & E.C. Prescott. (1997). “Postwar US Business Cycle: An Empirical Investigation”. *Journal of Money, Credit and Banking*, 29: 1-16.

Klein, Phillip A. & Geoffrey H. Moore. (1983). “The Leading Indicators Approach to Economic Forecasting: Retrospect and Prospect”. *Journal of Forecasting*, Vol. 2.

Klein, Philip A. (1990, editor). *Analyzing Modern Business Cycles: Essays Honoring Geoffrey H. Moore*. Armonk, New York: M.E. Sharper Inc.

- Mintz, Ilse. (1970). *Dating Postwar Business Cycle: Methods and Their Application to West Germany, 1950-67*. New York: NBER, Columbia University Press.
- Mintz, Ilse. (1974). "Dating United States Growth Cycles". *Explorations in Economic Research*. New York: NBER, Columbia University Press.
- Moore, Geoffrey H. (1950). "Statistical Indicators of Cyclical Revivals and Recessions". *Ocasional Paper 31*. National Bureau of Economic Research. New York, N.Y.
- Moore, Geoffrey H. & V. Zarnowitz. (1982). "Sequential signals of recession and recovery". *Journal of Business*, 55:57-85.
- Moore, Geoffrey H. (1983, editor). *Business Cycles, Inflation and Forecasting*, 2° ed. *Studies in Business Cycle* No. 24. National Bureau of Economic Research. Cambridge, Mass.: Ballinger Publishing Company.
- Moore, Geoffrey H. & Melita H. Moore. (1985). *International Economic Indicators: A Sourcebook*. Westport, Connecticut: Greenwood Press.
- Moore, Geoffrey H. (1987). "The Development and Use of International Economic Indicators". En *Statistical Methods for Cyclical and Seasonal Analysis*. Ed. R. P. Mentz, E. de Alba, A. Esasa y P. A. Morettín. Interamerican Statistical Institute, Panamá, 1989.
- Moore, Geoffrey H. (1990, editor). *Leading Indicators for the 1990s*. Homewood, Illinois: Dow Jones-Irwin.
- OECD Department of Economic and Statistics. (1987). *OECD Leading Indicators and Business Cycles in Member countries: 1960-1985*. Sources and Methods 39. Paris: OECD.
- The Conference Board. Business Cycle Indicators - General Information. <http://www.conference-board.org/economics/bci/general.cfm>.

3. Metodología del cálculo de Probabilidades Recursivas de Recesión

- Brier, G. W. (1950). "Verification of forecasts expressed in terms of Probability, *Monthly Weather Review*, 75 (January): 1-3
- Diebold, F. X. & G. D. Rudebusch. (1999). "Scoring the Leading Indicators".

Chapter 15 in *Business Cycles, Duration, Dynamics and Forecasting*. Princeton University Press.

Johnson, Norman L. (1970). *Continuous Univariate Distributions*, 2nd ed. New York: John Wiley & Sons.

Jorrat, Juan M. & Ana M. Cerro. (1999-a). “Probabilidad Mensual de Puntos de Giro de la Economía Argentina Conforme al Índice Líder: 1970-1999”. *Anales XXXIV Reunión Anual Asociación Argentina de Economía Política*. Rosario, Argentina: Universidad Nacional de Rosario.

Neftci S. (1982). “Optimal Prediction of Cyclical Downturns”. *Journal of Economic Dynamics and Control*, Vol. 4 N° 3, pp. 225.

Neftci, S. (1984). “Are economic time series asymmetric over the business cycle?”. *Journal of Political Economy*, 92: 307-328.

Phillips, K; L. Vargas & V. Zarnowitz. (1996). “New Tools for Analyzing The Mexican Economy: Indexes of Coincident and Leading Indicators”. *Economic Review Federal Reserve Bank of Dallas*, Second Quarter.

4. Metodología y Antecedentes del Ajuste Estacional de Series Económicas

Bell, W. R. y S. C. Hillmer. (1984). “Issues Involved with the Seasonal Adjustment of Economic Time Series”. *Journal of Business and Economic Statistics*, 2: 291-320.

Dagum, E. Bee (1988). “The X-11 ARIMA/88 Seasonal Adjustment Method: Foundations and User Manual”. *Time Series Research and Analysis Staff*. Ottawa, Canada: Statistics Canada.

Dagum, E. Bee & A. Capitanio. (1998). Smoothing Methods for short-term Analysis: Cubic splines and Henderson filters. *Statistica*, No. 1, pp. 5-24.

D’Urso Villar, Marcela A. (1995). *Métodos actuales de estimación estacional y su aplicación*. Tesis de Magister en Estadística Aplicada, Facultad de Ciencias Económicas, Universidad Nacional de Tucumán.

Findley, David F.; Brian C. Monsell; William R. Bell; Mark C. Otto & Chen Bor-Chung. (1998). “New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Progra”. *Journal of Business and Economic Statistics*, vol.16.

- Kallek, Shirley (1978). "An Overview of the Objections and Framework of Seasonal Adjustment, Seasonal Analysis of Economic Time Series". *Economic Research Report*, ER-1. Washington, DC: Department of Commerce.
- Lothian, J. & Morry, M. (1978). "A Set of Quality Control Statistics for the X11-ARIMA Seasonal Adjustment Method". *Time Series Research and Analysis Staff*. Ottawa, Canada: Statistics Canada.
- Macaulay, Frederick R. (1931). "The Smoothing of Time Series". *National Bureau of Economic Research*. New York: NBER.
- Monsell, B. C. (1988). "Supplement to Census Technical Paper N° 15: The Uses and Features of X-11.2 and X-11Q.2". *Statistical Research Division*. Washington, DC: Bureau of the Census.
- Piser, Leroy M. (1934). "The Adjustment of Time Data for the Influence of Easter". *Journal of the American Statistical Association*, vol. XXIX.
- Shiskin, Julius; Allan H. Young & John C. Musgrave. (1967). "The X-11 Variant of the Census Method II Seasonal Adjustments Program". Technical Paper N° 15, *Economic Research and Analysis Division*. Washington, DC: Bureau of the Census.
- U.S. Census Bureau (2004). *X-12-SEATS Reference Manual: Version 0.3 Beta*. Washington, DC, USA: May.
- Young, A. H. (1965). "Estimating Trading Day Variation in Monthly Economic Time Series". Technical Paper N° 12, *Economic Research and Analysis Division*. Washington, DC: Bureau of the Census.

5. Resultados del Programa Ciclos Económicos de Argentina, Universidad Nacional de Tucumán

- Jorrat, Juan M.; Nora M. Jarma & Claudia M. Hortt. (1994). "Estudio y Predicción del Ciclo Económico en Argentina". *Carta Económica*, diciembre de 1994/enero de 1995. Buenos Aires: Miguel A. M. Broda y Asoc.
- Jorrat, Juan M.; Nora M. Jarma; & Claudia M. Hortt. (1995). "Sistema de Indicadores y Ciclo Económico». *Kipukamayo*, No. 21: 8-15. Tucumán: Colegio de Graduados en Ciencias Económicas.

- Jorrat, Juan M. (1996). "Ciclos económicos y de crecimiento de Argentina y sistema de indicadores coincidente y líder: Enero 1970-Julio 1996". *Kipukamayo*, No. 25: 049-55, Número especial *XI Congreso de Profesionales en Ciencias Económicas*. Tucumán, Argentina: Colegio de Graduados de Ciencias Económicas.
- Jorrat, Juan M. (1996). "Ajuste Estacional y Corrección de los Factores Estacionales de los Índices de Producción Industrial de Argentina". *Anales XXXI Reunión Anual de la Asociación Argentina de Economía Política*, pp 321-336. Salta, Argentina: Universidad Nacional de Salta.
- Jorrat, Juan M.; Nora M. Jarma; & Claudia M. Hortt. (1997). "Ciclos Económicos y Sistema de Indicadores Coincidente y Líder". *Revista de la Sociedad Argentina de Estadística*, Vol. 1, N° 1: 61-72. Córdoba, argentina: Sociedad Argentina de Estadística.
- Jorrat, Juan M. (1998-a). "Coincident and Leading Indexes for Argentina: Growth Cycles and Forecast of Gross Domestic Product". *Workshop on Regional Economic Models*. Braga, Portugal: University of South Caroline and Universidade do Minho.
- Jorrat, Juan M. (1998-b). "Indicadores del Ciclo Económico de Argentina: Situación a Noviembre de 1998". *Revista del Instituto de Economía Aplicada*, No. 1:12-29. Tucumán, Argentina: Fundación Banco Empresario de Tucumán.
- Jorrat, Juan M. & Abel R. Viglione. (1998). "Pronóstico alentador: seguirá la recesión pero duraría poco". *Ambito Financiero*. Buenos Aires: 29 de diciembre, pp. 5-6.
- Jorrat, Juan M. & Susana E. Salvatierra. (1999). "Ajuste Estacional de Series de Tiempo Económicas de Argentina". *IV Congreso Latinoamericano de Sociedades de Estadística*. Mendoza, Argentina: Universidad Nacional de Cuyo.
- Jorrat, Juan M. & Abel R. Viglione. (1999). "El Valle del Ciclo Económico está muy cerca". *El Cronista*. Buenos Aires: 11 de agosto, pp. 4-5.
- Jorrat, Juan M. (2000-a). "La suba del PBI tendría un techo de 1,8%". *El Cronista*. Buenos Aires: 13 de junio, pp. 1 y 6.

- Jorrat, Juan M. (2000-b). "Indicadores del Ciclo Económico de Argentina: Informe a Julio del 2000". Mimeo (Agosto).
- Jorrat, Juan M. & Ana M. Cerro. (2000). "Computing Turning Point Monthly Probability of the Argentine Economy According to the Leading Index. 1973-2000". *Estudios de Economía*, Vol. 27, N° 2 (Diciembre), pp. 279-295. Santiago, Chile: Universidad de Chile. <http://www.econ.uchile.cl/ede/v27-2-f.pdf>.
- Jorrat, Juan M. & Víctor Iajya. (2002). "El Índice Mensual de Actividad Económica de Tucumán". *Indicadores Económicos y Sociales*. Año 1 N° 1 (Septiembre), pags. 15-18. Tucumán, Argentina: Universidad Nacional de Tucumán.
- Jorrat, J. M., L. Sal Paz y M. J. Catalán. (2002). "Ajuste Estacional de las Series Económicas de Argentina". *Anales XXXVII Reunión Anual Asociación Argentina de Economía Política*. En Internet: www.aaep.org.ar
- Jorrat, J. M. (2003). "Indicador Económico Regional: El Índice Mensual de Actividad Económica de Tucumán (IMAT)". *Anales XXXVIII Reunión Anual Asociación Argentina de Economía Política*. Mendoza, Argentina: Universidad Nacional de Cuyo. En Internet: www.aaep.org.ar.
- Martin Lucero, L. E. (2000). *Los Ciclos Industriales en Argentina: 1980-1999*. Tesis de Magister en Economía, Universidad de San Andrés, Buenos Aires, Argentina, director: Juan M. Jorrat.

APENDICE 1.

Series Coincidentes de Argentina. Periodo 1970 - 2005. Programa Ciclos Económicos de Argentina, director: Juan Mario Jorrat, UNT

(PIB) Producto Interno Bruto a Precios de Mercado.

En millones de pesos de 1993. Trimestral. Primer dato: 1956.1.

Fuentes: Empalme de series de distintas bases de: Base de datos DATAFIEL, Banco Central de la República Argentina (BCRA), Instituto Nacional de Estadística y Censos (INDEC) y Ministerio de Economía (MECON).

(IPI) Índice de Producción Industrial.

Índice Base 1997=100. Mensual. Primer dato: 1970.01.

Fuentes: Elaboración propia en base a series: (1) Mensualización del Índice del Volumen Físico de Producción Industrial (IVFPI), INDEC, 1970-1979; (2) Índice de Producción Industrial (IPIF), FIEL, Mensual, 1980-presente; y (3) Estimador Mensual Industrial (EMI), INDEC, mensual, 1994-presente. Índice mensual del IPIF y EMI, ponderadores estimados por regresión trimestral entre IVFPI como dependiente y regresores: IPIF y EMI.

(ISAC) Índice de Actividad del Sector Construcciones.

Índice Base 1997=100. Mensual. Primer dato: 1969.01

Fuentes: Estimaciones propias en base a datos de: (1) Asociación Argentina de Fabricantes de Cemento Pórtland (AFCP), mensual 1969-presente; (2) Base de datos DATAFIEL, datos históricos; y (3) Indicador Sintético de Actividad del Sector Construcciones de INDEC, mensual, 1993-presente.

(IMPD) Importaciones Reales Totales.

En millones de dólares norteamericanos de 1993, deflactado por Índice de Precios Mayoristas Todos los Bienes de EE.UU. (PPIACO). Mensual. Primer dato: 1969.01

Fuentes: DATAFIEL, INDEC, y Federal Reserve Economic Data (FRED), EE.UU. de Norteamérica.

(RECTOT) Recaudación Nacional Total.

En millones de pesos de 1993, deflactado por Índice de Precios al Consumidor (IPC). Mensual. Primer dato: 1970.01.

Fuente: Min. de Economía e INDEC.

(ISRTE) Índice de Salarios Reales Totales pagados en el Sector Industrial.

Índice Base 1997=100, deflactado por IPC. Trimestral. Primer dato: 1976.1.

Fuentes: Elaboración propia en base a datos Encuesta Industrial Mensual (EIM) de INDEC y Encuesta de Indicadores Laborales (EIL) del Ministerio de Trabajo.

(RRT) Remuneración Real Total de Asalariados en el SIJP.

Miles de pesos de 1993, deflactado por IPC. Mensual. Primer dato: 1994.07.

Fuente: Ministerio de Trabajo, Sistema Integrado de Jubilaciones y Pensiones (SIJP).

(NPT) Número de Asalariados en el SIJP.

Miles de empleados. Mensual. Primer dato: 1994.07.

Fuente: Ministerio de Trabajo, Sistema Integrado de Jubilaciones y Pensiones (SIJP).

(VMTOT) Ventas Minoristas.

En millones de pesos de 1993, deflactado por IPC. Mensual. Primer dato: 1995.01.

Fuente: Elaboración propia en base a datos de INDEC sobre ventas en centros de compras, ventas en supermercados, y ventas de electrodomésticos.

Series Líderes de Argentina. Periodo 1970 - 2005. Programa Ciclos Económicos de Argentina, director: Juan Mario Jorrat, UNT**(IBCBA) Índice General de Valor de la Bolsa de Comercio de Buenos Aires.**

Índice real, deflactado por Índices de Precios Internos Mayoristas (IPIM). Promedio mensual de datos de diarios de cierre. Primer dato: 1970.01.

Fuentes: Bolsa de Comercio de Buenos Aires, El Cronista, DATAFIEL e INDEC.

(RBANY) Razón Índices de las Bolsas de Buenos Aires y de Nueva York.
Índice real, cociente entre el IBCBA y el Índice Dow-Jones Industrial, promedio mensual de cierres diario deflactado por el PPIACO-EE.UU. Mensual. Primer dato: 1980.01.

Fuente: Bolsa de Comercio de Buenos Aires, El Cronista, INDEC y FRED.

(BM) Base Monetaria.

En millones de Pesos de 1993. Mensual. Primer dato: 1940.06.

Fuentes: Banco Central de la República Argentina e INDEC.

(NB) Número de Bancarrotas.

Número de Presentaciones a Quiebra y a Concurso Preventivo en Tribunales Federales de la Ciudad de Buenos Aires. Serie invertida.

Números. Mensual. Primer dato: 1969.01.

Fuente: El Cronista.

(VSPRMAN) Precio Relativo de Manufacturas.

Índice de Precios de Manufacturas respecto al Índice de Precios Internos Mayorista. Variación anual suavizada de seis meses. Mensual. Primer dato: 1975.01.

Fuentes: DATAFIEL e INDEC.

(RECIVA) Recaudación Real del IVA.

En millones de pesos de 1993, deflactado por IPC. Mensual. Primer dato: 1970.01.

Fuente: Min. de Economía e INDEC.

(IHPE) Índice de Horas Mensuales Promedio por Obrero Industrial.

Índice Base 1997=100. Trimestral. Primer dato: 1970.1

Fuente: DATAFIEL e INDEC.

(VSIPMLE) Índice de Productividad Media del Trabajo en la Industria.

Índice Base 1997=100. Variación anual suavizada de seis meses. Trimestral. Primer dato: 1970.1.

Fuentes: DATAFIEL e INDEC.

(SUPCBA) Sup. Cubierta Autorizada para Construcciones Privadas Nuevas y Ampliaciones.

Superficie cubierta autorizada en la Ciudad Autónoma de Buenos Aires.
Miles de m².

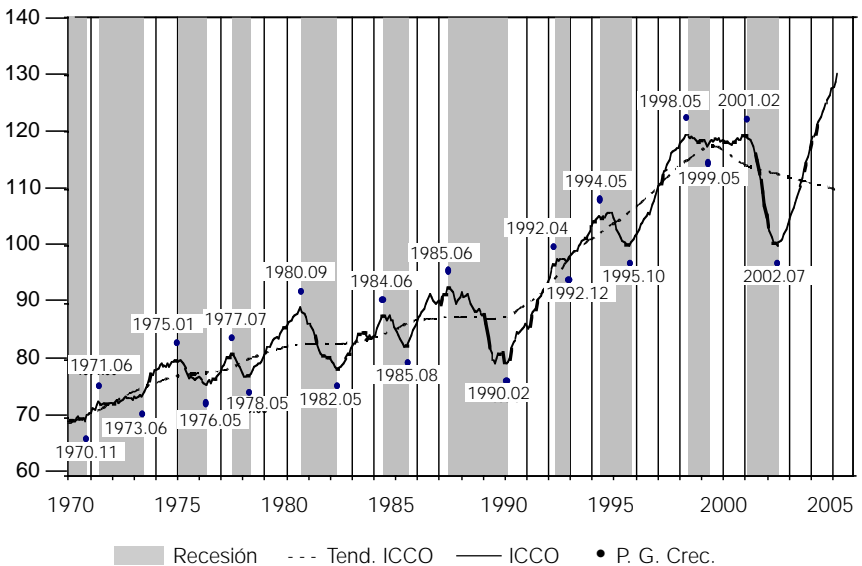
Fuente: Elaboración propia en base a datos de INDEC.

APENDICE 2.

Gráficos de los Índices Compuestos Coincidente y Líder en el Análisis de Crecimiento.

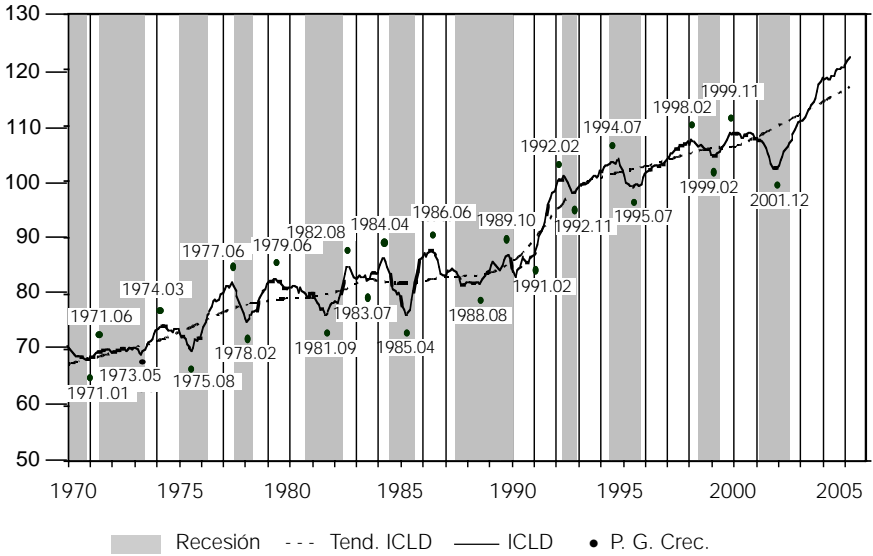
Índice Coincidente de Argentina y Ciclo de Crecimiento.

Base 1993 = 100. Datos mensuales. Período Enero 1970 - Marzo 2005.



Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat. - UNT

Índice Líder de Argentina y Ciclo de Crecimiento.
Base 1993 = 100. Datos mensuales. Período Enero 1990 - Abril 2005.

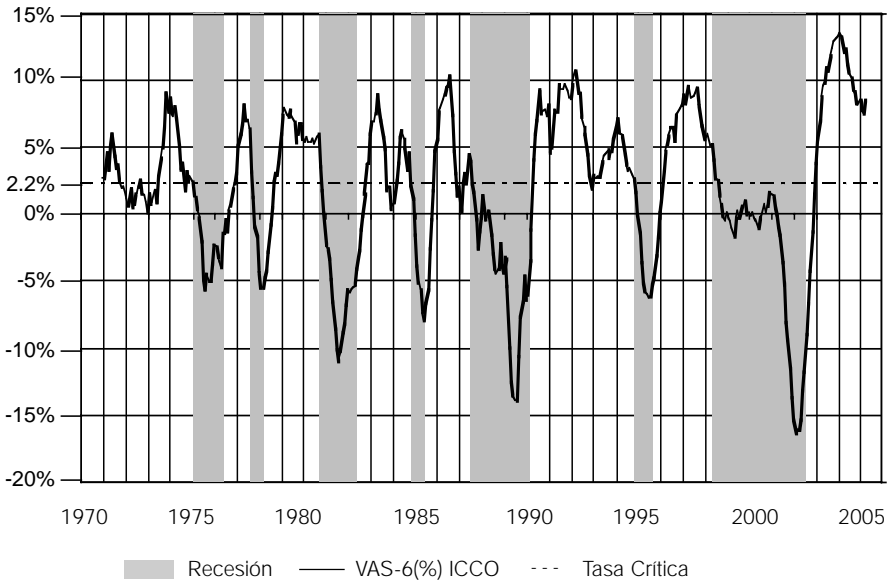


Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat. - UNT

APENDICE 3.

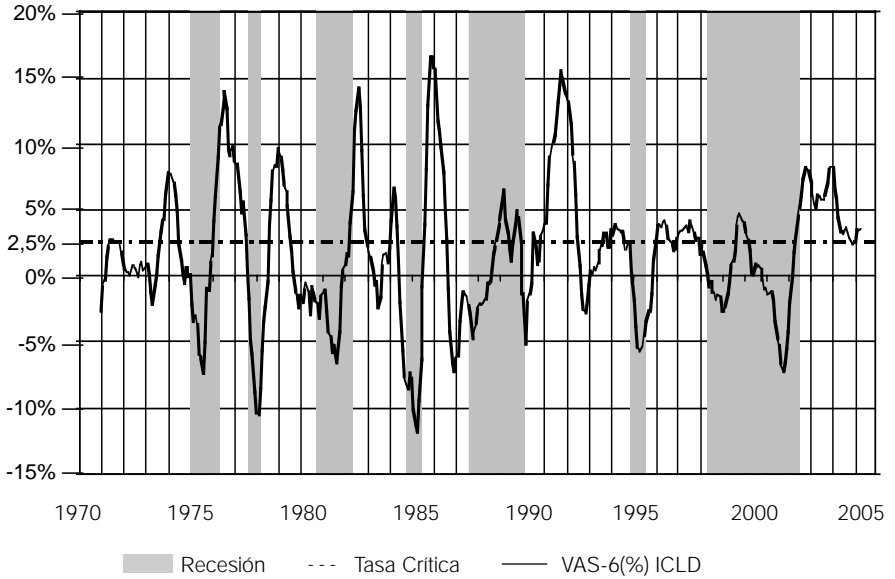
Gráficos de las Variaciones Anuales Suavizadas de los Índices Compuestos Coincidente y Líder en el Ciclo Económico.

**Índice Coincidente de Argentina:
Variación anual suavizada de seis meses.
Datos mensuales. Período Enero 1970 - Marzo 2005.**



Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat. - UNT

Índice Líder de Argentina:
Variación anual suavizada de seis meses.
Datos mensuales. Período Enero 1990 - Abril 2005.

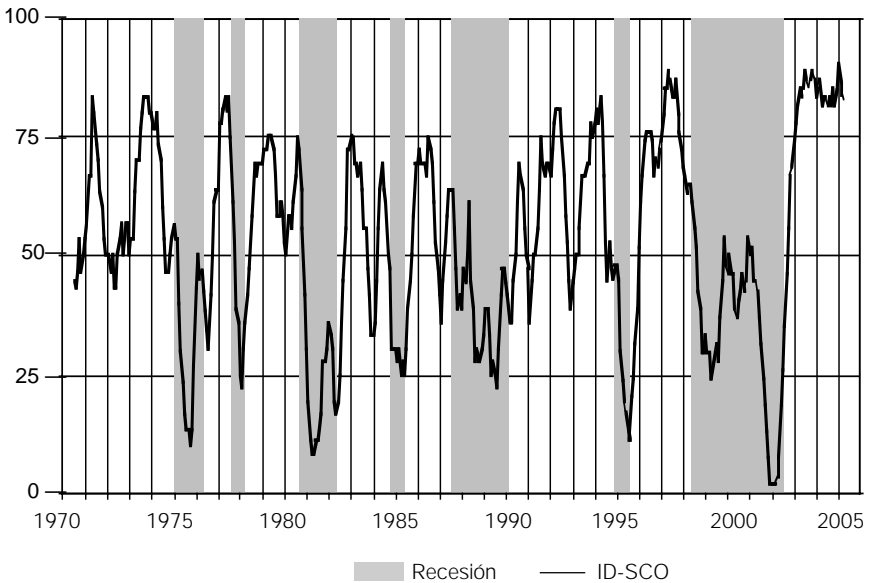


Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat. - UNT

APENDICE 4.

Gráficos de los Índices de Difusión de Series Coincidentes y Líderes en el Ciclo Económico.

**Índice de Difusión de Series Coincidentes de Argentina.
Datos mensuales. Período Enero 1970 - Marzo 2005.**



Fuente: Programa Ciclos Económicos de Argentina; director: Juan M. Jorrat. - UNT

PERSPECTIVAS Y AVANCES RECIENTES EN REGRESIÓN POR CUANTILES

WALTER SOSA ESCUDERO
UNIVERSIDAD DE SAN ANDRÉS*

1. INTRODUCCIÓN

El modelo de regresión lineal múltiple, el “automóvil de la estadística moderna”, a decir de Stigler (1999, pp.320), es la herramienta empírica mas utilizada en economía aplicada. Entre las varias razones que justifican su popularidad, una muy importante es que provee una forma simple y elegante de cuantificar el efecto que una variable tiene sobre otra, aisándolo del posible efecto concurrente que puedan tener otros factores. Una parte sustancial del desarrollo de la econometría moderna estuvo dedicado a elaborar modelos de regresión y métodos de estimación que tuviesen en cuenta en forma explícita las características específicas de los datos y teorías económicas, lo cual ha redundado en una enorme variedad de herramientas con variado grado de sofisticación.

* Este trabajo tiene su origen en mi presentación en el panel sobre “Progresos en Econometría” que tuvo lugar en la XXXIX Reunión Anual de la Asociación Argentina de Economía Política. Agradezco a los organizadores la invitación a participar en dicho evento. Tuve la fortuna de aprender estos métodos directamente de uno de sus creadores, Roger Koenker, durante mis días de estudiante doctoral en la Universidad de Illinois. Ciertamente muchos de los méritos de esta nota reflejan sus útiles enseñanzas, no así los errores y omisiones, que son de mi exclusiva responsabilidad. Agradezco a Roger Koenker y a Gilbert Bassett por un útil intercambio de e-mails sobre los temas de la última sección, a los asistentes al seminario de la Maestría en Biometría y Mejoramiento de la UBA, a Carlos Lamarche y Susana Perelman por sus útiles sugerencias, y a Yanina Azzolina por la tarea de edición final del documento. Correspondencia: Walter Sosa Escudero, Departamento de Economía, Universidad de San Andrés, Vito Dumas 284, B1644BID Victoria, Argentina, Ph/Fax: (54-11) 4725-7024, wsosa@udesa.edu.ar.

Este trabajo presenta una introducción informal a los métodos de *regresión por cuantiles*, los cuales han recibido recientemente un considerable interés en las más diversas áreas de la economía. ¿Qué tienen de distinto los modelos de regresión por cuantiles con respecto a los modelos “estándar” de regresión? No mucho. En todo caso, argumentaremos, estos últimos pueden verse como resúmenes, a veces relevantes y otras veces no, de los modelos de regresión por cuantiles. La sección 2 intenta establecer enfáticamente que los modelos de regresión estándar son en realidad modelos para las *esperanzas condicionales* de una variable. Los modelos de regresión por cuantiles surgen de forma natural cuando el objetivo consiste en caracterizar directamente la *distribución condicional* de una variable, lo cual se discutirá con más detalle en la sección 3, que presenta el modelo básico de regresión por cuantiles. La sección 4 ahonda en varias cuestiones relacionadas con el uso y la interpretación de los modelos de regresión por cuantiles. La sección 5 contiene una aplicación simple para el caso argentino, y la sección 6 orienta al lector interesado en algunas rutas de investigación reciente en esta literatura, tanto empírica como teórica. La última sección justifica porque hemos decidido traducir *quantile regression* como *regresión por cuantiles*. Finalmente, el Apéndice contiene algunos detalles técnicos del método de estimación del modelo, basado en técnicas modernas de programación lineal.

2. ESPERANZAS CONDICIONALES Y REGRESIÓN

Comencemos por considerar un modelo estándar de regresión lineal múltiple:

$$y = x'\beta + u,$$

en donde x es un vector de K variables explicativas, β es un vector de K coeficientes y u es una variable aleatoria que satisface $E(u/x)=0$. Consecuentemente

$$E(y/x) = x'\beta$$

La función $x'\beta$ recibe el nombre de *función de regresión*, y relaciona la esperanza condicional de y con las variables contenidas en el vector x . Una importante característica de esta especificación es que

$$\partial E(y|x) / \partial x_k = \beta_k \quad k = 1, 2, \dots, K,$$

lo cual provee una interpretación simple y elegante para los coeficientes desconocidos β_k : miden el efecto sobre $E(y|x)$ proveniente de alterar marginalmente la k -ésima variable explicativa, manteniendo todas las otras constantes.

Una cuestión que pasa a veces desapercibida en un tratamiento básico del tema es la siguiente. ¿Hasta qué punto es posible extrapolar que β_k mide el efecto marginal de x_k sobre y ? Trivialmente, si el término aleatorio u es funcionalmente independiente de x ,

$$\partial y / \partial x_k = \partial E(y|x) / \partial x_k = \beta_k \quad k=1, 2, \dots, K,$$

de lo que se deduce que si u y x no interactúan, el efecto que x_k tiene sobre la esperanza condicional de y coincide con el que x_k tiene sobre y directamente. Por el contrario, si x y u interactuasen, el efecto que x_k tiene sobre $E(y|x)$ podría ser visto como un resumen del efecto que x tiene sobre y , el cual puede ser relevante o no, como veremos en el siguiente ejemplo.

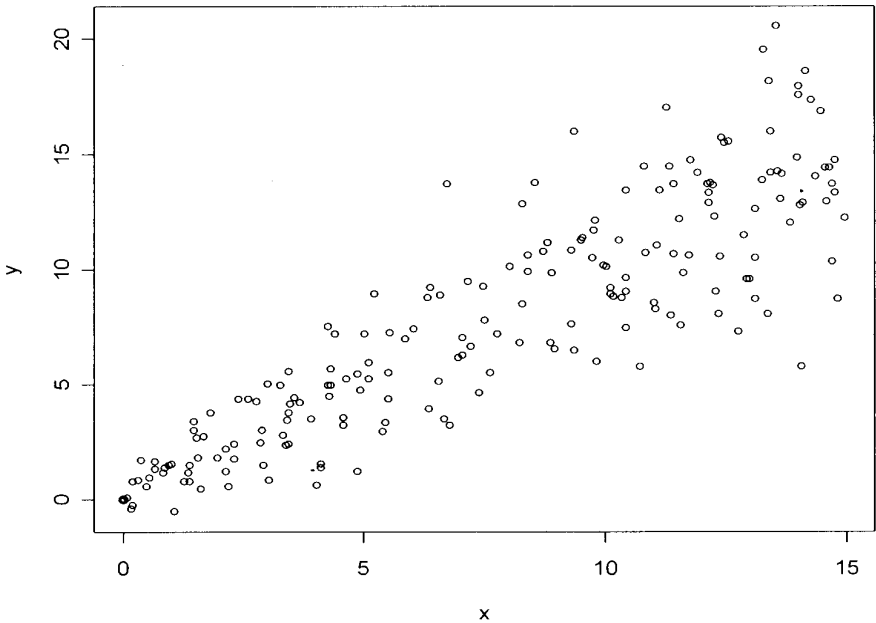
Consideremos un caso clásico de heterocedasticidad. La Figura 1 muestra gráficamente una nube de puntos prototípica de la mayoría de los textos básicos de econometría, en la cual, claramente, la varianza condicional de y aumenta con x .

Frente a este ejemplo, intentemos responder las siguientes preguntas: 1) ¿Qué efecto tiene x sobre $E(y|x)$?, 2) ¿Qué efecto tiene x sobre y ? La respuesta a la primer pregunta puede ser adecuadamente planteada en el contexto del modelo de regresión lineal simple $y = \beta_0 + \beta_1 x + u$, permitiendo que el término de error u sea heterocedástico. Dado que la presencia de heterocedasticidad no altera los supuestos sobre la esperanza condicional de u , $E(y|x) = \beta_0 + \beta_1 x$, el coeficiente β_1 mide en forma muy apropiada el efecto que x tiene sobre $E(y|x)$. La Figura 2 ilustra gráficamente estos resultados: la recta de regresión lineal provee un modelo adecuado para la relación entre $E(y|x)$ y x .

A los efectos de motivar la discusión sobre regresión por cuantiles, observemos cuidadosamente la Figura 3.

En esta Figura hemos agregado un conjunto de rectas de regresión para distintos niveles de la distribución condicional de y dado x . En este ejemplo la

Figura 1



presencia de heterocedasticidad hace que las rectas superiores tengan una pendiente cada vez mayor, de modo que el efecto que x tiene sobre y es notoriamente mayor “arriba” de la distribución: el efecto de x sobre y *no es homogéneo*. A los efectos de enfatizar el contraste, pensemos que estas rectas hubiesen sido paralelas si el término de error hubiese sido homocedástico, de modo que las pendientes de las mismas habrían sido todas iguales. Consecuentemente, en este ejemplo β_1 , si bien es una muy buena medida del efecto que x tiene sobre $E(y/x)$, es un resumen un tanto burdo del efecto que x tiene sobre y , en particular, tiende a subestimar el efecto que x tiene sobre y en la parte superior de la distribución condicional, y lo contrario por debajo.

A fines de dramatizar el ejemplo aún más, la figura 4 provee un ejemplo simple en donde x no tiene efecto alguno sobre $E(y/x)$ y así y todo tiene un fuerte efecto sobre y .

Figura 2

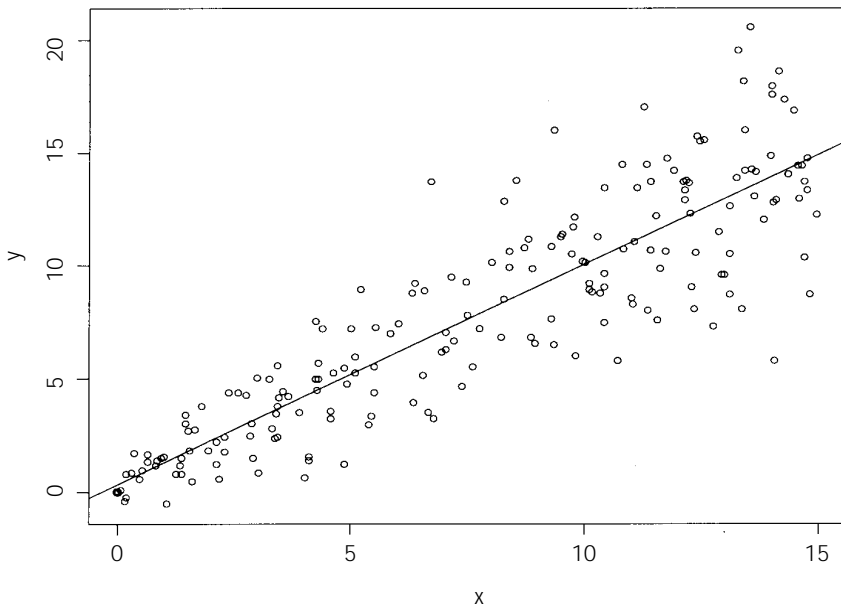
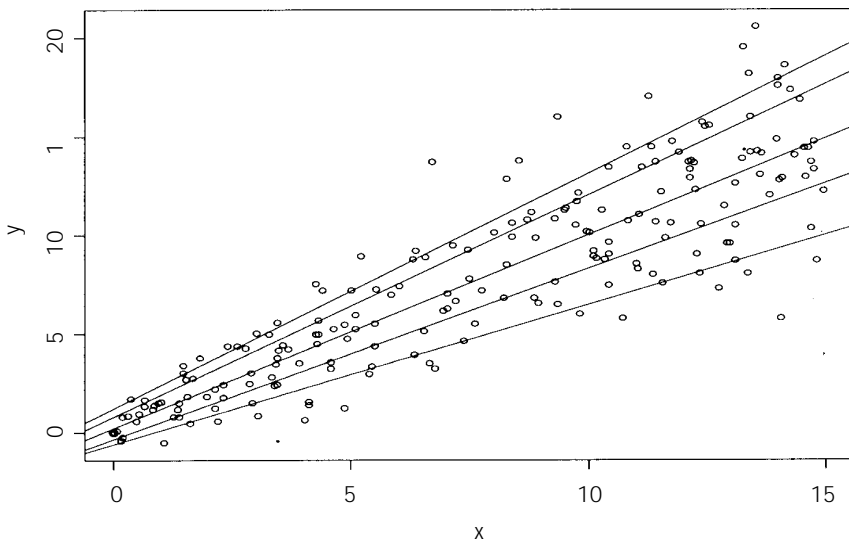
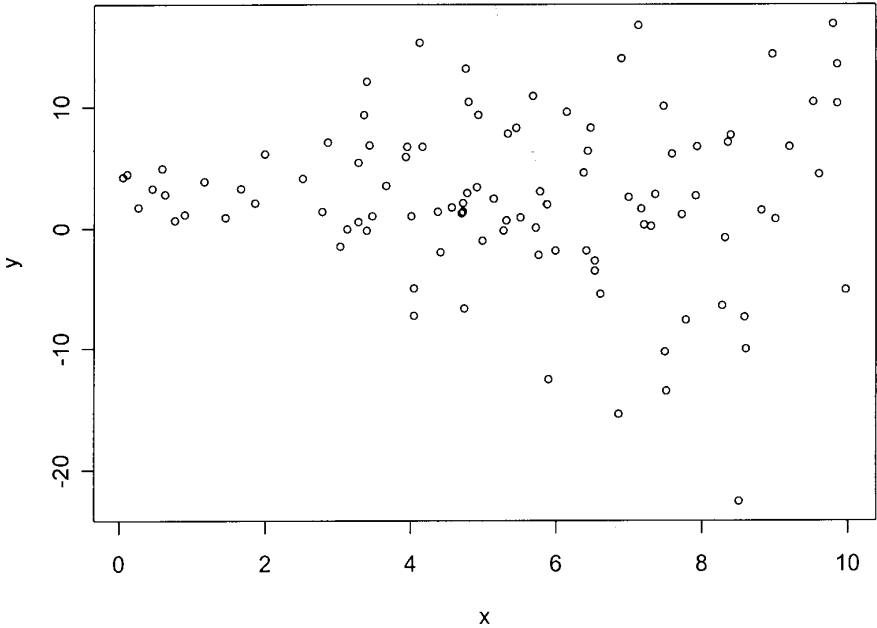


Figura 3



En este caso, claramente, $\beta_j = 0$, de modo que x no altera $E(y/x)$, pero es totalmente erróneo concluir que x no tiene efecto sobre y : x altera la dispersión de y .

Figura 4



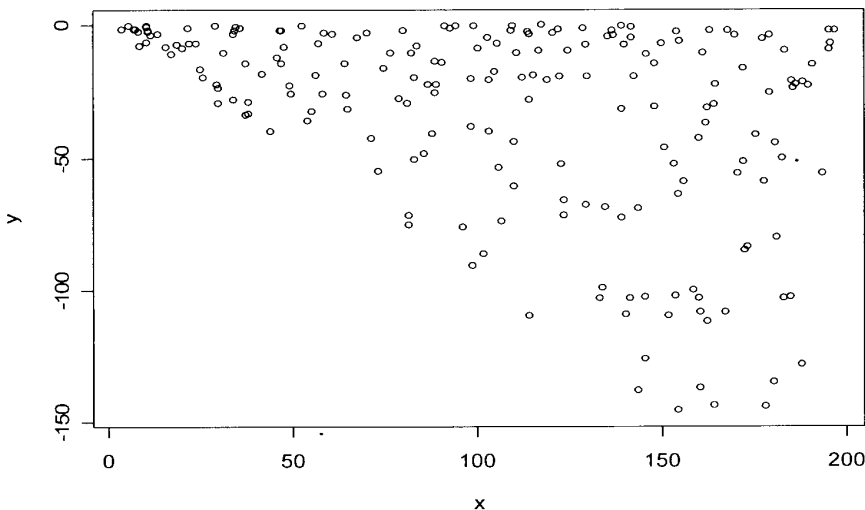
3. CUANTILES CONDICIONALES Y REGRESIÓN POR CUANTILES

Los modelos de regresión por cuantiles intentan modelar el efecto que x tiene sobre *toda* la distribución condicional de y , a diferencia de los modelos de regresión estándar que, como discutiésemos en la sección anterior, se concentran solo en la esperanza condicional.

En los ejemplos anteriores la presencia de heterocedasticidad rompe la independencia entre x y u , de una manera particular: si bien x y u no interactúan en la determinación de la esperanza condicional ($E(u/x)$ sigue siendo cero), sí lo hacen en la determinación de la varianza condicional. Consecuentemente, en el caso de la Figura 2, x hace crecer tanto la esperanza condicional como la

varianza condicional, lo cual ilustra cómo x puede tener un efecto sobre y que va mas allá de alterar $E(y/x)$. En términos generales, es deseable disponer de una estrategia empírica para explorar *cualquier* efecto que x pueda tener sobre la distribución condicional de y sin restringirnos a la media y la varianza. Por ejemplo, consideremos la siguiente configuración de los datos:

Figura 5



En este caso x reduce la esperanza condicional de y pero lo hace aumentando la *asimetría* de la distribución del término aleatorio.

Koenker y Bassett (1978), en un trabajo de crucial importancia, proponen el siguiente modelo de regresión para la distribución condicional de:

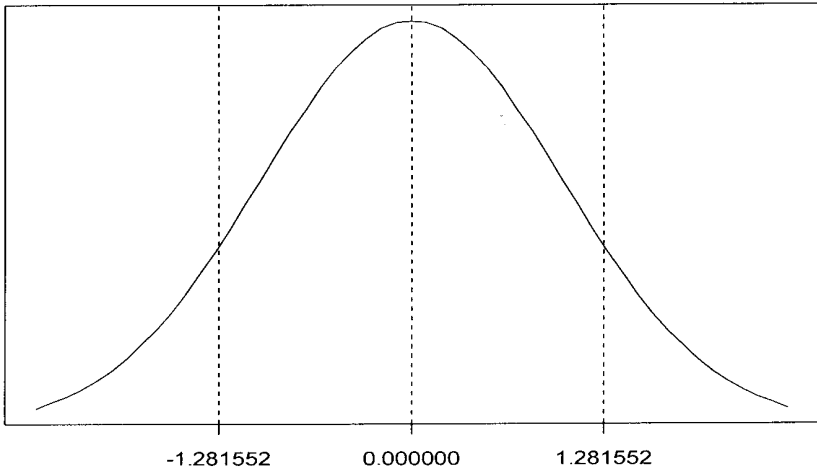
$$Q_{y/x}(t) = x'\beta(\tau) \tag{1}$$

en donde $\tau \in (0,1)$ y $\beta(\tau)$ es un vector de K coeficientes. La notación $Q_{y/x}(\tau)$ hace referencia al τ -ésimo cuantil de la distribución de y condicional en x . Recordemos que para una variable aleatoria Z con función de distribución acumulada $F(z)$ continua y monótona, el τ -ésimo cuantil es un número $Q_Z(\tau)$ que satisface:

$$F(Q_Z(\tau)) = \tau$$

o sea, el τ -ésimo cuantil es un número del soporte de la distribución tal que la probabilidad de que ocurran valores menores es τ ¹. Por ejemplo, cuando $\tau = 0.5$, $Q_Z(0.5)$ es un valor del soporte de la distribución de Z que deja a la izquierda la mitad de la distribución y es conocido como la *mediana* de Z . Análogamente, $Q_Z(0.25)$, el primer cuantil de Z , es un valor del soporte de la distribución, que deja a la izquierda el 25% de la distribución. La Figura 6 ilustra gráficamente algunos cuantiles de la distribución normal estándar ($\tau = 0.10, 0.50$ y 0.90).

Figura 6



A los efectos de entender la información contenida en el modelo especificado por (1), comencemos considerando el caso $\tau = 0.75$ para una sola variable explicativa más una constante. En este caso, el modelo adopta la siguiente forma:

$$Q_{y/x}(0.75) = \beta_0(0.75) + \beta_1(0.75) x$$

¹ La definición de cuantiles para funciones de distribución no invertibles es un poco más delicada. Ver Karr (1993) para mayores detalles.

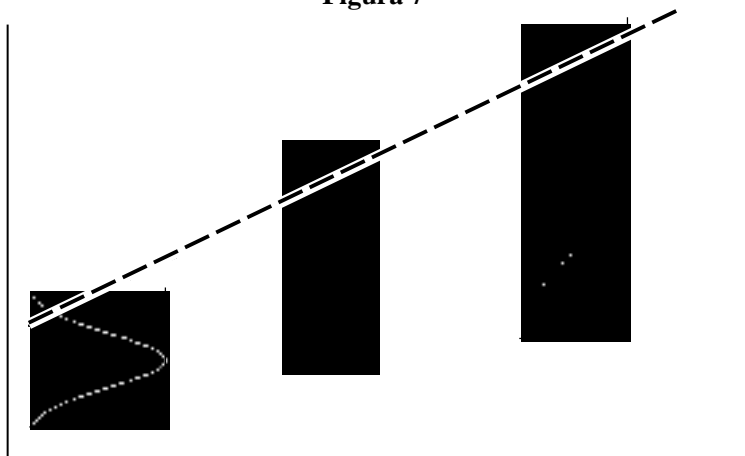
lo cual indica que la función que “une” los cuantiles de la distribución de y dado x a medida que x varía tiene que ser una función lineal. La Figura 7 ilustra la función $Q_{y/x}(0.75)$ para el ejemplo discutido en la sección anterior. A fines ilustrativos también se muestran las funciones de densidad condicional para algunos valores de x . La recta ilustra el hecho de que la misma une los cuantiles 0.75 de cada una de las distribuciones condicionales de y para todos los valores posibles de x .

La gran flexibilidad del modelo de regresión por cuantiles tiene que ver con que los coeficientes $\beta(\tau)$ no están restringidos a ser iguales entre sí para distintos valores de τ , por lo que el modelo (1) hace referencia a una familia de modelos de regresión, uno para cada cuantil de la distribución condicional, de ahí que el mismo es conocido como modelo de *regresión por cuantiles*².

Consideremos las derivadas de (1):

$$\partial Q_{y/x}(\tau) / \partial x_k = \beta_k(\tau)$$

Figura 7



² En términos generales, Manski (1991) define como ‘regresión’ a cualquier característica de la distribución de y condicional en x . Desde este punto de vista, los modelos de regresión estándar se refieren a las esperanzas condicionales y los de regresión por cuantiles, a los cuantiles condicionales.

Esto provee una interpretación para los coeficientes desconocidos del modelo. $\beta_k(\tau)$ se interpreta como el efecto que tiene una alteración marginal en x sobre el τ -ésimo cuantil condicional de la distribución de y . Notar que este coeficiente depende explícitamente de τ , por lo que x puede tener un efecto distinto para los diferentes cuantiles de la distribución de y . A modo de ejemplo, volvamos sobre el caso heterocedástico discutido anteriormente. La Figura 3 muestra las rectas de regresión por cuantiles para varios valores de τ . Notar que en este caso $\beta_l(\tau)$ es siempre mayor que cero (todas las rectas tienen pendiente positiva) y es creciente en τ (las rectas tienen pendiente mayor a medida que consideramos cuantiles superiores). Esto implica que el efecto que x tiene sobre y es mayor en los cuantiles superiores de la distribución condicional de y .

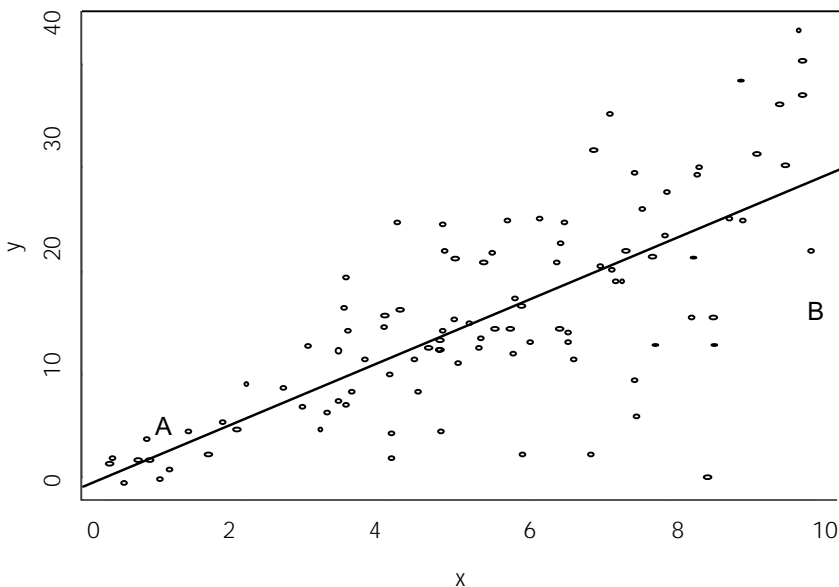
Suponiendo que la primer variable explicativa del modelo es una constante, el caso *homogéneo* corresponde a $\beta_k(\tau) = \beta_k$, $k=2, \dots, K$, es decir, al caso en donde todas las pendientes de las rectas de regresión por cuantiles son iguales para los distintos cuantiles y , consecuentemente, las rectas de regresión son paralelas, de modo que el efecto que x tiene sobre y es *homogéneo* a medida que cambiamos de cuantiles. Ciertamente, y desde este punto de vista, en este caso ideal, la contribución de los modelos de regresión por cuantiles es nula con respecto a la proporcionada por los modelos de regresión estándar, en donde el efecto de x sobre la esperanza condicional resume muy apropiadamente el efecto de x sobre y .

A esta altura es importante notar que un modelo de regresión por cuantiles propone distintas rectas de regresión para distintos niveles de la distribución *condicional* de y . Un error común es pensar que el modelo propone distintas rectas para distintos niveles de la distribución (no condicional) de y . A los efectos de remarcar esta diferencia, volvamos sobre la Figura 7, que ilustra la recta de regresión por cuantiles para $\tau = 0.75$. El sentido en el cual esta recta pasa “por arriba” de los datos es el siguiente: dado x , la probabilidad de observar valores por debajo de la recta es 0.75. Es decir, la recta $Q_{y|x}(0.75)$ tiende a pasar por la parte superior de la distribución condicional de y dado x y, claramente, no por la parte superior de la distribución incondicional de y . Por ejem-

plo, comparemos los puntos *A* y *B* en la Figura 8. *A* es un punto que está muy arriba en la distribución condicional de *y* pero muy abajo en la distribución no condicional de *y*. Lo contrario sucede con el punto *B*. Si pensamos que *y* mide el ingreso mensual en pesos y *x* el stock de educación de un individuo, el punto *A* corresponde a una persona que si bien tiene ingresos bajos comparados con toda la población de referencia, se trata de una persona de ingresos relativamente altos en comparación con aquellos que tienen su misma educación.

La principal contribución del trabajo seminal de Koenker y Bassett (1978) consiste en proveer un marco sólido para la estimación de los parámetros $\beta(\tau)$ en base a una muestra $(y_i, x_i), i=1, \dots, n$. Los detalles del método de estimación pueden ser de un nivel técnico levemente mayor que el elegido para este trabajo, por lo que los mismos son tratados en un Apéndice. Al lector interesado en las aplicaciones le basta con conocer que existen métodos y programas de computación específicos, lo suficientemente probados, que permiten estimar en forma confiable los coeficientes $\beta(\tau)$ y realizar varios ejercicios de

Figura 8



inferencia tales como construir intervalos de confianza o implementar tests de hipótesis.

4. CUESTIONES DE IMPLEMENTACIÓN, INTERPRETACIÓN Y PRESENTACIÓN

Volvamos a nuestro modelo simple para una sola variable explicativa:

$$Q_{y|x}(\tau) = \beta_0(\tau) + \beta_1(\tau) x$$

La Tabla 1 presenta estimaciones de los coeficientes $\beta_0(\tau)$ y $\beta_1(\tau)$ para algunos valores seleccionados de τ : 0.10, 0.30, 0.50, 0.70 y 0.90. La información contenida en cada par de filas es similar a la información relevante disponible cuando se estima un modelo simple de regresión: la primer fila contiene los coeficientes estimados y debajo se presentan estimaciones de un estadístico “ τ ” correspondiente a la hipótesis nula de que el coeficiente poblacional es cero, los cuales fueron obtenidos con los métodos discutidos en el Apéndice, las filas tres y cuatro contienen información similar para los coeficientes $\beta_1(\tau)$. Nótese que la Tabla contiene información para una *familia* de modelos de regresión, uno para cada uno de los cuantiles seleccionados en cada columna. La última columna presenta los resultados de estimar por el método de mínimos cuadrados ordinarios (MCO), que en este caso son muy similares a las obtenidas para el cuantil 0.50. Las rectas de regresión contenidas en la Figura 3 se corresponden a las estimaciones de esta Tabla. La elección de los valores τ : 0.10, 0.30, 0.50, 0.70 y 0.90. fue completamente arbitraria, en realidad podríamos haber elegido cualquier conjunto de valores del intervalo (0,1) y por razones de simplicidad en la presentación es habitual concentrar el análisis en un subconjunto relevante de valores³.

³ Potencialmente se podría pensar en estimar el modelo en el continuo τ en (0,1). Afortunadamente la cantidad de soluciones distintas es, generalmente $< 3n$, es decir, para un tamaño de muestra n la cantidad de rectas esencialmente distintas no es infinita, en el mismo sentido en que para una muestra de n observaciones hay n cuantiles muestrales esencialmente diferentes. Ver Portnoy (1991).

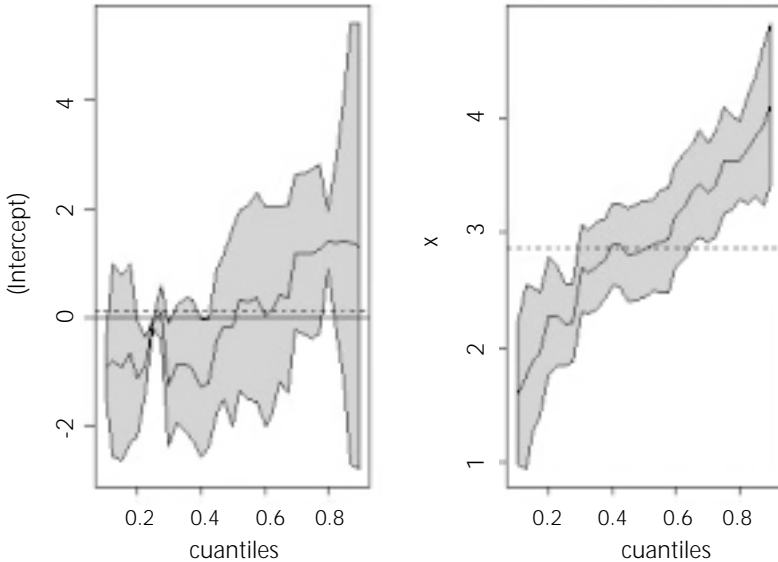
Tabla 1. Ejemplo simple

hetebasic	0.1	0.3	0.5	0.7	0.9	MCO
Intercepto	-0.936 (-3.271)	-1.221 (-1.797)	-0.176 (-0.161)	1.186 (1.368)	1.305 (0.528)	0.141 (0.088)
x	1.587 (4.324)	2.689 (11.593)	2.842 (11.207)	3.340 (12.470)	4.122 (9.804)	2.860 (10.686)

La Figura 9 presenta una forma habitual de mostrar los resultados de un modelo de regresión por cuantiles. Para nuestro ejemplo simple, se trata de dos gráficos, cada uno de ellos se refiere a cada uno de los coeficientes a estimar. La curva sólida central en el gráfico izquierdo representa los valores estimados de $\beta_0(\tau)$ para los distintos valores de τ , en este caso, para una grilla muy fina de posibles valores en el intervalo (0,1). La “sombra” representa un intervalo de confianza al 95%. Que la curva estimada $\hat{\beta}_0(\tau)$ presente un leve comportamiento creciente es una representación del hecho de que en nuestro caso las ordenadas al origen de las rectas de regresión por cuantiles en general lo son. El gráfico de la derecha hace lo propio con los coeficientes $\beta_1(\tau)$ estimados. En este caso, el aspecto creciente de la línea de estimación refleja el hecho de que las pendientes son crecientes en los cuantiles condicionales.

En lo que se refiere a la interpretación de los coeficientes, ya hemos aclarado que, al menos técnicamente, para un valor dado de τ , las pendientes tienen la interpretación de cambios en los cuantiles condicionales de y provenientes de cambios marginales en x , *ceteris paribus*. Consideremos un ejemplo estudiado por Manning, Blumberg y Moulton (1995). La variable explicada (y) es el consumo individual de alcohol por semana, y el vector de variables explicativas (x) incluye, entre otras, el precio promedio pagado por individuo, el ingreso familiar y un conjunto de características demográficas tales como la edad, sexo, tamaño de la familia, entre otras; el trabajo citado contiene una descripción detallada de estas variables, y el objetivo del mismo consiste en estimar la elasticidad precio de la demanda de alcohol. Los autores argumentan que el vector de características observables x “explican muy poco la varia-

Figura 9



bilidad en el consumo total”⁴, e implícitamente reconocen que mas allá de estos factores los individuos difieren notoriamente en sus preferencias por el alcohol, las cuales no son observables directamente. En términos geométricos, en el caso de un solo factor observable, una nube de puntos en donde y representa el consumo de alcohol y x el precio, los individuos cuyas observaciones se encuentran relativamente por arriba de la nube son aquellos consumidores cuyas preferencias por el alcohol son elevadas. En base a un modelo de regresión por cuantiles, estos autores encuentran que la elasticidad precio es mucho menor en los cuantiles superiores, lo que implica que, controlando por los factores observables, los consumidores con mayor grado de preferencia por el alcohol reaccionan relativamente poco frente a variaciones en el precio. Esto tiene consecuencias no triviales cuando el objetivo consiste en fijar impuestos, ya que si el análisis se hubiese basado en un modelo de regresión

⁴ Op.cit. pp.131, traducción del autor del original en inglés.

estándar, la elasticidad estimada para la media es notoriamente superior a la esperable para el grupo de focalización que se interesa gravar: los de alta preferencia por el alcohol.

Implícita en este análisis, y en la interpretación de los coeficientes del modelo de regresión por cuantiles, está la idea de que tiene sentido pensar que alterando x los individuos se mantienen en el mismo cuantil condicional. En este ejemplo de consumo de alcohol, el modelo tiene sentido si podemos pensar que los no factores no observables tienen la interpretación de “preferencia por el alcohol” y que la misma es una característica relativamente estable para un grupo de individuos, de modo que tenga sentido intuitivo hablar de “los individuos del cuantil superior de la distribución del consumo de alcohol condicional en los factores observables”. A favor, cabe decir que este supuesto de “*ceteris paribus*” está también implícito en cualquier modelo de regresión estándar, si es que tiene sentido práctico pensar en movimientos para un individuo de características no observables promedio.

Consideremos un segundo ejemplo, esta vez tomado de la ecología, basado en el trabajo de Cade, Terrell y Schroeder (1999). La “ley del mínimo” de Liebig enuncia que el crecimiento de una planta continúa siempre y cuando todos los factores requeridos (agua, luz, nitrógeno, etc.) estén presentes y que, consecuentemente, cuando uno de estos factores es limitante, el crecimiento se detiene. Aumentar la cantidad de este factor limitante permite restaurar el crecimiento hasta que ese u otro factor actúe nuevamente como limitante. Consecuentemente, un *factor limitante* es aquel relativamente menos disponible entre todos los factores que afectan el crecimiento, supervivencia o reproducción de un organismo. Cade et. al. (1999) utilizan métodos de regresión por cuantiles para estudiar la producción de bellotas en un bosque, utilizando como variable explicada la densidad de producción (número de bellotas anuales por hectárea) y como variable explicativa un índice de calidad del bosque de robles en donde se realiza el estudio, de modo que a mayor valor de este índice, mejor calidad. Los datos provienen de 43 parcelas experimentales en el estado de Missouri, EEUU, para los años 1989 a 1993; el trabajo citado contiene una descripción detallada de la forma en la cual fueron construidas estas variables. La relación entre densidad de producción y calidad del bosque es claramente

positiva y sigue un patrón similar al de la Figura 3: la calidad del bosque impacta sobre la densidad, pero lo hace en forma más intensa (mayor pendiente) en la parte superior de la distribución condicional. Cade et al. (1999) infieren que esto se debe a que justamente allí es cuando actúan menos intensamente los factores limitantes no observables y, consecuentemente, el verdadero efecto limitante de la calidad del bosque se aprecia mejor en los cuantiles superiores de la distribución, en donde los otros factores tienen un peso menor, y no en el centro, como se hubiese deducido de un modelo de regresión estándar, en donde actúan otros factores limitantes. Más concretamente, y siguiendo a Cade et al. (1999), la relación entre densidad (y) y calidad del bosque (x) podría describirse de la siguiente forma:

$$y = \beta_0 + \beta_1 x + g(x) u$$

en donde u representa a los factores no observables, y $g(\cdot)$ es una función arbitraria. Lo interesante de esta especificación es el término de interacción $g(x)u$ que hace que los no observables u interactúen con x . Por ejemplo, si $g(x) = \delta x u$, en donde δ es una constante positiva,

$$\partial y / \partial x = \beta_1 + \delta u$$

y, consecuentemente, el efecto de x sobre y es mayor cuanto mayor es u . Cade et al. (1999) indican que posiblemente la temperatura sea un componente no medido que es muy relevante y que por lo tanto es capturado por u . Lo interesante de la tecnología de regresión por cuantiles, en este ejemplo, es que es posible hacer inferencias acerca de la heterogeneidad en la forma en la que x afecta a y sin observar u y, sin especificar la función $g(\cdot)$. Ciertamente, si u y x no interactuasen (por ejemplo, si $g(x)$ es una constante no nula), el efecto hubiese sido homogéneo y el efecto de x sobre y podría medirse adecuadamente en el centro de la distribución condicional, como lo hace el modelo de regresión estándar.

5. UNA APLICACIÓN: RENDIMIENTO DE ESTUDIANTES UNIVERSITARIOS

Este ejemplo se basa en un trabajo de Di Gresia et al. (2005) sobre la performance de los estudiantes universitarios en las escuelas públicas⁵. Existe una extensa literatura que intenta estudiar como las características de un individuo, tales como la educación de sus padres o su edad, inciden en su performance educativa. Aunque informativos sobre muchos aspectos del proceso educativo, los modelos disponibles se encuentran todavía lejos de alcanzar este objetivo en forma precisa, lo cual suele reflejarse en una baja performance en términos de bondad del ajuste. Esto significa que aún luego de condicionar en muchos aspectos observables que determinan el rendimiento, los individuos todavía difieren sustancialmente a causa de estos factores no observables.

En consecuencia, la forma correcta de pensar en el efecto de un factor observable sobre el rendimiento de un estudiante es considerar como los cambios en el factor específico afectan la distribución condicional de rendimientos. Por ejemplo, si se considera el efecto de la educación del padre, la distribución de rendimientos condicional en los factores observables, incluyendo la educación del padre, aun presenta una importante variabilidad debido al rol no trivial que cumplen los inobservables, de modo que, aun dentro de un grupo de individuos con las mismas características observadas, encontraremos estudiantes con mala, regular o buena performance. Resulta natural esperar que toda la distribución condicional de rendimientos se desplace a la derecha cuando, a igualdad de otros factores, se consideran hijos con padres con mayor educación. En el caso extremo donde la educación adicional del padre desplace el total de la distribución condicional a la derecha en la misma proporción, el efecto de considerar individuos con padre con mayor educación sobre el rendimiento medio capturaría todo lo que se necesita conocer. En ese contexto, y bajo algunos supuestos simplificadores, un modelo de regresión estándar podría brindar la respuesta deseada: el coeficiente de la educación del padre en una

⁵ Agradezco a los coautores de este trabajo permitirme utilizar los resultados de nuestras investigaciones.

regresión lineal capturaría el efecto sobre el rendimiento esperado y, bajo estas circunstancias, sobre el rendimiento en general. Esta situación se daría en caso de que la educación del padre sea completamente independiente de los no observables, de modo que, los movimientos en la educación del padre impliquen desplazamientos de ubicación puros en la distribución condicional de rendimientos. Pero dado el rol no trivial que juegan los inobservables, vale considerar el caso en que los movimientos de la educación del padre interactúan de algún modo con los factores no incluidos en el modelo en una forma poco obvia. Como ejemplo extremo, podría conjeturarse que la educación del padre tuviera un rol importante en un hijo con una inclinación natural al estudio, y un efecto moderado sobre los hijos con menor interés. En cualquier caso, ‘el efecto medio’ de la educación de padre es positivo pero no representa a nadie en la población: sobreestima el efecto en el caso de hijos de baja propensión a estudiar, y subestima la situación de hijos mejor automotivados.

A fines de explorar la relación entre el rendimiento educativo y sus determinantes, consideremos el siguiente modelo para el rendimiento educativo, R ,

$$R_i = f(x_i, u_i), \quad i=1, \dots, n,$$

donde R_i es alguna medida del rendimiento del estudiante i , x es un vector de K determinantes observables del rendimiento y u es una variable aleatoria que representa determinantes no observables. Formulaciones como esta son consistentes con el enfoque de la *función de producción*, donde el rendimiento se ve como el producto de utilizar x y u como insumos en un proceso de “producción” del que se obtiene un resultado educativo. Ver Hanushek (1986) para esta interpretación. Bajo el supuesto usual de linealidad $R = x'\beta + u$, y si $E(u/x) = 0$, β puede ser estimado consistentemente utilizando el método de mínimos cuadrados ordinarios (MCO) basado en una muestra (y_i, x_i) , $i=1, \dots, n$ de n individuos.

Como se mencionase anteriormente, el rol de los inobservables está lejos de ser menor. Por ejemplo, en la mayoría de las estimaciones por MCO el

coeficiente R^2 no supera el 15% utilizando muestras de un tamaño considerable y un amplio conjunto de factores observables, de modo que el 85% de la variabilidad en el rendimiento educativo se debe a los componentes inobservables. Esto implica que aún luego de controlar por los observables, sigue existiendo una heterogeneidad relevante entre los individuos, lo que hace dudar de la representatividad de la esperanza condicional como resumen de la distribución condicional de rendimientos. Esto sugiere la conveniencia de acudir a un modelo de regresión por cuantiles a fines de explorar los efectos de los observables sobre *toda* la distribución condicional. Existen algunos antecedentes de efectos heterogéneos de los determinantes de la performance educativa. Eide y Showalter (1998) exploran el posible efecto heterogéneo de la calidad de la escuela sobre la performance. Levin (2001) analiza el efecto del tamaño de la clase.

Los datos utilizados en este ejemplo provienen del Censo de Estudiantes de universidades públicas realizado en 1994. Es importante reconocer el carácter pedagógico de este ejemplo, en donde varias cuestiones muy relevantes no serán tratadas. Di Gresia et al. (2005) contiene un análisis exhaustivo así como una descripción detallada de varias dificultades metodológicas implícitas en el análisis, y referiremos a este trabajo para el lector interesado en los detalles. Como medida del rendimiento educativo se utiliza el número de materias aprobadas desde el comienzo de la carrera dividido por la cantidad de años desde que ingresó a la universidad. Los factores explicativos observables incluyen: el sexo (1 si hombre, 0 si no), edad (en años), educ.p (educación del padre, en años), educ.m (educación de la madre en años), trab (1 si tiene trabajo, 0 si no), h.estd (horas de estudio semanales), sec.pri (1 si asistió a escuela secundaria privada, 0 si a pública). A los efectos de limitar el ejemplo, el mismo se concentra en la muestra de estudiantes de ciencias económicas de la Universidad de Buenos Aires ingresados en el año 1990. Estas carreras son de al menos 4 años, de modo que los mejores alumnos de esta cohorte están, al momento del censo, en cuarto año.

La Tabla 2 presenta los resultados de la estimación. Cada columna presenta estimaciones para cuantiles del 0.1 al 0.9. Para su comparación, las estima-

ciones por el método de mínimos cuadrados ordinarios (MCO) se incluyen en la última columna. La Figura 10 muestra gráficamente estos resultados, y se interpreta en forma análoga a la Figura 9 de la sección anterior.

El supuesto de homogeneidad se rechaza fuertemente en varios casos. El caso homogéneo correspondería a estimaciones de cuantiles representadas por una línea horizontal constante para cada variable explicativa. Este no es el caso para muchas de las variables. Una situación particularmente relevante corresponde al efecto de las horas de estudio. La estimación “media” por MCO implica que el incremento de una hora de estudio aumenta el rendimiento en 0.025. Los resultados del modelo de cuantiles muestran que este efecto es marcadamente decreciente a lo largo de los cuantiles de la distribución condicional de rendimiento. En particular, para individuos en los cuantiles inferior-

Tabla 2. Determinantes del rendimiento. Universidad de Buenos Aires

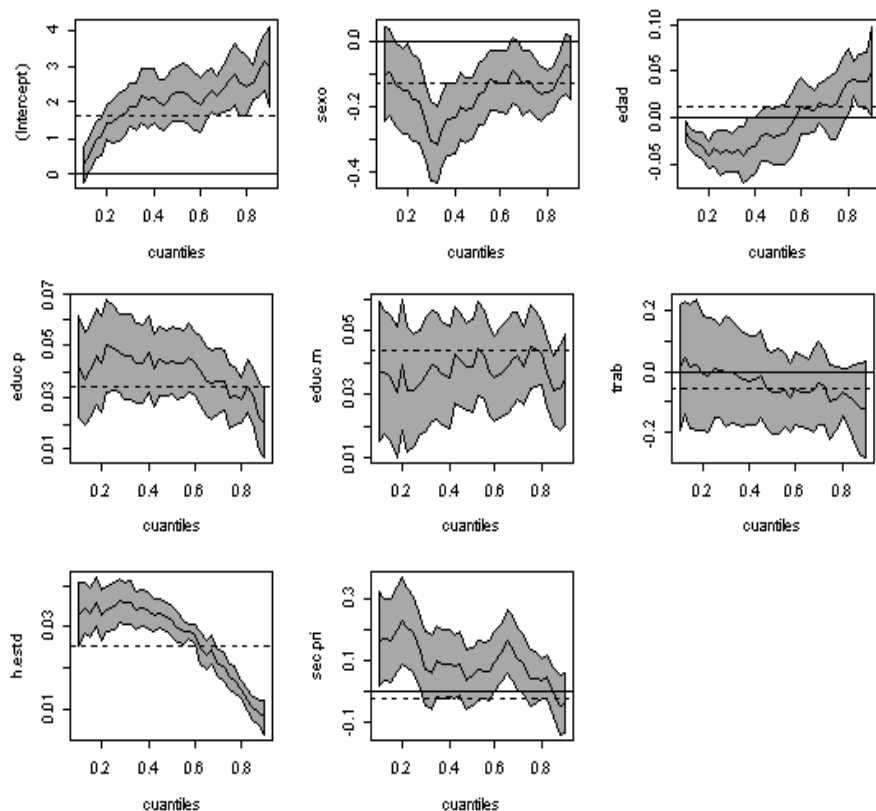
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	MCO
Intercepto	0.277 (0.909)	1.444 (4.717)	1.913 (5.266)	2.155 (4.645)	2.258 (4.926)	1.930 (4.299)	2.372 (5.963)	2.446 (4.873)	2.985 (4.425)	1.727 (6.751)
sexo	-0.101 (-1.154)	-0.148 (-1.765)	-0.307 (-4.196)	-0.237 (-3.691)	-0.161 (-2.784)	-0.125 (-2.202)	-0.130 (-2.198)	-0.158 (-3.953)	-0.082 (-1.414)	-0.143 (-2.940)
edad	-0.015 (-2.004)	-0.041 (-4.674)	-0.038 (-2.997)	-0.031 (-1.662)	-0.019 (-1.028)	0.012 (0.677)	0.014 (0.912)	0.039 (1.857)	0.051 (1.778)	0.007 (0.783)
educ.p	0.042 (3.457)	0.043 (3.766)	0.046 (4.594)	0.047 (5.274)	0.044 (5.531)	0.043 (5.732)	0.036 (4.687)	0.029 (5.564)	0.020 (2.552)	0.035 (5.095)
educ.m	0.037 (2.786)	0.039 (3.142)	0.036 (3.359)	0.035 (3.675)	0.039 (4.492)	0.035 (4.357)	0.042 (5.025)	0.043 (7.175)	0.035 (4.036)	0.042 (5.789)
trab	0.013 (0.107)	-0.005 (-0.044)	0.000 (0.003)	-0.032 (-0.353)	-0.073 (-0.883)	-0.054 (-0.756)	-0.037 (-0.450)	-0.067 (-1.470)	-0.123 (-1.294)	-0.045 (-0.664)
h.estd	0.033 (7.151)	0.033 (8.896)	0.036 (10.994)	0.034 (12.512)	0.031 (13.903)	0.028 (19.539)	0.021 (10.347)	0.015 (10.392)	0.008 (3.167)	0.025 (11.417)
sec.pri	0.166 (1.835)	0.226 (2.726)	0.071 (0.975)	0.092 (1.447)	0.050 (0.878)	0.094 (1.647)	0.105 (1.731)	0.039 (0.986)	-0.036 (-0.630)	0.106 (2.178)

Notas: cada columna presenta estimaciones de los efectos de las variables explicativas para distintos cuantiles condicionales del rendimiento. La última columna presenta las estimaciones por MCO. Debajo de cada estimación puntual se presentan entre paréntesis los estadísticos «t».

res el efecto es 0.033, y decrece en forma monótona a un tercio (0.012) en el cuantil superior estimado. Las estimaciones medias son claramente no representativas del efecto de las horas de estudio sobre el rendimiento, ya que si bien el efecto es positivo para todos los cuantiles de la distribución condicional de rendimientos, el mismo es más fuerte en la cola inferior. Es interesante notar que, consecuentemente, al aumentar las horas de estudio, la distribución de rendimientos se vuelve más homogénea.

Otro caso interesante es el efecto de la edad sobre el rendimiento. La estimación por MCO no es significativa, indicando que la edad no altera el rendimiento esperado, pero los resultados por cuantiles sugieren un efecto significa-

Figura 10



tivo que se mueve monotónicamente de negativo a positivo entre cuantiles. Esto es consistente con un fuerte *efecto escala* en el rendimiento: los grupos de mayor edad tienen un rendimiento esperado similar al del grupo más joven, pero con mayor dispersión. Sin duda es erróneo concluir que la edad no tiene efecto en el rendimiento, en forma similar a lo discutido en referencia a la Figura 4 de la sección 2.

El efecto de la educación de la madre es interesante ya que la estimación por MCO es positiva y significativa (0.042), pero las estimaciones por cuantiles son también positivas y significativas y relativamente constantes, con valores cercanos al del efecto medio. Entonces, en este caso el efecto sobre el promedio provee un resumen representativo del efecto de la educación de la madre sobre toda la distribución condicional: parece provocar un desplazamiento puro en la distribución de rendimientos.

El efecto de haber asistido a una escuela secundaria privada sigue un patrón interesante. Comienza siendo significativo y positivo y decrece en forma monótona hasta dejar de ser significativo en los cuantiles por encima del 0.4, lo que sugiere que la educación previa privada es un factor relevante solo para aquellos estudiantes en la parte inferior de la distribución.

6. AVANCES Y PERSPECTIVAS EN REGRESIÓN POR CUANTILES

La idea de ajustar distintas curvas para distintos niveles de la distribución condicional de una variable posiblemente haya estado implícita en varios procedimientos estadísticos simples, como, por ejemplo, dibujar para una variable y y varios “diagramas de caja” para los distintos niveles de otra variable x y, ciertamente, en forma explícita, tal como lo hacen notar Mosteller y Tukey (1977) en su texto clásico. Pero no es hasta la aparición del trabajo de Roger Koenker y Gilbert Bassett (1978) que este problema recibió un tratamiento adecuado, en el sentido de explicitar claramente el modelo probabilístico subyacente, una estrategia de estimación concreta, y una teoría asintótica sólida que sentara la base para la implementación práctica de estos métodos. A partir de este trabajo seminal, los primeros avances estuvieron dirigidos a consolidar

este marco teórico (Bassett y Koenker (1982, 1986)) y a establecer algunos vínculos con métodos econométricos estándar, tales como el nexo entre la estimación por cuantiles y el problema de la evaluación de heterocedasticidad (Koenker y Bassett, 1982).

Un fuerte quiebre en la tendencia de la literatura sobre regresión por cuantiles se da hacia mediados de la década del noventa, con la aparición de los trabajos de Chamberlain (1994) y Buchinsky (1994), quienes utilizan este tipo de métodos para estudiar la relación entre los salarios individuales y sus determinantes. En el marco de la muy extensa literatura sobre retornos a la educación, estos trabajos establecen la utilidad conceptual de cuantificar el efecto de la educación no solo en la esperanza condicional, sino en distintos niveles de la distribución condicional de los salarios (en logaritmos). Los resultados de estos trabajos sugieren que el retorno en la media que surge de los modelos log-lineales estándar estimados por el método de mínimos cuadrados ordinarios, es un resumen un tanto impreciso de una realidad más heterogénea, en donde los retornos a la educación crecen monótonamente en los distintos cuantiles de la distribución condicional del ingreso, de modo que los individuos en los cuantiles condicionales de la distribución de los salarios experimentan retornos más altos.

La aparición de estos trabajos le dio un fuerte impulso a la literatura de regresión por cuantiles, tanto a nivel teórico como empírico. En el primero de los aspectos, una proporción sustancial de la producción intelectual en este área estuvo destinada a desarrollar métodos de cuantiles para varios problemas estándar en econometría, dicho de otro modo, a encontrar análogos “para cuantiles” de varios métodos de uso común. Una primera aplicación relevante se refiere a la conveniencia de utilizar métodos de cuantiles para el problema de estimación de modelos lineales con datos censurados. Bajo el supuesto, ciertamente restrictivo, de normalidad y homogeneidad, los modelos *tobit* estimados por el método de máxima verosimilitud proveen una solución simple y elegante a este problema. En un trabajo de gran relevancia, Powell (1986) muestra que el problema de censura puede ser adecuadamente tratado sin recurrir a supuestos distributivos ni de homogeneidad a través del método de

regresión por cuantiles. A los fines de brindar una intuición de la utilidad de estos métodos para el problema de censura, pensemos que así como la mediana muestral no se altera si modificamos el valor de una observación originariamente por debajo de la mediana, análogamente, la recta de regresión por cuantiles correspondiente al cuantil 0.5 tampoco se altera si alteramos la posición relativa de un punto originariamente por arriba de la recta. Consecuentemente, si entendemos que una observación censurada es en realidad una observación que originariamente estaba por debajo de la recta y que luego de operar el mecanismo de censura es alterada, la recta estimada no debería alterarse y, consecuentemente, la estimación de la mediana condicional debería proveer una estrategia de estimación consistente aun bajo la presencia de un mecanismo que altere, si bien no en forma drástica, la posición relativa de los puntos. Esta idea es posteriormente extendida en varias direcciones por Buchinsky y Hahn (1998). Fitzenberger (1997) presenta una útil revisión de la aplicabilidad de las regresiones por cuantiles para lidiar con datos censurados. Algunas referencias recientes son Portnoy (2003), Honore, Khan y Powell (2002) y Chernozhukov y Hong (2002). Los métodos de regresión por cuantiles resultan particularmente útiles en contextos de alta censura tales como los problemas de duración, ver Koenker y Biliás (2001) y Koenker y Geiling (2001) para aplicaciones en esta literatura.

El supuesto de linealidad, mantenido a lo largo de este trabajo puede ser relajado en varias dimensiones. Koenker, Ng y Portnoy (1994) proponen modelos de regresión no lineal por cuantiles, basados en métodos de *smoothing splines*, y una versión local, en el espíritu de los modelos de regresión no-paramétrica, es estudiada por Yu y Jones (1998). Buchinsky (1995) y Machado y Mata (2000) muestran que los modelos del tipo Box-Cox, que permiten acomodar una variedad de especificaciones no-lineales, pueden ser fácilmente implementados en un contexto de regresión por cuantiles. Esto se deriva de la propiedad de *equivariancia* de los cuantiles condicionales ante transformaciones monótonas, es decir:

$$h(Q_{y/x}(\tau)) = Q_{h(y)/x}(\tau)$$

para cualquier función monótona $h(\cdot)$, propiedad que, trivialmente, no posee la esperanza condicional, a menos que $h(\cdot)$ sea una función lineal. Lee (2003) estudia modelos parcialmente lineales de regresión por cuantiles.

Kordas (2002) extiende los resultados de Horowitz (1992) y Manski (1985), proponiendo un marco de análisis para datos binarios desde una perspectiva de cuantiles condicionales. Abadie, Angrist e Imbens (2002) exploran el vínculo entre regresión por cuantiles y el problema de evaluación de programas.

Recientemente la literatura de regresiones por cuantiles para series temporales ha recibido considerable atención. Koenker y Xiao (2001) extienden la noción de modelos autorregresivos para cuantiles condicionales (QAR). Koenker y Zhao (1994) estudian un modelo ARCH, también desde una perspectiva de cuantiles. La literatura sobre métodos de regresión por cuantiles para paneles es todavía muy incipiente, ver Koenker (2004).

Como se señala en el Apéndice, la estimación de estos modelos no es computacionalmente trivial y la mayoría de las veces involucra procesos de optimización no diferenciables. Koenker y D'Orey (1987) proponen un algoritmo útil para estimar modelos lineales y no lineales. Koenker y Park (1996) proponen un algoritmo basado en métodos de punto interior. Portnoy y Koenker (1997) presentan una útil comparación de la performance computacional de los métodos basados en mínimos valores absolutos y en mínimos cuadrados. La teoría asintótica ha sido refinada en varias direcciones. Koenker (1994) presenta una visión detallada de varios métodos para construir intervalos de confianza asintóticos. De particular relevancia es el nexo entre la estimación de modelos de regresión por cuantiles y la teoría de rangos de Hajek y Sidak (1967), lo cual abrió una fructífera vía para el análisis asintótico. Fitzenberger (1997) explora el uso de técnicas de bootstrap para regresión por cuantiles. Koenker y Machado (1999) proponen una medida de bondad de ajuste similar al R^2 de un modelo de regresión estándar. Koenker y Xiao (2002) proveen un marco analítico para realizar inferencia sobre el proceso empírico $\hat{\beta}(\tau)$.

En lo que respecta a las aplicaciones, como se comentase anteriormente, la aparición de los trabajos antes mencionados de Chamberlain (1994) y Buchinsky (1994) le da un fuerte impulso a la literatura de regresión por

cuantiles para estudiar los determinantes de los ingresos individuales, comparando una larga serie de estudios en líneas de investigación similares, en las cuales se enmarcan los trabajos de Arias, Hallock y Sosa Escudero (2001), Machado y Mata (2001), García, Hernández y López Nicolás (2001) entre otros.

El trabajo de Manning et al. (1995), discutido en la sección 4, también pertenece a este grupo de trabajos pioneros. Progresivamente, varias áreas de la economía aplicada han sido estudiadas desde una perspectiva de cuantiles condicionales, tales como la estimación de modelos de frontera estocástica (Bernini et al., 2004), el problema de convergencia en el crecimiento (Canarella y Pollard, 2004), el análisis del tamaño inicial de firmas (Machado 2000 o la muy debatida cuestión del efecto del tamaño de clase sobre la performance de los estudiantes (Levin, 2001). Para aplicaciones en finanzas, ver Park (2002) o Bassett et al. (2002). En esta misma línea resulta relevante resaltar el trabajo de Engle y Manganelli (2004), que utilizan métodos de regresión por cuantiles para estudiar cuestiones de valor en riesgo.

Los modelos de regresión por cuantiles han resultado de particular utilidad para estudiar el efecto de variables observables sobre algunas características de la distribución del ingreso, tales como la desigualdad o la pobreza, tales como los estudios recientes de Machado y Mata (2000), Pereira (2000), y Arias y Sosa Escudero (2000). Belluzo (2004) es una muy interesante aplicación al problema de evaluación de proyectos. Gasparini y Sosa Escudero (2005) utilizan estos métodos para estimar modelos hedónicos de alquileres de viviendas e imputar las rentas implícitas en la posesión de inmuebles.

Si bien tienen su origen en la econometría, las técnicas de regresión por cuantiles se han implementado exitosamente en varias disciplinas científicas, tal como se ilustra en el trabajo de Cade et al. (1999) discutido en la sección 4. Cade y Noon (2003) contiene una interesante lista de aplicaciones recientes de los métodos de cuantiles en ecología y biología.

La mayoría de los estudios descriptos en esta sección parecen tener en común el interés explícito en estudiar el efecto que tienen uno o varios factores sobre un resultado *distributivo* y que por lo tanto, la focalización en la

esperanza condicional implícita en los modelos de regresión estándar puede resultar poco relevante cuando no es adecuada. Otra característica distintiva tiene que ver con el interés en focalizar en algún aspecto en particular de la distribución condicional (la cola superior, como en el trabajo de Cade, o la inferior, como en los trabajos de finanzas o en varios problemas de datos censurados). Como señala Matzkin (2004), los modelos de regresión por cuantiles son esencialmente modelos de interacción entre factores observables y no observables.

Existen varios trabajos que revisan, con variado grado de detalle, distintos aspectos de la literatura sobre regresión por cuantiles. El carácter semiparamétrico y la naturaleza marcadamente no estándar de los métodos de estimación muchas veces introducen complejidades formales que pueden constituir una barrera a la entrada, aún para no especialistas con una sólida formación en econometría. Koenker y Hallock (2001) presentan una revisión informal y general. Buchinsky (1998) enfatiza el uso de estos métodos para el estudio de cuestiones relacionadas con el mercado laboral. Cade y Noon (2003) es una introducción muy fácil de leer, dirigida a biólogos, y Koenker (2001) es una excelente referencia desde una perspectiva histórica. El texto de Koenker (2005) contiene un tratamiento detallado y exhaustivo sobre el tema y parece ser la referencia definitiva sobre el tema. Llamativamente la literatura de regresión por cuantiles todavía no ha encontrado el espacio que se merece en los libros de texto intermedio o avanzado y, a la fecha de elaboración de este trabajo, aquellos que hacen referencia explícita al tema son muy pocos, siendo Peracchi (2001), Pagan y Ullah (1997) y el reciente texto de Cameron y Trivedi (2005) algunas excepciones. Deaton (1997) ilustra el uso de estos métodos para problemas de ingresos y curvas de Engel.

En términos de software, *Stata* tiene varias rutinas para modelos simples. Las rutinas disponibles para *Splus* y *R*, en particular el paquete *quantreg*⁶, desarrollado por Roger Koenker, son herramientas más sofisticadas, en un entorno altamente programable al cual ha gravitado recientemente una parte considerable de la estadística.

⁶ Ver <http://cran.r-project.org/src/contrib/Descriptions/quantreg.html>

7. CODA: UNA DIGRESIÓN SEMÁNTICA

Una dificultad inesperada en la elaboración de este artículo consistió en buscar una traducción apropiada para el término *quantile regression*, nombre con el que se hace referencia a esta familia de modelos en el idioma inglés. Mi reacción instintiva consistió en buscar trabajos en español que utilizaran estas técnicas, pero a la fecha, lamentablemente, los dedos de una mano sobran para contar los mismos. Mi búsqueda produjo solo dos referencias publicadas. Zarate (2002) utiliza el vocablo “regresión cuantílica” mientras que Zamudio Carrillo (2001) prefiere “regresión cuantil”. Originariamente, Koenker y Bassett (1978) utilizaron “regression quantiles” para hacer referencia al vector de coeficientes $\hat{\beta}(\tau)$ pero en la década del 90 el término fue reemplazado por “quantile regression”, haciendo referencia al proceso general de estimar por cuantiles, no necesariamente restringiéndose a modelos lineales. Muy humildemente y sin mayor justificación etimológica ni semántica, propongo *regresión por cuantiles*, que se basa en vocablos de uso común en nuestro idioma, a la vez de ser consistente con la idea de ajustar modelos de regresión para distintos lugares (cuantiles) de la distribución condicional de una variable⁷.

REFERENCIAS

- Abadie, A., Angrist, J. e Imbens, G., 2001, Instrumental Variables Estimates of subsidized training on the quantile of trainee earnings, *Econometrica*, 70, 91-117.
- Abrevaya, J., 2001, The effects of demographics and maternal behavior on the distribution of birth outcomes, *Empirical Economics*, 26, 247-57.

⁷ www.freetranslation.com propone ‘retroceso de quantile’ y www.dictionary.com, ‘regresión del quantile’. www.foreignword.com, muy prudentemente, se niega a ofrecer una traducción. La Economía parece una disciplina mucho más exacta y promisoría a la luz de la ciencia de la traducción.

- Arias, O., Hallock, K. y Sosa Escudero, W., 2001, Individual heterogeneity in the returns to schooling: instrumental variables Quantile Regression using twins data, *Empirical Economics*, 26, 7-40.
- Arias, O. y Sosa Escudero, W., 2004, Subjective and objective poverty in Bolivia: a binary quantile regression approach, mimeo, The World Bank.
- Basset, G. y Chen, Hsiu-Lang 2001, Quantile style: return-based attribution using regression quantiles, *Empirical Economics*, 26, 293-305.
- Bassett, G. y Koenker, R., 1982, An empirical quantile function for linear models with iid errors, *Journal of the American Statistical Association*, 77, 407-415.
- Bassett, G. y Koenker, R., 1986, Strong consistency of regression quantiles and related empirical processes, *Econometric Theory*, 2, 191-201.
- Belluzzo, W., 2004, Semiparametric approaches to welfare evaluations in binary response models, *Journal of Business and Economic Statistics*, 22, 322-330.
- Bernini, C., Freo, M. y Gardini, A., 2004, Quantile estimation of frontier production function, *Empirical Economics*, 29, 373-81
- Buchinsky, M., 1994, Changes in US wage structure 1963-1987: an application of quantile regression, *Econometrica*, 62, 405-58.
- Buchinsky, M., 1995, Quantile regression, Box-Cox transformation model and the US wage structure, 1963-1987, *Journal of Econometrics*, 65, 109-154.
- Buchinsky, M., 1997, The dynamics of changes in the female wagw distribution in the USA: a quantile regression approach, *Journal of Applied Econometrics*, 13, 1-30.
- Buchinsky, M., 1998, Recent advances in quantile regression models: a practical guide for empirical research, *Journal of Human Resources*, 33, 88-126.
- Buchinsky, M. y Hahn, J., 1998, An alternative estimator for the censored quantile regression model, *Econometrica*, 66, 653-71.
- Cade, B. y Noon, B., 2003, A gentle introduction to quantile regression for ecologists, *Frontiers in Ecology and the Environment*, 1, 412-420.
- Cade, B., Terrell, J. y Schroeder, R., 1999, Estimating effects of limiting factors with regression quantiles, *Ecology*, 80, 311-323.
- Cameron, C. y Trivedi, P., 2005, *Microeconometrics*, Cambridge University Press, Cambridge.

- Canarella, G. y Pollard, S., 2004, Parameter heterogeneity in the neoclassical growth model: a quantile regression approach, *Journal of Development Economics*, 29, 1-31.
- Chamberlain, G., 1994, Quantile regression, censoring and the structure of wages, en *Advances in Econometrics*, Sims, C. (ed), 171-209, Elsevier, New York.
- Chernozhukov, V. y L. Umantsev, 2001, Conditional Value-at-Risk: aspects of modeling and estimation, *Empirical Economics*, 26, 271-92.
- Chernozhukov, V. y Hong, H., 2002, Three-step censored quantile regression and extramarital affairs, *Journal of the American Statistical Association*, 97, 459, 872-82.
- Deaton, A., 1997, The analysis of household surveys, *Baltimore: John Hopkins*.
- Di Gresia, L, Fazio, M., Porto, A., Ripani, L. y Sosa Escudero, W., 2005, Rendimiento y Productividad de los Estudiantes. El Caso de las Universidades Públicas Argentinas, en Porto, A. (editor) *Economía de la Educación Universitaria: Argentina-Brasil-Perú*, Editorial de la Universidad Nacional de La Plata, La Plata.
- Eide, E. y Showalter, M., 1998, The effect of school quality on student performance: a quantile regression approach, *Economics Letters*, 58, 345-50.
- Engle, R. y Manganelli, S. 1999, CaViar: conditional autoregressive value at risk by regression quantiles, *University of California San Diego*, Department of Economics, Working Paper 99/20, October.
- Fitzenberger, B., 1997, A guide to censored quantile regressions, en Maddala, G.S. y Rao, C.R., eds. *Handbook of Statistics*, Vol. 15, Cap. 15, Elsevier.
- Fitzenberger, B., 1998, The Moving Blocks Bootstrap and Robust Inference in Linear Least Squares and Quantile Regressions, *Journal of Econometrics*, 82, 235-287.
- Garcia, J., Hernandez, P. y Lopez-Nicolas, A., 2001, How wide is the gap? An investigation of gender wage differences using quantile regression, *Empirical Economics*, 26, 149-67.
- Gasparini, L. y Sosa Escudero, W., 2003, Implicit Rents from Own-housing and Income Distribution: Econometric Estimates for Greater Buenos Aires, *Journal of Income Distribution*, 2.

- Gosling, A., Machin, S. y Meghir, C., 2000, The changing distribution of male wages in the U.K., *Review of Economic Studies*, 67, 635-66.
- Gutenbrunner, C., y Jureckova, J., 1992, Regression quantile and regression rank score process in the linear model and derived statistics, *Annals of Statistics*, 20, 305-330.
- Hajek, J. y Sidak, Z., 1967, *Theory of Rank Tests*, Academia: Prague.
- He, X., 1997, Quantile curves without crossing, *American Statistician*, 51, 186-192.
- Honore, B., Khan, S. y Powell, J., 2002, Quantile regression under random censoring, *Journal of Econometrics*, 109, 67-105.
- Horowitz, J., 1992, A smoothed maximum score estimator for the binary regression model, *Econometrica*, 60, 505-531.
- Karr, A., 1993, *Probability*, Springer-Verlag, New York.
- Koenker, R., 1994, Confidence intervals for regression quantiles, en Mandl, P. y Huskova, M. (eds.), *Asymptotic Statistics: Proceedings of the 5th Prague Symposium*, Physica-Verlag, Heidelberg, 349-59.
- Koenker, R., 2001, Galton, Edgeworth, Frisch, and prospects for quantile regression in econometrics, *Journal of Econometrics*, 95, 347-74.
- Koenker, R., 2005, *Quantile Regression*, Cambridge University Press, Cambridge.
- Koenker, R., 2004, Quantile regression for longitudinal data, *Journal of Multivariate Analysis*, 91, 74-89.
- Koenker, R. y Bassett, G., 1978, Regression quantiles, *Econometrica*, 46(1), 33-50.
- Koenker, R. y Bassett, G., 1982, Robust tests for heteroscedasticity based on regression quantiles, *Econometrica*, 50, 43-61.
- Koenker, R. y Biliias, Y., 2001, Quantile regression for duration data: a reappraisal of the Pennsylvania Reemployment bonus experiments, *Empirical Economics*, March, 26:1, pp.199-220.
- Koenker, R. y D'Orey, V., 1987, Computing regression quantiles, *Applied Statistics*, 36, 383-393.
- Koenker, R. y Geiling, O., 2001, Reappraising medfly longevity: a quantile regression survival analysis, *Journal of the American Statistical Association*, 96, 458-468.

- Koenker, R. y Machado, J., 1999, Goodness of fit and related inference processes for quantile regression, *Journal of the American Statistical Association*, 94, 1296-1310.
- Koenker, R. y Park, B., 1996, An interior point algorithm for non-linear quantile regression, *Journal of Econometrics*, 71, 265-283.
- Koenker, R. y Schorfeide, F., 1994, Quantile spline models for global temperature change, *Climatic Change*, 28, 395-404.
- Koenker, R., Ng, P. y Portnoy, S., 1994, Quantile smoothing splines, *Biometrika*, 81, 673-680.
- Koenker, R. y Xiao, Z, 2002, Inference on the Quantile Regression Process, *Econometrica*, 70(4), 1583-1612.
- Koenker, R. y Zhao, Q., 1996, Conditional Quantile Estimation and Inference for ARCH Models, *Econometric Theory*, 12, 793-813.
- Kordas, G., 2002, Smoothed binary regression quantiles, *Journal of Applied Econometrics*, en prensa.
- Lee, S., Efficient semiparametric estimation of a partially linear quantile regression model, *Econometric Theory*, 19, 1-31.
- Levin, J., 2001, For whom the reductions count: A quantile regression analysis of class size and peer effects on scholastic achievement, *Empirical Economics*, 26, 221-246.
- Mata, J. y Machado, J., 1996, Firm start-up size: A conditional quantile approach, *European Economic Review*, 40, 1305-1323.
- Machado, J. y Mata, J., 2000, Box-Cox quantile regression and the distribution of firm sizes, *Journal of Applied Econometrics*, 15, 253-274.
- Manski, C., 1985, Semiparametric analysis of discrete response: asymptotic properties of the maximum score estimator, *Journal of Econometrics*, 32, 65-108.
- Manski, C., 1991, Regression, *Journal of Economic Literature*, XXIX, 34-50.
- Manning, W., Blumberg, L. y Moulton, L., 1995, The demand for alcohol: the differential response to price, *Journal of Health Economics*, 14, 123-148.
- Martins, P. y Pereira, P., 2004. Does education reduce wage inequality? Quantile regression evidence from 16 countries, *Labour Economics*, 11, 355-371.

- Mata, J. y Machado, J., 1996, Firm start-up size: a conditional quantile approach, *European Economic Review*, 40, 1305-1323.
- Matzkin, R., 2003, Nonparametric Estimation of Nonadditive Random Functions, *Econometrica*, 71, 1339-1375.
- Mosteller, F. y Tukey, J., 1977, *Data Analysis and Regression: A Second Course in Statistics*, Addison-Wesley, Massachusetts.
- Pagan, A. y Ullah, A., 1999, *Nonparametric Econometrics*, Cambridge University Press, Cambridge.
- Park, B., 2002, On the quantile regression based tests for asymmetry in stock return volatility, *Economic Journal*, 16, 175-91.
- Peracchi, F., 2001, *Econometrics*, Wiley, Chichester, UK.
- Portnoy, S., 1991, Asymptotic behavior of the number of regression quantile breakpoints, *SIAM Journal*.
- Portnoy, S., 2003, Censored regression quantiles, *Journal of the American Statistical Association*, 98, 464, 1001-12.
- Portnoy, S. y Koenker, R., 1997, The gaussian hare and the laplacian tortoise: computability of squared errors vs. absolute-error estimators, with discussion, *Statistical Science*, 12, 279-300.
- Powell, J., 1986, Censored regression quantiles, *Journal of Econometrics*, 32, 143-155.
- Stigler, S., 1999, *Statistics on the Table*, Harvard University Press, Cambridge.
- Taylor, J.W., 1999, A quantile regression approach to estimating the distribution of multiperiod returns, *Journal of Derivatives*, 7, 64-78.
- Taylor, J.W., 2000, A quantile regression neural network approach to estimating the conditional density of multiperiod returns, *Journal of Forecasting*, 19, 299-311.
- Yu, K. y Jones, M., 1998, Local linear quantile regression, *Journal of the American Statistical Association*, 93, 441, 228-237.
- Zamudio Carrillo, A., 2001, Educación y la distribución condicional del ingreso: una aplicación de regresión cuantil, *El Trimestre Económico*, 68, 39-70.
- Zarate, H., 2002, Cambios en la Estructura Salarial: Una Historia desde la Regresión Cuantílica, Banco Central de la República, serie Borradores de Economía, Bogota.

APÉNDICE.

Estimación e inferencia en modelos de regresión por cuantiles

Recordemos el tipo de problema que resuelve el método de mínimos cuadrados:

$$\min_{b_0, b_1 \in \mathbb{R}^2} \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

Enfrentado a este problema, cualquier estudiante de econometría básica puede visualizar una nube de puntos y una recta que pasa por el medio de la misma representando la solución al buscada. La naturaleza de la solución (que sea una recta y que pase por el medio) tiene que ver con que el método define como “error” a la distancia vertical entre el punto (y_i, x_i) y la recta implícitamente definida por los puntos $(x_i, b_0 + b_1 x_i)$. Que a los efectos del problema los errores aparezcan en la función a minimiza elevados al cuadrado tiene que ver con que los errores positivos importan tanto como los negativos. Consecuentemente, la solución al problema es una recta que pasa ‘por el medio’ de la nube de punto.

Consideremos ahora la siguiente variación del problema anterior:

$$\min_{b_0, b_1 \in \mathbb{R}^2} \sum_{i=1}^n |y_i - b_0 - b_1 x_i|$$

La única diferencia es que hemos reemplazado los cuadrados por valores absolutos. Consecuentemente, ninguna de las intuiciones acerca de la naturaleza de la solución cambia: por construcción la solución al problema produce una recta que pasa “por el medio” de la nube de puntos ya que los errores positivos y negativos también son tratados en forma simétrica. Las diferencias entre ambos métodos tienen que ver con el *grado* en el cual se penalizan los errores, el primero de ellos lo hace en forma cuadrática mientras que el segun-

do lo hace en forma lineal, pero, ciertamente, ambos lo hacen en forma simétrica.

De acuerdo a lo discutido en la sección 3, el método de regresión por cuantiles tiene como objetivo pasar rectas de regresión para distintos niveles de la distribución condicional de la variable explicada, consecuentemente, es natural conjeturar que la solución al problema pasara por tratar a los errores en forma *asimétrica* en los problemas anteriormente descriptos. Por ejemplo, si la función de penalidad castigase relativamente mas los errores positivos que los negativos, la recta tendería a ubicarse ‘por arriba’ de la nube de puntos, de modo de cometer pocos errores positivos.

A los efectos de simplificar la notación, definamos $e_i = y_i - b_0 - b_1 x_i$, y consideremos el siguiente problema de optimización:

$$\min_{b_0, b_1 \in \mathbb{R}^2} \sum_{i=1}^n e_i (t - 1 [e_i < 0]),$$

en donde t pertenece a $(0, 1)$, y $1[.]$ es una función indicadora que vale $1[e_i < 0] = 1$ si $e_i < 0$ y 0 en caso contrario. Es fácil observar que si $t=0.5$, $t - 1[e_i < 0]$ es igual a 0.5 si $e_i > 0$ e igual a -0.5 en caso contrario, de modo que el problema de optimización se reduce a:

$$\min_{b_0, b_1 \in \mathbb{R}^2} \sum_{i=1}^n 0.5 |e_i|,$$

que, más allá de una constante, es idéntico al problema planteado anteriormente, en donde los errores positivos y negativos son penalizados de la misma forma. Consideremos ahora el caso en donde $t > 0.5$, por ejemplo $t = 0.75$. Ahora $t - 1[e_i < 0]$ es igual a 0.75 si $e_i > 0$ e igual a -0.25 en caso contrario. Intuitivamente, lo que está sucediendo es que la función objetivo está penalizando los errores positivos el triple de lo que penaliza los errores negativos, consecuentemente, la recta de regresión resultante tendera a pasar por la parte superior de la nube de puntos, cometiendo pocos errores positivos, que son fuertemente penalizados. Por el contrario, si comenzamos con un valor muy

pequeño de τ , por ejemplo, $\tau=0.1$, los errores positivos reciben una penalización muy chica mientras que los negativos son muy penalizados: la recta de regresión pasara por la parte inferior de la nube de puntos. Entonces, es fácil deducir que comenzando por un valor de τ muy cercano a 0, a medida que incrementemos τ tendera a pasar cada vez mas arriba de la nube de puntos, y en el caso intermedio $\tau=0.5$ la recta tendera a pasar por el centro de la misma.

En el *paper* fundacional de la literatura de regresión por cuantiles, Roger Koenker y Gilbert Bassett (1978) muestran que, efectivamente, el método descrito anteriormente provee estimaciones consistentes de los coeficientes $\beta(\tau)$ que definen las rectas de regresión por cuantiles.

Consideremos el caso general de K variables explicativas. El problema de optimización sera:

$$\min_{b(\tau) \in \mathbb{R}^k} \sum_{i=1}^n \rho_{\tau}(y_i - x'_i b(\tau))$$

con $\rho = z(\tau - I[y_i - x'_i b(\tau)])$. Una observación crucial para el estudio de las propiedades de método de estimación y para el desarrollo de métodos computacionales es que en realidad el mismo es reformulable como un problema estándar de programación lineal, introduciendo $2n$ variables de holgura complementaria. Sea $e_i = v x'_i b(\tau)$ y e un vector columna que contiene los e_i 's y I_n un vector columna de n unos, el problema de optimización anterior puede ser reexpresado como:

$$\min_{(b(\tau), u, v) \in \mathbb{R} \times \mathbb{R}_+^{2n}} \{ \tau I'_n u + (1 - \tau) I'_n v \mid I'_n b(\tau) + u - v = y \}$$

Esta apreciación fue de gran relevancia a la hora de desarrollar métodos computacionales rápidos y precisos para la implementación practica de estos métodos, en base a avances recientes en programación lineal, en particular en la literatura de algoritmos de punto interior, que permiten resolver dichos problemas con notorias ventajas computacionales. Los principales avances teóricos en relación a las propiedades asintóticas de este tipo de estimadores provi-

nieron del nexo existente entre el problema dual de este programa lineal y la teoría asintótica de rangos establecida por Hajek y Sidak (1967).

Desde un punto de vista práctico, el principal aporte del trabajo de Koenker y Bassett (1978) fue establecer un marco analítico del cual se deriva una teoría sólida para la implementación de los métodos de regresión por cuantiles, que permiten no solo producir estimaciones puntuales de los parámetros de interés sino también realizar otros ejercicios de inferencia tales como construir intervalos de confianza o evaluar hipótesis. Denotemos con $\hat{\beta}(t)$ al estimador de $b(t)$ que surge como solución al problema de programación lineal anteriormente descrito. En la práctica es habitual estimar $\beta(t)$ para un conjunto de m cuantiles τ_1, \dots, τ_m , por ejemplo, en el ejemplo de la sección 4 estimamos coeficientes para los cuantiles 0.1, 0.3, 0.5, 0.7, 0.9, por lo que el proceso de estimación involucra cinco vectores de dos coeficientes cada uno (la ordenada al origen y la pendiente, para cada uno de los cuantiles elegidos). En términos generales, el proceso de estimación puntual produce estimadores $\beta(\tau_i)$, $i=1, \dots, m$. Denotemos con $\beta(\tau_i)$ a los vectores de K coeficientes poblacionales para los cuantiles τ_i , $i=1, \dots, m$, y definamos los vectores $\hat{\beta}(\tau) = (\hat{\beta}(\tau_1)' \dots \hat{\beta}(\tau_m)')$ y $\hat{\beta}(\tau) = (\beta(\tau_1)' \dots \beta(\tau_m)')$. Bajo el supuesto de que la muestra es independiente (pero no necesariamente idénticamente distribuida) y bajo condiciones de regularidad estándar, es posible mostrar que:

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(O, \Lambda)$$

en donde Λ es una matriz compuesta por submatrices $\Lambda_{j,p}$, $j=1, \dots, m$, $p=1, \dots, m$ en donde:

$$\Lambda_{j,p} = (\min \{ \tau_j, \tau_p \} - \tau_j, \tau_p) (E [f_{\tau_j}(0/x) xx'])^{-1} E [xx'] (E [f_{\tau_j}(0/x) xx'])^{-1}$$

en donde $f_{\tau_s}(0/x)$ es la función de densidad de y condicional en $x'\beta(\tau_s)$. La implementación empírica de este resultado requiere estimar la matriz Λ , que eventualmente depende de poder estimar la magnitud $f_{\tau_s}(0/x)^{-1}$ (llamada *sparsity*, en inglés). Existen dos alternativas para realizar inferencias tales como eva-

luar hipótesis simples o construir intervalos de confianza. Una consiste en apelar a métodos de *bootstrap* y una segunda consiste en apelar a invertir tests de hipótesis basados en la teoría de rangos, una discusión detallada de estos procedimientos excede los propósitos de esta nota. Buchinsky (1998) presenta una útil comparación de métodos directos basados en la estimación de Λ y en *bootstrap*. Koenker (1994) compara la performance de varias estrategias, incluyendo los métodos basados en rangos.