

# monetaria

CENTRO DE ESTUDIOS MONETARIOS LATINOAMERICANOS

VOLUMEN XXXI, NÚMERO 3, JULIO-SEPTIEMBRE

2008

# CEMLA

## ASAMBLEA

Bancos Centrales Asociados (*vox et votum*) y Miembros Colaboradores (*vox*)

## JUNTA DE GOBIERNO, 2007-2009

**Presidente:** Banco Central de Reserva del Perú □ **Miembros:** Banco Central de la República Argentina □ Banco Central de Brasil □ Banco Central de Honduras □ Banco de Jamaica □ Banco de México (permanente) □ Banco Central de Venezuela.

AUDITORÍA EXTERNA  
Banco de México

## PERSONAL DIRECTIVO

**Director General:** Kenneth Coates □  
**Subdirector General:** Luiz Barbosa  
□ *Coordinadora Institucional de Capacitación:* Jimena Carretero Gordon  
□ *Coordinador Institucional de Programas y Reuniones Técnicas de Banca Central:* Fernando Sánchez Cuadros □ *Coordinadora Institucional de Servicios de Información:* Ana-Laura Sibaja Jiménez □ *Coordinador Institucional de Administración y Finanzas:* Javier Granguillhome Morfin.

**monetaria** es una publicación trimestral del Centro de Estudios Monetarios Latinoamericanos, Durango n° 54, México, D. F., 06700. Impresa en los talleres de Alejandro Duplancher, Mariano Escobedo n° 114-3B, México, D. F., 11320. 380 ejemplares. ISSN 0185-1136.

# monetaria

VOLUMEN XXXI, NÚMERO 3, JULIO-SEPTIEMBRE DE 2008

James J. Heckman

**291** Causalidad econométrica

Daniel Barráez

Carolina Pagliacci

**339** Identificación de segmentos de precios en el mercado de fondos *overnight* usando modelos ocultos de Markov

Manfred Esquivel Monge

**361** Histéresis en dolarización: evidencias de la economía costarricense

Alberto Humala

**405** South American disinflation and regime switches: unobserved volatility components?

*Los trabajos firmados son responsabilidad de los autores y no coinciden necesariamente con el criterio del Centro de Estudios Monetarios Latinoamericanos.*



*James J. Heckman*

# Causalidad econométrica

## I. INTRODUCCIÓN

Los economistas y estadísticos hacen inferencias causales y recurren, en parte, a un conjunto común de instrumentos. Los economistas enfocan la causalidad desde la perspectiva de la evaluación de políticas. En la economía, los parámetros causales y las inferencias de causalidad están motivados por temas de políticas. Como los distintos temas de política requieren distintos parámetros, no existe un parámetro causal universal ni un

*Traduce y publica el CEMLA, con la debida autorización, la investigación de J. J. Heckman (Departamento de Economía, Universidad de Chicago), Econometric Causality, NBER Working Paper, n° 13934 (<http://www.nber.org/papers/w13934>). Este trabajo fue presentado en la ISI Conference en Seúl, Corea el 27 de agosto del 2001. Se ha enriquecido de las discusiones con Jaap Abbring, Pedro Carneiro, Steve Durlauf, Seong Moon, Salvador Navarro, T. N. Srinivasan, John Trujillo, Ed Vytlačil, Jordan Weil e Yu Xie; así como de comentarios del editor, Eugene Seneta y dos árbitros anónimos. Esta investigación fue apoyada por NIH R01-HD043411, NSF SES-0241858 y el Geary Institute, University College, Dublín, Irlanda. Las opiniones expresadas son las del autor y no necesariamente de los auspiciantes nombrados ni del Bureau of Economic Research. © 2008 por James J. Heckman. Todos los derechos reservados. Podrán citarse sin permiso expreso pequeños trozos de texto de hasta dos párrafos, siempre y cuando se atribuya enteramente a la fuente incluyendo la notificación de ©.*

conjunto de ellos. Este trabajo informa a los estadísticos de los avances en la economía que resultan útiles para atender los problemas de las políticas.

El enfoque econométrico desarrolla modelos explícitos de resultados donde se investigan las causas de los efectos y se analizan los mecanismos que rigen la elección del tratamiento. Se estudia la relación entre los resultados del tratamiento y los mecanismos de elección del tratamiento. Una cuidadosa justificación de las no observables en las ecuaciones de resultados y de selección de tratamientos facilita el diseño de estimadores. Se consideran evaluaciones tanto objetivas como subjetivas, donde las evaluaciones subjetivas son aquellas tanto de la persona recibiendo el tratamiento como de quienes lo asignaron. Se analiza las diferencias entre resultados esperados y realizados, tanto subjetivos como objetivos. Se desarrollan modelos para analizar los efectos de tratamiento simultáneos. Se realiza una distinción cuidadosa entre los modelos para potenciales resultados y los métodos empíricos para identificar los efectos del tratamiento.

El trabajo se desarrolla de la siguiente manera. La sección II distingue entre tres problemas diferentes en el análisis de los modelos causales y define el enfoque econométrico. La sección III discute una variedad de asuntos sobre la evaluación de políticas y los parámetros causales utilizados por los economistas. La sección IV analiza los contrafácticos, la causalidad y los modelos econométricos estructurales, y contrasta el enfoque econométrico con el enfoque adoptado en la estadística. La sección V presenta una síntesis de ambos enfoques.

## II. EL ENFOQUE ECONOMÉTRICO

Los contrafácticos son los posibles resultados en los distintos estados hipotéticos del mundo. Un ejemplo serían los diferentes resultados para la salud de una persona asociados con ingerir o no un medicamento. Las comparaciones causales suponen contrastar entre los resultados en los estados posibles definidos de tal forma que sólo la presencia o ausencia del medicamento varía entre los estados. La persona que recibe el medicamento es la misma que no la recibe, salvo por su estado ante el tratamiento y, quizás, el resultado asociado a su estado ante el tratamiento. El problema de inferencia causal consiste en evaluar si, manteniendo constantes todos los demás factores, la aplicación del tratamiento afecta el resultado. El concepto de causalidad que

se desarrolla en este trabajo y en la literatura estadística sobre los efectos de tratamiento se basa en la idea de variación controlada, variación en el tratamiento manteniendo los demás factores constantes. Se diferencia de otros conceptos de causalidad basados en la predicción (por ejemplo, Granger 1969 y Sims, 1972). Holland (1986) realiza algunas distinciones útiles entre definiciones de causalidad comúnmente invocadas. Cartwright (2004) analiza diversas definiciones de causalidad desde una perspectiva filosófica.

El enfoque econométrico de inferencia causal distingue cuidadosamente tres problemas: *i*) definición de los contrafácticos; *ii*) identificación de los modelos causales a partir de datos idealizados acerca de la distribución de poblaciones (infinitas muestras sin variación de muestreo), y *iii*) identificación de los modelos causales a partir de datos empíricos, en donde la variabilidad de la muestra se convierte en un factor. El contraste entre (ii) y (iii) surge de la diferencia entre distribuciones empíricas basadas en datos muestrales y las distribuciones poblacionales que generan los datos. El cuadro 1 delinea los tres problemas distintos:

**CUADRO 1.** TRES TAREAS DISTINTAS QUE SURGEN DEL ANÁLISIS DE MODELOS CAUSALES

<i>Tarea</i>	<i>Descripción</i>	<i>Requisitos</i>
1	Definir el conjunto de eventos hipotéticos o contrafácticos	Una teoría científica
2	Identificar los parámetros causales a partir de datos hipotéticos de la población	Análisis matemático de identificación puntual o de conjunto
3	Identificar los parámetros a partir de datos reales	Teoría de estimación y de prueba

El primer problema conlleva la aplicación de ciencia, lógica e imaginación. Es también en parte un tema de costumbres. Un modelo de contrafácticos será más aceptado en la medida que más ampliamente sean aceptados sus ingredientes, es decir, las reglas empleadas para derivar un modelo, incluyendo si sigue o no las reglas de la lógica y la matemática, y como concuerde con las teorías establecidas. Los modelos son descripciones de mundos hipotéticos que se obtienen mediante variaciones, hipotéticamente, de los factores que determinan los resultados. Los modelos no son declaraciones empíricas ni descripciones del mundo real. Sin embargo, suelen usarse para hacer predicciones

acerca del mundo real y son a menudo, representaciones abstractas de descripciones empíricas.

El segundo problema (i) se trata de inferir a partir de muestras muy grandes. ¿Se puede recuperar contrafácticos (o promedios o distribuciones de contrafácticos) a partir de datos que estén libres de cualquier variación de muestreo? Este es el problema de identificación.

El tercer problema es sobre inferencia práctica. ¿Se puede recuperar un modelo dado o un contrafáctico deseado a partir de un conjunto de datos dado? Las soluciones a este problema implican temas de inferencia y pruebas en muestras del mundo real. Este es el problema con el cual están más familiarizados los estadísticos y los científicos sociales empíricos. La frontera entre los problemas (ii) y (iii) es permeable, según como se definan *los datos*.

Parte de la controversia en torno a la construcción de contrafácticos y modelos causales es parcialmente una consecuencia de la falta de claridad acerca de estos tres problemas distintos por parte de los analistas, que suelen confundirlos. En ciertos círculos se han asociado algunos métodos particulares de estimación (por ejemplo, fusión estadística de bases de datos o estimación por variable instrumental) con la *inferencia causal*, y aún las definiciones de ciertos *parámetros causales* se han consolidado debido a problemas de definición, identificación y estimación.

El enfoque econométrico de la evaluación de políticas separa estos problemas y enfatiza la naturaleza provisional del conocimiento causal. Algunos estadísticos rechazan la noción de la naturaleza provisional del conocimiento causal y buscan un enfoque de la inferencia causal libre de supuestos (ver, por ejemplo, Tukey, 1986). Sin embargo, el conocimiento humano avanza mediante el desarrollo de modelos teóricos y su examen frente a los datos. Los modelos utilizados son inevitablemente provisionales y dependientes de supuestos *a priori*. Aun la aleatorización, debidamente ejecutada, no puede dar respuestas a todas interrogantes causales relevantes.

Muchos *modelos causales* en estadística son una guía incompleta para la interpretación de datos o para sugerir respuestas a preguntas específicas sobre políticas. Están motivados por el ideal del experimento. No especifican claramente los mecanismos que determinan cómo se dan los contrafácticos hipotéticos, ni cómo se instrumentan las intervenciones hipotéticas, salvo en la comparación de intervenciones generadas *aleatoriamente* o *no-aleatoriamente*. Se enfocan únicamente en los resultados, dejando el modelo para elección de resultados especificado sólo



implícitamente. La construcción de los resultados contrafácticos se basa en apelaciones a la intuición y no a modelos formales.

Debido a que los mecanismos determinantes de la selección de resultados no se modelan en el enfoque estadístico, a menudo se adopta la metáfora de la *asignación aleatoria*. Dicho énfasis en la aleatorización o sus sustitutos, como la fusión estadística o las variables instrumentales, elimina una variedad de canales alternativos de identificación de contrafácticos a partir de datos poblacionales o muestrales. El énfasis sobre la aleatorización tiene consecuencias prácticas debido a la consolidación de la tarea 1 con las tareas 2 y 3 en el cuadro 1. Dado que la aleatorización se utiliza para definir los parámetros de interés, en la práctica esto a veces lleva a la confusión de que ésta es la única, o al menos la mejor, manera para identificar parámetros causales a partir de datos reales. Las versiones más extremas de este enfoque niegan el estado causal a cualquier intervención que en principio no pueda instrumentarse mediante un experimento práctico en el mundo real.

Una razón por la cual muchos modelos estadísticos son incompletos es porque no especifican las fuentes de aleatoriedad que generando variabilidad entre agentes, es decir, no especifican por qué las personas hacen elecciones distintas y tienen resultados diferentes, siendo que en todo lo demás son idénticas desde el punto de vista de observación. No distinguen entre el contenido del conjunto de información de los agentes y el contenido del conjunto de información del observador estadístico, aun cuando esta distinción es básica para justificar las propiedades de todo estimador para la resolución de problemas de selección y evaluación. No distinguen la incertidumbre desde el punto de vista del agente cuyo comportamiento está siendo analizado, de la variabilidad como es investigada por el analista observador. También son incompletos porque son recursivos. No permiten simultaneidad en la elección de resultados de tratamiento, que constituye un aspecto central de la teoría de juegos y de los modelos de interacción social y contagio (ver, por ejemplo, Brock y Durlauf, 2001; Tamer, 2003).

Desde Haavelmo (1943, 1944) los economistas han reconocido el valor de modelos precisos para la construcción de contrafácticos, para responder a preguntas de *causalidad* y para dirigirse a temas más generales acerca de la evaluación de políticas. El marco econométrico es explícito acerca como se generan modelos de contrafácticos, de las fuentes de las intervenciones (las reglas de asignación de *tratamientos*) y de las fuentes de no observables en las asignaciones de tratamientos, sus resultados y

la relación entre ellos. En lugar de dejar implícitas las reglas que rigen la selección de tratamientos, el enfoque econométrico emplea relaciones explícitas entre no observables en los resultados y los mecanismos de selección para identificar los modelos causales a partir de los datos y aclarar la naturaleza de los supuestos de identificación.

La finalidad de la literatura económica, así como la meta de toda ciencia, es comprender las causas que producen efectos para poder utilizar versiones empíricas de los modelos para pronosticar los efectos de intervenciones nunca antes experimentadas, calcular diversos contrafácticos de políticas y emplear la teoría científica como guía para la selección de estimadores e interpretación de las evidencias. Estas actividades requieren el desarrollo de una teoría más elaborada que la visualizada en la literatura estadística actual sobre inferencia causal.

Muchos modelos causales en la estadística son dispositivos tipo *caja negra* diseñados para investigar el impacto de *tratamientos*, con frecuencia paquetes complejos de intervenciones, sobre resultados observados en un entorno dado. Rara vez se intenta separar los componentes de los tratamientos complejos. Los modelos científicos explícitos penetran la caja negra para explorar el o los mecanismos que producen los efectos. Según la terminología de Holland (1986), se debe distinguir entre comprender *los efectos de las causas* (la meta de la literatura sobre efectos de tratamiento, según la definen un importante número de estadísticos) y comprender *las causas de los efectos* (la meta de la literatura econométrica que construye modelos explícitos).

Al enfocar un asunto de caja negra limitado, la literatura sobre efectos de tratamiento evita muchos de los problemas enfrentados por la literatura econométrica que construye modelos explícitos de contrafácticos y mecanismos de asignación. Esta es su gran virtud. Al mismo tiempo produce parámetros de aplicación más limitada. A falta de supuestos adicionales, dichos parámetros no se prestan a extrapolaciones fuera de muestra ni a precisar pronósticos del impacto de otras políticas además de las que están bajo investigación empírica. Al no ser explícito acerca del contenido de la caja negra (comprender las causas de los efectos), la literatura sobre efectos de tratamiento se ata las manos en cuanto al uso de la información acerca de parámetros de comportamiento básicos obtenidos en otros estudios, así como de la intuición científica para complementar la información disponible al momento. Carece de la capacidad para proveer explicaciones para los *efectos* estimados de acuerdo con la teoría. Cuando los componentes de los tratamientos varían entre

investigaciones, el conocimiento no se acumula a través de los estudios de efectos de tratamiento, mientras que sí se acumulan entre los estudios que estiman modelos generados a partir de parámetros comunes dentro del enfoque econométrico.

### III. TEMAS DE EVALUACIÓN DE POLÍTICAS Y CRITERIOS DE INTERÉS

Esta sección presenta tres cuestiones centrales en la evaluación de políticas. Se definen los efectos de tratamiento a nivel individual y se trata el problema de la evaluación en términos generales.

#### 1. Los tres problemas de evaluación de políticas considerados en este trabajo

La economía considera tres clases amplias de problemas en la evaluación de políticas. El primer problema para la evaluación de políticas es:

- *Problema 1:* Evaluar el impacto de intervenciones históricas sobre los resultados, incluyendo su impacto en términos del bienestar sobre los receptores del tratamiento y la sociedad en general.

Por históricas, se entiende que son intervenciones documentadas. Se emplea una variedad de resultados y criterios para formar estas evaluaciones, según el asunto al momento. Los economistas distinguen los resultados objetivos o públicos, que en principio pueden ser medidos por cualquier observador, de los resultados *subjetivos* que consisten en las evaluaciones de los agentes sujetos al tratamiento (por ejemplo, pacientes) o de los agentes que prescriben el tratamiento (por ejemplo, los médicos). Los resultados objetivos son de naturaleza intrínsecamente *ex post* (posteriores al hecho). La literatura estadística acerca de la inferencia causal enfoca exclusivamente los resultados objetivos *ex post*. Los resultados subjetivos pueden ser *ex ante* (previstos) o *ex post*. De tal forma, el resultado de un ensayo médico produce a la vez una tasa de curación junto al dolor y sufrimiento del paciente. El dolor y sufrimiento previstos *ex ante* pueden diferir del dolor y sufrimiento experimentados *ex post*. Los agentes pueden tener también evaluaciones *ex ante* de los resultados objetivos que difieran de sus evaluaciones *ex post*. Por impacto, me refiero a la construcción de contrafácticos sean a nivel individual o poblacional, y sus valoraciones. Por bienestar,

quiero decir las valoraciones de los resultados obtenidos de la intervención por parte de los agentes sujetos de análisis u otra parte (por ejemplo, los padres del agente o la sociedad en general). Pueden ser *ex ante* o *ex post*. P1 es el problema de la *validez interna*. Es el problema de identificar un parámetro dado del tratamiento (o un conjunto de parámetros) en un entorno dado.

La mayor parte de la evaluación de políticas es diseñada con vistas al futuro y para tomar decisiones informadas acerca de políticas nuevas y la aplicación de políticas anteriores a entornos nuevos. En el análisis de políticas es útil identificar un segundo problema.

- *Problema 2*: Pronosticar los impactos (construir estados contrafácticos) de las intervenciones instrumentadas en un entorno dado sobre los demás entornos, incluyendo sus impactos en términos del bienestar.

Estas intervenciones comprenden las políticas descritas mediante características genéricas (por ejemplo, tasas impositivas o de beneficios, o terapia empleada, incluyendo intensidad) que son aplicadas a distintos grupos de personas o en distintos momentos de los estudiados en la instrumentación de políticas para las cuales hay datos disponibles. Este es el problema de la *validez externa*: trasladar un parámetro de un tratamiento o conjunto de parámetros estimados a otro entorno (ver, por ejemplo, Shadish y Cook, 2007). El entorno incluye las características de los individuos y los tratamientos.

Al final, el problema más ambicioso consiste en pronosticar el efecto de una nueva política, nunca antes experimentada.

- *Problema 3*: Pronosticar los impactos de las intervenciones históricamente no experimentadas (construir estados contrafácticos asociados con intervenciones) en diversos entornos, incluyendo su impacto en términos de bienestar.

Este problema requiere la utilización del pasado para pronosticar las consecuencias de políticas nuevas. Es un problema fundamental del conocimiento. P3 es un problema que los economistas deben resolver a diario. A continuación se presenta un marco dentro del cual los analistas pueden atender estos problemas de una manera sistemática. También es un marco que puede usarse para la inferencia causal.

## 2. Definición de efectos de tratamiento a nivel individual

El evaluar consiste en valorar, y comparar valores entre posibles

resultados. Estas son dos tareas distintas. Se definen resultados correspondientes al estado (política, tratamiento)  $s$  para un agente  $\omega$  como  $Y(s, \omega)$ ,  $\omega \in \Omega$ . El agente puede ser un hogar, un paciente, una empresa o un país. Se puede pensar en  $\Omega$  como un universo de agentes. Supongamos que  $\Omega = [0,1]$ .  $Y(\cdot, \cdot)$  puede ser un vector de valores, pero para simplificar el argumento trabajo con resultados escalares. (Ver Heckman y Vytlačil, 2007a, para el análisis con resultados en vectores).

Los  $Y(s, \omega)$  son resultados que se obtienen después de que los tratamientos son elegidos. Antes del tratamiento los agentes pueden desconocer los  $Y(s, \omega)$ , pero pueden realizar pronósticos acerca de éstos. Estos pronósticos pueden influir sus decisiones de participar en el programa, o pueden influir sobre los agentes quienes deciden si un individuo participa o no en el programa. La selección en el programa se basa en los componentes reales o previstos de los resultados dados, da lugar al problema de selección en la literatura de evaluación.

Definamos  $S$  como el conjunto de tratamientos posibles con elementos denotados por  $s$ . Para simplificar la exposición, suponemos que este conjunto es igual para todo  $\omega$ . Para cada  $\omega$  se obtiene una colección de posibles resultados dada por  $\{Y(s, \omega)\}_{s \in S}$ . Por simplicidad, se supone que el conjunto  $S$  es finito (Heckman y Vytlačil, 2007a, consideran casos más generales). Por ejemplo, si  $S = \{0,1\}$  existen dos tratamientos, uno de los cuales puede ser un estado sin tratamiento (por ejemplo,  $Y(0, \omega)$  es el resultado para el agente  $\omega$  quien no recibe un tratamiento, que puede ser un medicamento, educación o acceso a una nueva tecnología, mientras que  $Y(1, \omega)$  es el resultado del estado 1 del tratamiento para el agente  $\omega$  quien sí recibe el medicamento, la educación o el acceso). Aunque la literatura teórica dedica mayor atención al entorno caracterizado por dos tratamientos, en la práctica es más común toparse con entornos de múltiples tratamientos.

Cada *estado* (tratamiento) puede consistir en un compuesto de estados subcomponentes. En tal caso, se podría definir al propio  $s$  como un vector (por ejemplo,  $s = (s_1, s_2, \dots, s_K)$  para  $K$  componentes) correspondiente a los distintos componentes que comprende el tratamiento. Así, un protocolo médico típicamente consiste en un paquete de tratamientos. Uno podría interesarse en el paquete de uno (o más) de sus componentes. De tal forma  $s_1$  podría ser los meses de tratamiento con un medicamento,  $s_2$  la calidad de los médicos, y así sucesivamente. No se pierde generalidad al suponer que  $s$  es escalar, ya que a cada tratamiento distinto se le puede adjudicar una etiqueta distinta.

Pueden usarse subíndices para los resultados, con  $Y_t(s, \omega)$  correspondiente con los resultados del tratamiento medidos en distintos momentos. El conjunto de índices para  $t$  puede ser de enteros correspondientes a tiempo discreto, o de intervalos correspondientes a tiempo continuo. En principio, podrían colocarse índices  $t$  para  $S$ , definido ya sea por enteros correspondientes a tiempo discreto o a intervalos correspondientes a tiempo continuo. Los  $Y_t(s, \omega)$  son resultados ya realizados o *ex post* (posteriores al tratamiento). Al elegir un tratamiento, puede que estos valores no se conozcan. Gill y Robins (2001), Van der Laan y Robins (2003), Abbring y Van den Berg (2003), Abbring y Heckman (2007, 2008) y Heckman y Navarro (2007) desarrollan modelos para contrafácticos dinámicos, en donde los  $S$  con índices por tiempo o por  $\omega$  van surgiendo a medida que la información se acumula. El subíndice del tiempo se mantiene implícito a lo largo de este ensayo.

El *efecto de tratamiento individual* para el agente  $\omega$ , comparando los resultados objetivos del tratamiento  $s$  con los resultados objetivos del tratamiento  $s'$ , es:

$$(1) \quad Y(s, \omega) - Y(s', \omega), \quad s \neq s',$$

para dos elementos  $s, s' \in S$ . Esto también se conoce como el *efecto causal a nivel individual*. El efecto causal es el cambio de resultados *ceteris paribus* Marshalliano (1890) para un agente entre estados  $s$  y  $s'$ . Solo  $s$  y  $s'$  varían.

Existen otras comparaciones de interés en la evaluación de un programa. Los economistas están interesados en el bienestar de los participantes, además de los resultados objetivos (ver Heckman y Smith, 1998). Aunque los estadísticos suelen razonar en términos de mecanismos de asignación, los economistas reconocen que son las preferencias de los agentes las que suelen regir las elecciones reales. Se pueden llevar a cabo comparaciones entre resultados en términos de utilidad (personal,  $R(Y(s, \omega), \omega)$ ) o en términos de las preferencias de un planificador o de un médico,  $R_G$ , o pueden hacerse ambos tipos de comparaciones para el mismo resultado, y evaluando su concordancia u oposición). Las funciones de utilidad producen valoraciones subjetivas de los resultados por los agentes con tratamiento o el planificador.

Para simplificar la notación y al mismo tiempo permitir posibilidades más generales para los argumentos de la función de valoración, expresemos  $R(Y(s, \omega), \omega)$  como  $R(s, \omega)$ , suprimiendo la dependencia explícita de  $R$  sobre  $Y(s, \omega)$ . Con esta notación se puede plantear si  $R(s, \omega) > R(s', \omega)$ , o no (¿el agente ha mejorado

como resultado del tratamiento  $s$  en relación con el tratamiento  $s'$ ?). La diferencia entre resultados subjetivos es  $R(s, \omega) - R(s', \omega)$ , y es un tipo de efecto de tratamiento. Manteniendo  $\omega$  fijo se mantienen fijas todas las características del agente salvo el tratamiento asignado,  $s$ . Dado que las unidades de utilidad  $R(s, \omega)$  son arbitrarias, se podría anotar un indicador para cada  $s$  y  $\omega$  si el resultado en  $s$  fuera mayor o menor que el resultado en  $s'$ , es decir, si  $R(s, \omega) > R(s', \omega)$ , o no. Este también es un tipo de efecto de tratamiento. Puede ser que los agentes que toman decisiones acerca del tratamiento estén parcialmente informados al momento de decidir sobre las recompensas resultantes. La modelación de la diferencia entre los resultados esperados y realizados es una parte integral del enfoque econométrico de la causalidad y la evaluación de políticas. Una característica central del enfoque econométrico de la evaluación de programas es la evaluación de las valoraciones subjetivas como son percibidas por los tomadores de decisiones, y no sólo de las valoraciones objetivas.

El término *tratamiento* se emplea de distintas maneras en distintas literaturas. En su uso más común, un mecanismo de asignación de tratamientos es una regla  $\tau: \Omega \rightarrow \mathcal{S}$  que establece un tratamiento para cada individuo  $\omega$ . Las consecuencias de la asignación son los resultados  $Y(s, \omega)$ ,  $s \in \mathcal{S}$ ,  $\omega \in \Omega$ . La colección de estas posibles reglas de asignación es  $\mathcal{T}$  donde  $\tau \in \mathcal{T}$ . Hay dos aspectos de una política según esta definición. La política decide qué le corresponde a quién. En términos más precisos, selecciona a los individuos  $\omega$  y especifica los tratamientos  $s \in \mathcal{S}$  recibidos.

La literatura econométrica ofrece una definición con más matices de la asignación de tratamientos, que reconoce explícitamente el elemento de elección del agente  $\omega$  en la elaboración de la regla de asignación de tratamientos. El tratamiento puede comprender la participación en actividades como educación, capacitación, una terapia médica, la adopción de una tecnología específica y afines. Muchas veces la participación en el tratamiento es a elección del agente. La modelación de este proceso de elección es una característica distintiva del enfoque econométrico. Según esta definición más comprensiva de tratamiento, a los agentes se les asigna incentivos tales como impuestos, subsidios, dotaciones y elegibilidad que afectan sus elecciones, pero es el agente quien elige el tratamiento seleccionado. Las preferencias del agente, los sistemas de entrega de programas, las estructuras de mercado y aspectos similares pueden afectar la elección del tratamiento. El mecanismo de selección de tratamiento puede involucrar a múltiples actores y múltiples decisiones que

resultan en la asignación de  $s$  a  $w$ . Por ejemplo,  $s$  puede ser educación mientras que  $Y(s, w)$  son los ingresos según el nivel de escolaridad del agente  $w$ . Una política puede consistir en un conjunto de pagos que estimulen la escolaridad, como en el programa PROGRESA en México, y en tal caso el tratamiento es la elección de escolaridad con sus consecuencias para los ingresos. La  $s$  también puede consistir en un protocolo médico que requiere cumplimiento por parte del paciente (elección de comportamiento) para ser efectivo.

La siguiente descripción de asignación de tratamientos reconoce las elecciones y restricciones individuales y es más apropiada para la evaluación de políticas que admiten el papel de la elección por parte de los agentes. Especifiquemos reglas de asignación  $a \in \mathbf{A}$  que mapean a los individuos  $w$  con restricciones (beneficios)  $b \in \mathbf{B}$  según distintos mecanismos. En esta notación, un mecanismo de asignación de restricciones  $a$  es un mapa  $a: \Omega \rightarrow \mathbf{B}$  definido sobre el espacio de los agentes. Las restricciones pueden incluir dotaciones, elegibilidad, impuestos, subsidios y demás incentivos que puedan afectar las elecciones de tratamiento por parte de los agentes. Los elementos de  $b$  pueden ser los parámetros de programas de impuestos y beneficios que afectan los incentivos individuales. En tanto una estructura más general es posible, donde programas específicos para  $w$  se asignan a la persona  $w$ , el costo de esta generalidad es una notación más compleja. Para simplicidad, limito el enfoque a un conjunto fijo, aunque posiblemente muy grande, de parámetros definidos para todos los agentes. El mapa  $a$  define la regla utilizada para asignar  $b \in \mathbf{B}$ . Puede incluir reglas determinísticas que producen programas que mapean  $w$  a  $\mathbf{B}$  tales como programas de impuestos o de elegibilidad. También puede incluir mecanismos de asignación aleatoria que asignan  $w$  a un elemento de  $\mathbf{B}$ . Los mecanismos de asignación aleatoria agregan elementos adicionales de aleatoriedad al entorno. Abusando de la notación, cuando se emplea la asignación aleatoria debe redefinirse  $\Omega$  para incluir esta nueva fuente de aleatoriedad.

Dado  $b \in \mathbf{B}$  asignados por un mecanismo de asignación de restricciones  $a \in \mathbf{A}$ , los agentes escogen sus tratamientos. Definamos el mecanismo de asignación de tratamientos  $\tau: \Omega \times \mathbf{A} \times \mathbf{B} \rightarrow \mathbf{S}$  como un mapa que lleva el agente  $w \in \Omega$  enfrentándose a restricciones  $b \in \mathbf{B}$  asignados por el mecanismo  $a \in \mathbf{A}$  con un tratamiento  $s \in \mathbf{S}$  (se utiliza una notación redundante para aclarar los conceptos). En los entornos con elección,  $\tau$  es la regla de elección empleada por los agentes donde  $\tau \in \mathbf{T}$ , que es un conjunto de reglas de elección posibles. Lo convencional es suponer



que los tomadores de decisiones relevantes eligen un único  $\tau \in \mathbf{T}$ , aunque ello no se requiere en esta definición. Se define como un régimen de políticas  $p \in \mathbf{P}$  una pareja  $(a, \tau) \in \mathbf{A} \times \mathbf{T}$ , que mapea los agentes denotados por  $\omega$  con los elementos de  $s$ . En esta notación,  $\mathbf{P} = \mathbf{A} \times \mathbf{T}$ .

La incorporación de la elección al análisis de los efectos de tratamiento es un ingrediente esencial y distintivo del enfoque econométrico de la evaluación de programas sociales. El análisis tradicional de tratamiento-control en la estadística considera como equivalentes a los mecanismos  $a$  y  $\tau$ . En dicha literatura, una asignación se traduce en una asignación de tratamiento, no en una asignación de incentivos y elegibilidad para tratamiento donde el agente es quien elige su tratamiento. En esta notación, el enfoque tradicional posee un solo mecanismo de asignación mientras que el incumplimiento se trata como un problema en lugar de una fuente de información acerca de las preferencias de los agentes, que a su vez es característica central del enfoque econométrico (Heckman y Smith, 1998). De tal forma que, con cumplimiento total,  $a: \Omega \rightarrow \mathbf{S}$  y  $a = \tau$ , donde  $\mathbf{B} = \mathbf{S}$ .

La invarianza de políticas es un supuesto clave para todo estudio de evaluación de políticas. Permite a los analistas caracterizar los resultados sin especificar cómo se obtienen tales resultados. La invarianza de políticas tiene dos aspectos. El primer aspecto es que para un  $b \in \mathbf{B}$  dado (un programa de incentivos), el mecanismo  $a \in \mathbf{A}$  por el cual se le asigna a cada  $\omega$  una  $b$  (por ejemplo, asignación aleatoria, coerción a punta de pistola, etc.) y el incentivo  $b \in \mathbf{B}$  se suponen irrelevantes para los valores de los resultados obtenidos para cada  $s$  elegido. En segundo lugar, para una  $s$  dada en el caso de un agente  $\omega$ , el mecanismo  $\tau$  por el cual  $s$  es asignado al agente según el mecanismo de asignación  $a \in \mathbf{A}$  es irrelevante para los valores que se asumen para los resultados realizados. Ambos supuestos definen que entienden los economistas por invarianza de políticas. Los primeros en definir y formalizar la invarianza fueron Marschak (1953) y Hurwicz (1962).

La invarianza de políticas permite describir los resultados como  $Y(s, \omega)$  e ignorar las características del entorno de políticas y elecciones en la definición de resultados. En cambio, si deben tomarse en cuenta los efectos de incentivos y mecanismos de asignación sobre los resultados, se debe trabajar con  $Y(s, \omega, a, b, \tau)$  en lugar de  $Y(s, \omega)$ . La descripción más compleja es el resultado asociado con el estado de tratamiento  $s$  para la persona  $\omega$ , a quien fue asignado el paquete de incentivos  $b$  por medio del mecanismo  $a$  que son los argumentos de la regla de asignación  $\tau$ . Para definiciones precisas de invarianza véase Heckman y Vytlačil (2007a).

Los supuestos de invarianza establecen que, para iguales tratamiento  $s$  y agente  $\omega$ , los distintos mecanismos de asignación de restricciones  $a$  y  $a'$  y asignaciones de estado de restricciones asociadas  $b$  y  $b'$  producen igual resultado. Por ejemplo, excluye la posibilidad de que el acto de aleatorizar o el acto de apuntar una pistola al agente para asegurar su cooperación con las intenciones del planificador, tengan un efecto sobre los resultados, dado el caso que el agente termina en  $s$ . Este es un supuesto fuerte.

El segundo supuesto de invarianza invocado en la literatura es que para  $a$  y  $b$  fijos, los resultados son iguales, independientemente del mecanismo de asignación de tratamientos. Este supuesto establece que el mecanismo real empleado para asignar tratamientos no afecta los resultados. Entre otras cosas excluye la interacción social, el contagio y los efectos de equilibrio general. Heckman (1992), Heckman y Smith (1998), Heckman, LaLonde y Smith (1999), y Heckman y Vytlacil (2007b) discuten las evidencias en contra de este supuesto, y Heckman, Lochner y Taber (1998a,b,c) muestran cómo suavizarlo.

Si también son considerados los efectos de tratamiento basados en evaluaciones subjetivas, lo que distingue al enfoque econométrico, es necesario ampliar los supuestos de invarianza hasta producir invarianza de recompensas para ciertas políticas y mecanismos de asignación (ver Heckman y Vytlacil, 2007a). Los supuestos de invarianza requeridos establecen, por ejemplo, que la utilidad no se ve afectada por la aleatorización ni por el mecanismo de asignación de las restricciones. Heckman (1992), Heckman, LaLonde y Smith (1999), y Heckman y Vytlacil (2007b) presentan evidencias en contra de este supuesto. Otro supuesto de invarianza elimina las interacciones sociales tanto en resultados objetivos como subjetivos. Resulta útil distinguir entre invarianza de resultados objetivos e invarianza de resultados subjetivos. La aleatorización puede afectar las evaluaciones subjetivas por medio de su efecto de agregar incertidumbre al proceso decisorio, pero no puede afectar las valoraciones objetivas. Los modelos del enfoque econométrico demuestran como los mecanismos de asignación y las interacciones sociales afectan a las ecuaciones de elección y resultado, en lugar de postular *a priori* que los postulados de invarianza de resultados sean siempre satisfactorios.

#### *a) Criterios más generales*

Son muchas las comparaciones que pueda hacer el analista (por ejemplo, ver Heckman, Smith y Clements, 1997). Se puede

comparar resultados en conjuntos diferentes que están ordenados. Defínase  $\operatorname{argmax}_{s \in \mathcal{S}} \{Y(s, \omega)\}$  como el valor de  $s$  que produce el elemento maximal de  $Y(s, \omega)$  para  $s \in \mathcal{S}$ . De tal forma que si  $Y(s, \omega)$  es ingreso escalar y se comparan resultados para  $s \in \mathcal{S}_A$  con resultados para  $s' \in \mathcal{S}_B$  donde  $\mathcal{S}_A \cap \mathcal{S}_B = \emptyset$ , entonces podrían compararse  $Y_{s_A}$  con  $Y_{s_B}$ , donde:

$$s_A = \operatorname{argmax}_{s \in \mathcal{S}_A} \{Y(s, \omega)\} \text{ y } s_B = \operatorname{argmax}_{s \in \mathcal{S}_B} \{Y(s, \omega)\},$$

donde se suprime la dependencia de  $s_A$  y  $s_B$  sobre  $\omega$ . Esto compara lo mejor de un conjunto de elecciones con lo mejor de otro. Otro contraste puede comparar la mejor elección con la siguiente mejor elección. Para hacerlo, se define  $s' = \operatorname{argmax}_{s \in \mathcal{S}} \{Y(s, \omega)\}$  y  $\mathcal{S}_B = \mathcal{S} \setminus \{s'\}$  y se define el efecto de tratamiento como  $Y_{s'} - Y_{s_B}$ . Esta es la comparación del resultado más alto en  $\mathcal{S}$  con el siguiente mejor resultado. Podrían construirse más comparaciones a nivel individual, y podrían calcularse utilizando preferencias personales,  $R(\omega)$ , usando las preferencias del planificador,  $R_G$ , o usando las preferencias del planificador sobre las preferencias de los agentes. Heckman (2005) y Heckman y Vytlačil (2007a,b) presentan una detallada discusión de los criterios alternativos.

### 3. El problema de evaluación

Con la ausencia de una teoría, no existen reglas bien definidas para construir contrafácticos o estados hipotéticos, ni para construir las reglas de asignación de tratamientos. Lewis (1974), sin una teoría articulada, define estados contrafácticos admisibles como “los mundos más cercanos posibles”. Su definición falla ante la ausencia de una métrica o topología significativa para medir *cercanía* entre mundos posibles. Las teorías científicas articuladas suministran algoritmos para generar el universo de estados contrafácticos internamente consistentes y teórico-consistentes. Estos estados hipotéticos son mundos posibles. Son el producto de una actividad puramente mental. Distintas teorías producen diferentes  $Y(s, \omega)$  y diferentes mecanismos de asignación.

El *problema de evaluación* es que el analista observa a cada agente en sólo uno de  $\bar{S}$  estados posibles. Se desconoce el resultado para dicho agente en otros estados que no se realizaron, y por lo tanto no se pueden formar directamente efectos de tratamiento a nivel individual. El *problema de selección* surge porque sólo se puede observar a ciertos agentes en un estado. Así, se

observa  $Y(s, \omega)$  sólo para aquellos agentes que optaron por (o fueron elegidos para) estar en dicho estado. En general, los resultados para agentes encontrados en  $S = s$  no son representativos de lo que serían los resultados de los agentes si fuesen asignados aleatoriamente a  $s$ .

El problema de evaluación es un problema de identificación que surge al construir los estados contrafácticos y las reglas de asignación de tratamientos producidos por estos modelos abstractos que usan datos. Este es el segundo problema del cuadro 1. El problema no se establece con precisión sino hasta que los datos disponibles para el analista se han definido en forma precisa. Distintas áreas del conocimiento suponen el acceso a distintos tipos de datos.

Para cada régimen de políticas, en cualquier momento dado se observa al agente  $\omega$  en algún estado pero en ninguno de los otros. Por lo tanto no se puede observar  $Y(s', \omega)$  para el agente  $\omega$  si se observa  $Y(s, \omega)$ ,  $s \neq s'$ . Dejamos  $D(s, \omega) = 1$  si se observa al agente  $\omega$  en estado  $s$  bajo el régimen de política  $p$ , donde mantengo implícito el régimen de política  $p$  para no complicar la notación. Con esta notación,  $D(s, \omega) = 1$  implica que  $D(s', \omega) = 0$ , para  $s \neq s'$ .

Se observa  $Y(s, \omega)$  si  $D(s, \omega) = 1$ , pero no  $Y(s', \omega)$ , para  $s \neq s'$ . Se puede definir el  $Y(\omega)$  observado como:

$$(2) \quad Y(\omega) = \sum_{s \in S} D(s, \omega) Y(s, \omega)$$

La construcción de una contraparte empírica del efecto causal a nivel individual (1) es imposible a partir de los datos sobre  $(Y(\omega), D(\omega))$ ,  $\omega \in \Omega$ , donde  $D(\omega) = \{D(s, \omega)\}_{s \in S}$ , sin supuestos adicionales. Esta formulación del problema de evaluación es conocida como el modelo de regresión cambiante de Quandt (Quandt, 1958) y en la estadística es atribuida a Neyman (1923), Cox (1958) y Rubin (1978). Existe una versión formulada por Haavelmo (1943) en un contexto de ecuaciones lineales para un continuo de tratamientos. El modelo de Roy (1951) es otra versión de este marco con dos resultados de tratamiento posibles ( $S = \{0, 1\}$ ) más una medición de resultado escalar y un mecanismo de asignación particular  $\tau$  que consiste en que  $D(1, \omega) = \mathbf{1}[Y(1, \omega) > Y(0, \omega)]$ , donde  $\mathbf{1}[\cdot]$  significa  $\mathbf{1}[\cdot] \geq 1$  si el argumento “.” es cierto, y de no ser así, = 0. De tal forma que  $\tau(\omega) = 1$  para el  $\omega$  que satisface  $Y(1, \omega) \geq Y(0, \omega)$  y  $\tau(\omega) = 0$  para el  $\omega$  que satisface  $Y(1, \omega) < Y(0, \omega)$ . El mecanismo de selección depende de los resultados potenciales. Como los agentes eligen el sector de mayores ingresos, el mecanismo real de selección no es aleatorio.

Se analizan versiones de este modelo con mecanismos de auto-selección más generales en Heckman (1990), Heckman y Smith (1998), Heckman y Vytlačil (2007a, b) y Abbring y Heckman (2007).

La literatura macroeconómica de evaluación analiza políticas de cobertura universal para un momento dado en el tiempo (por ejemplo, una política impositiva o la seguridad social) tales que  $D(s, \omega) = 1$  para algún  $s$  y todo  $\omega$ . Utiliza datos de series de tiempo para evaluar el impacto de políticas para diferentes períodos y típicamente emplea resultados promedio (o utilidad promedio) para evaluar las políticas.

El problema de la autoselección es un aspecto esencial del problema de evaluación cuando los datos son generados por medio de las elecciones de los agentes. Los agentes que toman las decisiones pueden no ser los mismos agentes que reciben el tratamiento (por ejemplo, padres que toman decisiones para sus hijos). Dichas elecciones pueden incluir el cumplimiento con los protocolos de un experimento social, así como las elecciones usuales que hace la gente en su diario vivir. Como consecuencia de la autoselección, la distribución observada de los  $Y(s, \omega)$  no coincide con la distribución de población de una muestra aleatoria de los  $Y(s, \omega)$ .

En un modelo de Roy prototípico, la elección de tratamiento (incluyendo las decisiones de no darse de baja del programa) provee información acerca de la evaluación relativa de los  $Y(s, \omega)$ . Este punto es más general, y suele ser enfatizado en la literatura econométrica (por ejemplo, Heckman y Smith, 1998; Heckman y Vytlačil, 2007a). Las elecciones de los agentes suministran información acerca de las evaluaciones subjetivas, que son de interés independiente.

El problema de la evaluación surge de la ausencia de información acerca de los resultados para el agente  $\omega$  más allá del resultado observado. Aun un experimento social llevado a cabo con toda perfección no puede resolver este problema (Heckman, 1992). La aleatorización identifica sólo un componente de  $\{Y(s, \omega)\}_{s \in S}$  para cualquier agente. Además, aún con muestras grandes y una aleatorización válida, algunos de los  $s \in S$  podrían no ser observados cuando se busca evaluar nuevas políticas que nunca se han experimentado.

Existen dos vías principales para escapar de este problema. El primer camino, caracterizado en los modelos econométricos explícitamente formulados, a menudo identificado como *análisis estructural econométrico*, consiste en modelar  $Y(s, \omega)$  explícitamente en términos de sus determinantes según los especifica la teoría.

Esto conlleva una descripción de las variables aleatorias que caracterizan a  $\omega$  y distinguir cuidadosamente entre lo que conocen los agentes y lo que conoce el analista. Este enfoque también modela  $D(s, \omega)$  y la dependencia entre  $Y(s, \omega)$  y  $D(s, \omega)$  que se produce a partir de variables comunes a  $Y(s, \omega)$  y  $D(s, \omega)$ . El modelo de Roy modela esta dependencia explícitamente. Ver Heckman y Honoré (1990) y Heckman (2001) para una discusión de este modelo. Heckman (1990), Heckman y Smith (1998), Carneiro, Hansen y Heckman (2003) y Cunha, Heckman y Navarro (2005) formulan extensiones del modelo de Roy. Este enfoque enfatiza la comprensión de los factores subyacentes a los resultados y la elección de las ecuaciones de resultados y su dependencia. Los modelos empíricos basados explícitamente en la teoría científica asumen este camino de investigación. Algunos estadísticos llaman a esto el *enfoque científico* y se muestran sorpresivamente hostiles. Ver Holland (1986).

Un segundo camino de escape, que ha sido utilizado por la literatura reciente sobre efectos de tratamiento, consiste en redirigir la atención de los determinantes de  $Y(s, \omega)$  a la estimación de alguna versión poblacional correspondiente con (1), con frecuencia el promedio, sin modelar los factores que producen el resultado ni la relación entre los resultados y el mecanismo que los selecciona. Las valoraciones de los resultados por parte de los agentes no se toman en cuenta. La literatura estadística sobre efectos de tratamiento se enfoca exclusivamente en el problema de políticas P1 para el subconjunto observado de resultados. Ignora el problema de pronosticar una política nueva para un entorno nuevo (problema P2) o una política nunca antes experimentada (problema P3). El pronóstico de los efectos de las políticas nuevas es una tarea central de la ciencia.

#### **4. Nuevos parámetros de tratamientos a nivel poblacional presentes en la economía**

Los economistas y los estadísticos a menudo recurren al mismo conjunto de parámetros de tratamientos a nivel poblacional tales como el efecto de tratamiento promedio (*average treatment effect*, ATE),  $(E(Y(s)-Y(s')))$ , tratamiento sobre los tratados (*treatment on the treated*, TT),  $E(Y(s)-Y(s')) | D(s)=1$ , y tratamiento sobre los no tratados (*treatment on the untreated*, TUT),  $E(Y(s)-Y(s')) | D(s')=1$ ). En esta subsección, mantenemos implícito el " $\omega$ " para simplificar la notación. En ciertas discusiones estadísticas, el ATE es elevado a un sitio de privilegio como *el* parámetro causal. Los economistas utilizan distintos parámetros

causales para distintos problemas de políticas. La distinción entre rendimiento marginal y promedio es un concepto central de la economía. También es de interés en la medicina, donde el efecto de un tratamiento sobre el paciente marginal es un tema importante. Suele ser de interés evaluar el impacto de una extensión (o contracción) marginal de un programa o régimen de tratamiento. El análisis de costo-beneficio incremental se lleva a cabo en términos de ganancias y beneficios marginales.

El *efecto de tratamiento para personas en el margen de indiferencia* (EOTM) entre  $j$  y  $k$ , dado que son las dos mejores opciones disponibles, se define en términos de preferencias personales y en relación con los costos específicos de cada opción  $C(j)$ . En términos formales, se hace explícita la dependencia de la recompensa sobre  $Y(s)$ ,  $C(s)$ , es decir, se denota  $R(Y(s), C(s))$  como la recompensa para el estado  $s$ :

$$(3) \text{EOTM}^R(j, k) = E \left[ Y(j) - Y(k) \left| \begin{array}{l} R(Y(j), C(j)) = R(Y(k), C(k)); \\ R(Y(j), C(j)) \geq R(Y(k), C(k)) \\ R(Y(k), C(k)) \geq R(Y(l), C(l)) \\ l \neq j, k \end{array} \right. \right]$$

Este es el beneficio promedio para los agentes que son indiferentes entre tratamientos  $j$  y  $k$ , dado que estos son las dos mejores opciones disponibles. De forma paralela, se puede definir  $\text{EOTM}^{R_G}(Y(j) - Y(k))$  empleando las preferencias de otro agente (por ejemplo, el padre de un niño, un burócrata paternalista, etc.). Podría ser, por ejemplo, una evaluación subjetiva por parte de un médico. Se pueden definir parámetros análogos para comparaciones *setwise* del promedio (ver Heckman y Vytlačil, 2005, 2007a, b). Una generalización de este parámetro conocido como el efecto de tratamiento marginal (*marginal treatment effect*, ETM) introducida en la literatura de evaluación por Bjorklund y Moffite (1987) y desarrollado posteriormente en Heckman y Vytlačil (1999, 2005, 2007b) y Heckman, Urzua y Vytlačil (2006), juega un papel central en la organización e interpretación de una gran variedad de estimadores econométricos y estadísticos como promedios ponderados de efectos de tratamiento marginales.

Se pueden definir muchos otros parámetros de tratamiento promedio dependiendo de la elección del conjunto de condiciones. Pueden suministrarse definiciones análogas para la mediana y otras versiones cuantiles de estos parámetros (ver Heckman, Smith y Clements, 1997; Abadie, Angrist y Imbens, 2002).

Aunque los promedios son convencionales, también son de considerable interés las distribuciones de los parámetros de tratamiento. En la subsección siguiente se analizan los parámetros de distribución.

Resulta ser de especial interés para el análisis de políticas el *efecto de tratamiento relevante de política* (*policy relevant treatment effect*, PRTE). Es el efecto sobre resultados agregados de un régimen de políticas  $p \in P$  comparado con el efecto de otro régimen de políticas. Según supuestos de invarianza:

$$\text{PRTE: } E_p(Y(s)) - E_{p'}(Y(s)), \quad \text{donde } p, p' \in P,$$

donde las expectativas se toman en distintos espacios de las reglas de asignación de políticas. Heckman y Vytlačil (2007b) enseñan como identificar este parámetro.

Los efectos de tratamiento promedio juegan un papel especial en el enfoque estadístico de la causalidad. Constituyen la pieza central del modelo de Holland (1986) – Rubin (1978) y en muchos otros trabajos en estadística y epidemiología. Estos promedios pueden ser identificados por medio de experimentos sociales con pleno cumplimiento y sin trastorno, debido a una propiedad matemática especial de los promedios. Si se pueden identificar el promedio de  $Y(j)$  y el promedio de  $Y(k)$  mediante un experimento en el cual  $j$  es el tratamiento y  $k$  es la línea de base, se puede formar el efecto de tratamiento promedio para  $j$  comparado con  $k$ . Estos pueden formarse para dos grupos distintos de agentes. Mediante un argumento similar, TT o TUT pueden ser formados por una aleatorización sobre subconjuntos particulares de la población (aquellos quienes seleccionarían el tratamiento y quienes no, respectivamente), suponiendo pleno cumplimiento y la ausencia de sesgo en la aleatorización. Ver Heckman (1992), Heckman y Vytlačil (2007b), y Abbring y Heckman (2007).

El argumento a favor de la aleatorización se debilita cuando al analista le interesan otras mediciones que resumen la distribución, o la distribución misma. En general, la aleatorización no es un procedimiento efectivo para identificar ganancias en el promedio, o la distribución de ganancias o muchos otros parámetros clave. El elevar los promedios de población a categoría de principales parámetros *causales* promueve la aleatorización como método ideal de estimación.

## 5. Distribución de los contrafácticos

Aun cuando los promedios son tradicionales, las respuestas a



muchas preguntas interesantes sobre evaluación requieren conocimiento de las características de la distribución de las ganancias del programa más allá del promedio. De tal forma, es de interés conocer si alguna fracción de la población se beneficia de un tratamiento, aun cuando en promedio se da un beneficio cero o un resultado de promedio negativo. Sea  $s_p$  una notación acotada para la asignación de  $w$  al resultado  $s$  con la política  $p$  y el conjunto asociado de mecanismos de asignación. Para cualesquiera regímenes  $p$  y  $p'$ , la proporción que se beneficia es  $\Pr(Y(s_p(w), w) \geq Y(s_{p'}(w), w))$ . Esto se conoce como el *criterio de votación*. (Heckman, Smith y Clements, 1997). Requiere conocimiento de la distribución conjunta de ambos argumentos de la desigualdad. Sin supuestos adicionales, los experimentos sólo pueden identificar las distribuciones marginales y no las distribuciones conjuntas necesarias para identificar el criterio de votación (Heckman, 1992).

Para tratamientos particulares dentro de un régimen de políticas  $p$  también resulta interesante determinar la proporción que se beneficia de  $j$  comparado con  $k$  como  $\Pr(Y(j, w) \geq Y(k, w))$ . Uno podría estar interesado en las cuantiles de  $Y(s_p(w), w) - Y(s_{p'}(w), w)$  o de  $Y(j, w) - Y(k, w)$  para  $s_p(w) = j$  y  $s_{p'}(w) = k$  o el porcentaje que gana participando en  $j$  (comparado con  $k$ ) con la política  $p$ . Los análisis más completos incluirían costos y beneficios. Los criterios de distribución son especialmente sobresalientes si los beneficios del programa no son transferibles o si existen restricciones a las redistribuciones sociales factibles que impiden el logro de los objetivos redistributivos. Abbring y Heckman (2007) presentan una revisión completa de los enfoques para la identificación de distribuciones conjuntas de resultados contrafácticos.

## 6. Incorporando la incertidumbre

La consideración sistemática de la incertidumbre introduce consideraciones adicionales que son centrales para el análisis económico, pero que en gran medida son ignorados en la literatura estadística actual sobre efectos de tratamiento. Las personas desconocen los resultados asociados con estados aún no experimentados. Si existen resultados potenciales desconocidos al momento de tomar decisiones en cuanto al tratamiento, lo mejor que pueden hacer los agentes es pronosticarlos según alguna regla. Aún si, *ex post*, los agentes conocen su resultado en un estado de comparación y referencia, es posible que no lo conozcan *ex ante* y quizás siempre mantengan la incertidumbre acerca

de lo que hubiesen experimentado en un estado alternativo. Esto da lugar a una distinción adicional: aquella entre evaluaciones *ex post* y *ex ante* de los resultados tanto subjetivos como objetivos. La literatura de inspiración económica sobre evaluación de políticas formula esta distinción. En cambio, no así la literatura estadística sobre efectos de tratamiento.

Dado que los agentes típicamente no poseen información perfecta, es posible que un criterio de votación simple que supone una previsión perfecta acerca de los resultados de las políticas no pronostique las elecciones con precisión, y por tanto requiera de modificaciones. Sea  $I_w$  el conjunto de información al alcance del agente  $w$ . Él o ella evalúa la política  $j$  frente a  $k$  utilizando esa información. Aplicando un criterio de utilidad esperada, el agente  $w$  prefiere la política  $j$  sobre la  $k$  si

$$E(R(Y(j,w),w)|I_w) \geq E(R(Y(k,w),w)|I_w) d\mu(I_w)$$

La proporción de gente que prefiere  $j$  es:

(4)

$$PB(j|j,k) = \int 1(E[R(Y(j,w),w)|I_w]) \geq E(R(Y(k,w),w)|I_w) d\mu(I_w),$$

donde  $\mu(w)$  es la distribución de  $w$  en la población cuyas preferencias acerca de los resultados están siendo estudiadas. El criterio de votación presentado en la sección anterior es el caso especial en donde el conjunto de información  $I_w$  contiene  $(Y(j,w), Y(k,w))$ , de forma que no existe incertidumbre acerca de  $Y(j)$  y  $Y(k)$ . Cunha, Heckman y Navarro (2005), Cunha, Heckman y Navarro (2006) y Abbring y Heckman (2007) ofrecen ejemplos de la aplicación de este criterio. Véase Cunha, Heckman y Navarro (2005, 2006) para cálculos en relación con ambos tipos de distribución conjunta.

El incorporar la incertidumbre al análisis hace que sea esencial distinguir entre evaluaciones *ex ante* y *ex post*. *Ex post*, parte de la incertidumbre acerca de los resultados de la política se resuelve aunque, en general, los agentes no tienen información completa acerca de lo que hubieran sido sus resultados potenciales en regímenes de políticas que no han experimentado, y es posible que sólo tengan información parcial acerca de la política que han experimentado (por ejemplo, la política puede tener consecuencias de largo plazo posteriores al momento de evaluación).

Antes de elegir una actividad, los agentes pueden estar inciertos acerca de los resultados que habrán de producirse.

También pueden tener incertidumbre en relación con el total de los costos que enfrentarán. En general, la información del agente no es la misma que la del analista, y pueden no estar anidadas. El agente puede saber cosas de antemano de las cuales jamás se enterará el analista. Por el contrario, con el beneficio de la retrospectión, el analista puede tener alguna información que el agente desconozca al hacer su elección.

Sea  $I_a$  el conjunto de información que enfrenta el agente en el momento de realizar sus elecciones, y antes de que se produzcan los resultados. Los agentes sólo pueden estimar de forma imperfecta las consecuencias de sus elecciones. Empleando una notación algo no estándar, se puede representar la evaluación de  $s$  por parte de un agente como:

$$R(s, I_a) = \mu_R(s, I_a) + v(s, I_a),$$

reflejando que las valoraciones *ex ante* se hacen con base en información *ex ante*, donde  $\mu_R(s, I_a)$  es determinado en función de variables conocidas por el econometrista y  $v(s, I_a)$  son componentes conocidos por el agente pero no por el econometrista. Las evaluaciones *ex post* también pueden realizarse empleando un conjunto de información diferente  $I_{ep}$  que refleja la incorporación de información posterior al momento de realizar la elección. Es posible que:

$$\operatorname{argmax}_{s \in S} \{R(s, I_a)\} \neq \operatorname{argmax}_{s \in S} \{R(s, I_{ep})\},$$

en cuyo caso puede haberse dado un arrepentimiento o una euforia *ex post* en relación con la elección formulada.

La distinción *ex ante* versus *ex post* es esencial para entender los comportamientos. En los entornos de incertidumbre, las elecciones de los agentes se hacen con base en cálculos *ex ante*. Sin embargo, la literatura sobre efectos de tratamiento generalmente informa acerca de rendimientos *ex post*.

El econometrista puede tener aun otro conjunto de información diferente,  $I_e$ . Las probabilidades de selección calculadas con base en un conjunto de información generalmente difieren de las obtenidas con base en otro. Operando en retrospectiva el econometrista puede tener acceso a información que no está disponible a los agentes cuando ejercen sus opciones.

Evaluaciones *ex post* de un programa a través de encuestas hechas a los agentes que lo han completado pueden discrepar con las evaluaciones *ex ante* del programa. Ambas pueden estar reflejando valoraciones sinceras del programa (Hensher,

Louviere y Swait, 1999; Katz, Gutek, Kahn y Barton, 1975). Éstas son comunicadas cuando los agentes poseen información distinta, o cuando sus preferencias han sido alteradas como resultado de su participación. Antes de participar, los agentes pueden tener dudas acerca de las consecuencias de su participación. Un agente que ha completado el programa  $j$  podrá conocer  $Y(j, w)$  pero solo puede especular acerca del resultado alternativo  $Y(k, w)$  que no ha experimentado. En tal caso, la *satisfacción ex post* de  $j$  relativa a  $k$  para el agente  $w$  quien participa únicamente en  $k$  es sinónima de la siguiente desigualdad:

$$(5) \quad R(Y(j, w), w) \geq E(R(Y(k, w), w) | I_w),$$

en donde la información es postratamiento. Las encuestas con preguntas acerca de la satisfacción del *cliente* acerca del programa logran captar elementos subjetivos de la experiencia del programa que suelen escapar a las mediciones *objetivas* de los resultados que usualmente excluyen los costos y beneficios síquicos. Heckman, Smith y Clements (1997), Heckman y Smith (1998) y Heckman, LaLonde y Smith (1999) presentan evidencias al respecto de este tema. Carneiro, Hansen y Heckman (2001, 2003), Cunha, Heckman y Navarro (2005, 2006) y Heckman y Navarro (2007) desarrollan modelos econométricos para distinguir entre evaluaciones *ex ante* y *ex post* de programas sociales. Ver Abbring y Heckman (2007) para una extensa revisión de esta literatura. Heckman y Vytlačil (2007a) analizan los datos necesarios para identificar estos criterios, y presentan ejemplos de modelos de Roy y sus extensiones que permiten reglas de decisión más generales, así como información imperfecta por parte de los agentes. Muestran cómo usar modelos económicos para la formación de parámetros de tratamiento.

## 7. Un modelo específico

Para cristalizar la discusión en esta sección es útil presentar un modelo econométrico prototípico para la evaluación de políticas. Un paciente puede recibir tratamientos "1" y "0" con resultados  $Y_1(w)$  y  $Y_0(w)$ . Omitimos la notación  $w$  para simplificar la notación.

$Y_1$  es un índice de bienestar del paciente con tratamiento;  $Y_0$  sin tratamiento. En cualquier momento dado, una persona está con tratamiento o sin tratamiento. La decisión de someter a tratamiento puede hacerse con base en los resultados esperados  $E(Y_1|I)$  y  $E(Y_0|I)$  y los costos  $E(C|I)$  donde las expectativas

pertenecen al tomador de decisiones relevante; el paciente, el médico o posiblemente los padres si el paciente es un niño. Los costos podrían ser el dolor y sufrimiento del paciente o los costos médicos directos para el paciente. En cualquier caso, los costos  $C$  y las expectativas  $I$  corresponden al tomador de decisiones relevante que determina quién recibe el tratamiento.

Desde el punto de vista del paciente, la utilidad o el valor esperado del tratamiento es  $E(Y_1|I) - E(C|I)$ . El valor de ningún tratamiento es  $E(Y_0|I)$ . El valor esperado neto es

$$(6) \quad E(Y_1|I) - E(C|I) - E(Y_0|I)$$

Entonces, para los pacientes que eligen un tratamiento basándose en la ganancia máxima:

$$(7) \quad D = \begin{cases} 1, & \text{si } [E(Y_1|I) - E(C|I) - E(Y_0|I)] \geq 0; \\ 0, & \text{de otra forma} \end{cases}$$

o, de forma más sucinta,  $D = 1[(E(Y_1|I) - E(C|I) - E(Y_0|I)) \geq 0]$ . Este es el modelo generalizado de Roy desarrollado por Cunha *et al.* (2005). Ver Heckman y Vytlacil (2007a) para una visión general de tales modelos.

Si es el médico quien toma la decisión de tratar, entonces los  $C$  e  $I$  relevantes son los del médico. En lugar de un medicamento, el tratamiento puede ser educación, migración o la instalación de una tecnología, y los resultados potenciales son los contrafácticos con o sin tratamiento. El efecto de tratamiento *ex post* es  $Y_1 - Y_0$ . El efecto *ex ante* es  $E(Y_1|I) - E(Y_0|I)$ .

La construcción de  $(Y_0, Y_1)$  y la decisión de asignar tratamientos está motivada por la teoría científica o de comportamiento. El modelo más básico en la economía es el modelo de Roy ya mencionado. La información del tomador de decisiones es perfecta. No hay costos directos de tratamiento ( $C = 0$ ) y la regla de decisión es:

$$(8) \quad D = 1(Y_1 \geq Y_0)$$

Los que reciben tratamiento son los que se benefician de él. De tal forma que los tratados constituyen una muestra no aleatoria de la población en general, y se produce un sesgo de selección al utilizar la muestra tratada para inferir lo que experimentarían una persona promedio seleccionada al azar.

El enfoque econométrico modela la dependencia entre  $Y$  observado  $= D(Y_1) + (1-D)Y_0$  y  $D$  para sugerir estimadores alternativos para la identificación de parámetros causales. Hay trabajos

recientes que identifican varios efectos de tratamiento promedio, distribuciones de efectos de tratamiento y el costo de tratamiento, incluyendo el dolor y sufrimiento de los pacientes. Algunas especificaciones comúnmente usadas son:

$$(9) \quad \begin{aligned} Y_1 &= X\beta_1 + U_1 \\ Y_0 &= X\beta_0 + U_0 \\ C &= Z\gamma + U_C \end{aligned}$$

donde  $(X, Z)$  son observados por el analista y  $Y(U_1, U_0, U_C)$  no son observados. El paciente puede conocer más o menos que el analista. Los modelos econométricos permiten al paciente saber más (observar más) que el analista, y analizar la selección de pacientes para los tratamientos tomando en cuenta la asimetría de información entre el paciente y el analista. (El cotejamiento supone que, condicionado a  $X$  y  $Z$ ,  $D$  es independiente de  $Y_0, Y_1$  y por tanto asume que hay mucha información disponible para el analista). El modelo de Roy fija  $\gamma = 0$ ,  $U_C = 0$  y supone normalidad para  $(U_0, U_1)$ . Estos supuestos paramétricos y de distribución se relajan en la literatura econométrica reciente (ver Heckman y Vytlačil, 2007a para un resumen).

El enfoque estadístico no modela la regla de asignación de tratamientos ni su relación con los resultados potenciales. El enfoque econométrico hace de la ecuación de asignación de tratamientos la pieza central de su enfoque y considera valoraciones tanto objetivas como subjetivas, así como resultados *ex ante* ( $E(Y_1|I)$ ,  $E(Y_0|I)$ ,  $E(C|I)$ ) y *ex post* ( $Y_1, Y_0, C$ ). Para este modelo, el EOTM es  $E(Y_1 - Y_0 | E(Y_1|I) - E(Y_0|I) - E(C|I) = 0)$ , es decir el beneficio para las personas en el punto de indiferencia entre tratamiento y no tratamiento.

#### IV. CONTRAFÁCTICOS, CAUSALIDAD Y MODELOS ECONOMÉTRICOS ESTRUCTURALES

La literatura estadística sobre evaluación de políticas a veces compara enfoques econométricos *estructurales* con modelos de *efectos de tratamiento* o *causales* (ver, por ejemplo, Angrist y Imbens, 1995; Angrist, Imbens y Rubin, 1996). La comparación no es clara porque los términos no están definidos con precisión. Heckman (2005) y Heckman y Vytlačil (2007a) definen formalmente los modelos estructurales y los emplean como mecanismos para la generación de contrafácticos. Consideran ecuaciones tanto de resultados como de elección de tratamientos.

Esta sección presenta una breve introducción al enfoque econométrico y lo compara con los modelos de inferencia causal en la estadística.

## 1. Generando contrafácticos

El modelo tradicional de la econometría es el modelo de *todas las causas*. Escribe los resultados como un mapeo determinístico de insumos a productos:

$$(10) \quad y(s) = g_s(x, u_s),$$

donde  $x$  y  $u_s$  son variables fijas especificadas por la teoría económica relevante. La notación nos anticipa la distinción entre observable ( $x$ ) y no observable ( $u_s$ ) que es muy importante en la instrumentación empírica. El papel de los dos tipos de variables en (10) es simétrico. Esta notación permite que distintas no observables  $u_s$  afecten diferentes resultados.  $\mathbf{D}$  es el dominio del mapeo  $g_s: \mathbf{D} \rightarrow \mathbf{R}^y$ , donde  $\mathbf{R}^y$  es el ámbito de  $y$ . Puede haber múltiples variables de resultado. Todos los resultados tienen una explicación en el sentido funcional mediante los argumentos de  $g_s$  en (10). Si se modelan las realizaciones *ex post* de los resultados, es enteramente razonable invocar un modelo de *todas las causas* ya que las realizaciones son conocidas (*ex post*) y toda incertidumbre ha sido resuelta. En la definición de una función está implícito el requisito de que  $g_s$  sea *estable* o *invariable* frente a cambios en  $x$  y  $u_s$ . La función  $g_s$  se mantiene estable frente a los cambios en sus argumentos. La invarianza es una propiedad clave del modelo causal.

La ecuación (10) es una función de producción que vincula insumos (factores) con productos.  $g_s$  mapea  $(x, u_s)$  al ámbito de  $y$  o la imagen de  $\mathbf{D}$  según  $g_s$ , en donde el dominio de definición  $\mathbf{D}$  puede diferir del soporte empírico. De tal forma, la ecuación (10) mapea los insumos admisibles hacia posibles resultados *ex post*. Esta notación permite que distintas no observables de una lista en común  $u$  aparezcan en diferentes ecuaciones de resultado.

Una versión *estructural profunda* de (10) modela la variación de los  $g_s$  en términos de  $s$  como un mapa construido por medio de la generación de características  $q_s$ ,  $x$  y  $u_s$  en los resultados:

$$(11) \quad y(s) = g(q_s, x, u_s),$$

donde ahora el dominio de  $g$ ,  $\mathbf{D}$ , se define para  $q_s$ ,  $x$ ,  $u_s$  tal que  $g: \mathbf{D} \rightarrow \mathbf{R}^y$ . Los componentes  $q_s$  proveen la base para generar los

contrafácticos a través de tratamientos procedentes de un conjunto básico de características. La función  $g$  mapea  $(q_s, s, u_s)$  a el ámbito de  $y$ ,  $g: (q_s, x, u_s) \rightarrow \mathbf{R}^y$ , donde el dominio de definición  $D$  de  $g$  puede diferir del soporte empírico. En esta especificación, los distintos tratamientos  $s$  se caracterizan por diferentes paquetes compuestos por los elementos de un conjunto de características comunes a todos los tratamientos. Este marco suministra la base para resolver el problema de política P3 dado que las políticas nuevas (tratamientos) se generan a partir de características comunes, y todas las políticas son puestas en una base común. Si una política nueva se caracteriza por transformaciones conocidas de  $(q_s, x, u_s)$  que yacen en el dominio de definición de  $g$ , el problema P3 de pronóstico de políticas puede ser resuelto. El argumento de los mapas  $g$  y  $g_s$  son parte de la especificación *a priori* de un modelo causal. Los analistas pueden discrepar en torno a los argumentos apropiados para incluir en estos mapas.

Un beneficio de un enfoque de efectos de tratamiento que se centra en el problema P1 es que funciona exclusivamente con resultados en lugar de insumos. Sin embargo, no aporta nada para resolver los problemas P2 y P3, ni suministra una base para interpretar los efectos de tratamiento a nivel de la población.

Considérense modelos alternativos de resultados escolares para alumnos donde  $s$  cataloga el tipo de educación (por ejemplo, pública normal, pública especial, privada laica, privada religiosa). Las  $q_s$  son las características observadas de los colegios de tipo  $s$ . Las  $x$  son las características observadas del alumno. Las  $u_s$  son las características no observadas tanto de los colegios como del alumno. Si se pudiera caracterizar un nuevo tipo de colegio propuesto como un paquete nuevo de diferentes niveles de los mismos ingredientes  $x$ ,  $q_s$  y  $u_s$ , al tiempo que se pudiera identificar (11) sobre el dominio de la función definida por el nuevo paquete, se puede resolver el problema P3. Si el mismo insumo escolar (mismo  $q_s$ ) se aplicara a distintos estudiantes (aquellos con distinto  $x$ ) y se pudiese identificar (10) u (11) sobre el nuevo dominio de definición, se puede resolver el problema P2. Al profundizar en las *causas de los efectos* se puede hacer más que simplemente comparar entre sí, los efectos de los tratamientos existentes. Además, el modelar el  $u_s$  y su relación con las no observables en la ecuación de elección de tratamiento es altamente informativo en cuanto a la elección de las estrategias de identificación apropiadas.

Las ecuaciones (10) y (11) se conocen a veces como funciones causales Marshallianas. Suponiendo que los componentes de  $(x, u_s)$  o  $(q_s, x, u_s)$  están libres de variación, una característica que



puede o no responder a la teoría relevante, se puede variar cada argumento de estas funciones para lograr un efecto causal, *ceteris paribus*, del argumento sobre el resultado. (Ver Heckman y Vytlačil, 2007a para una definición precisa de libre de variación.) Algunos componentes pueden estar libres de variación, y otros no. Estos experimentos mentales se efectúan para variaciones hipotéticas. Recuérdese que la teoría *a priori* especifica los argumentos de las funciones causales y la lista de los elementos que se mantienen fijos cuando se manipula una variable.

Cuando se modifica una variable mientras las demás se mantienen fijas, se produce un efecto causal Marshalliano, *ceteris paribus*, de un cambio en dicha coordenada sobre las variables de resultado. La variación de  $q_s$  fija distintos niveles de tratamiento. Variaciones de  $u_s$  entre agentes explica por qué personas con las mismas características  $x$  responden de manera distinta al mismo tratamiento  $s$ . No es necesario que la variación, *ceteris paribus*, sea para una sola variable de la ecuación. Un tratamiento generalmente consiste en un paquete de características, y si se varía el paquete de  $q_s$  a  $q_{s'}$  se obtienen diferentes efectos de tratamiento.

A continuación se emplea la convención de que los valores en minúscula se usan para definir valores fijos mientras que la mayúscula denota variables aleatorias. Al definir (10) y (11) hemos trabajado explícitamente con variables fijas manipuladas de forma hipotética como en el álgebra de la física elemental. En un mundo puramente determinístico, los agentes responden ante estas variables no estocásticas. Aun si el mundo es incierto, *ex post* y luego de la realización de la incertidumbre, los resultados de insumos inciertos son determinísticos. Algunos componentes de  $u_s$  pueden ser choques aleatorios que se producen con posterioridad a la toma de decisiones acerca del tratamiento.

De tal forma que si el entorno se caracteriza por esta incertidumbre, (10) y (11) pueden interpretarse como realizaciones *ex post* del contrafáctico a medida que la incertidumbre se resuelve. Las versiones *ex ante* pueden ser diferentes. Desde el punto de vista del agente  $w$  con el conjunto de información  $I_w$ , el valor esperado *ex ante* de  $Y(s, w)$  es:

$$(12) \quad E(Y(s, w) | I_w) = E(g(Q(s, w), X(w), U(s, w)) | I_w),$$

donde  $Q(s, w)$ ,  $X(w)$ ,  $U(s, w)$  son variables aleatorias generadas por una distribución que depende del conjunto de información del agente indexada por  $I_w$ .

La expectativa podría calcularse utilizando el conjunto de información del tomador de decisiones relevante (por ejemplo, los padres para el caso de resultados del hijo) que no tiene por qué ser el agente cuyos resultados son objeto de medición. Estas variables aleatorias se extraen de la distribución subjetiva del agente  $w$ . Esta distribución puede diferir de la distribución producida por la realidad o la naturaleza si las expectativas del agente difieren de la realidad objetiva. En la presencia de incertidumbre intrínseca, el tomador de decisiones relevante actúa sobre (12) pero el contrafáctico *ex post* es:

$$(13) \quad Y(s, w) = E(Y(s, w) | I_w) + v(s, w),$$

donde  $v(s, w)$  satisface  $E(v(s, w) | I_w) = 0$ . Con esta interpretación, el conjunto de información del agente  $w$  es parte de la especificación del modelo pero las realizaciones provienen de una distribución de probabilidades, y el conjunto de información incluye la tecnología  $g$ . Esta representación clarifica la distinción entre resultados determinísticos *ex post* y resultados intrínsecamente aleatorios *ex ante*. Abbring y Heckman (2007) presentan un resumen de los modelos econométricos de evaluación que incorporan la incertidumbre.

Esta reexpresión del modelo básico determinístico concilia el modelo de *todas las causas* (10) y (11) con el modelo de incertidumbre intrínseca preferido por algunos estadísticos (ver, por ejemplo, Dawid, 2000, y la discusión que siguió a su trabajo). Hay incertidumbre *ex ante* a nivel del agente  $w$  individual, pero no *ex post*. Las realizaciones de  $v(s, w)$  son ingredientes del modelo de todas las causas *ex post*, pero no son parte del modelo subjetivo de todas las causas *ex ante*. La ley de probabilidad usada por el agente para computar las expectativas de  $g(Q(s, w), X(w), U_s(w))$  puede ser distinta de la distribución objetiva que genera los datos observados. En el modelo de todas las causas *ex ante*, las manipulaciones de  $I_w$  definen los efectos causales *ex ante*.

De tal forma que, desde el punto de vista del agente, se pueden variar los elementos de  $I_w$  para producir funciones de respuesta causales *ex ante* Marshallianas. El efecto de tratamiento *ex ante* desde el punto de vista del agente para los tratamientos  $s$  y  $s'$  es:

$$(14) \quad E(Y(s, w) | I_w) - E(Y(s', w) | I_w)$$

Los datos utilizados para determinar estas funciones pueden ser de apoyo limitado. En tal caso los analistas no pueden

identificar plenamente las relaciones teóricas en los dominios hipotéticos de definición. Además, en materia de soporte, los componentes de  $X$ ,  $U_s$  y  $I_w$  pueden no estar libres de variación, aun si lo están en el dominio hipotético de definición de la función. Si las  $X$  de la muestra son funcionalmente dependientes, no será posible identificar la función causal Marshalliana con respecto a variaciones en  $x$  sobre el soporte disponible aun si uno puede imaginar hipotéticamente como variar los componentes de  $x$  sobre los dominios de definición de las funciones (10) u (11).

Continúo con una distinción importante entre fijar y condicionar factores que va directamente al corazón de la distinción entre modelos causales y relaciones de correlación. Este aspecto es independiente de cualquier problema con los soportes de la muestra comparados con los dominios de definición de las funciones.

## 2. Fijar vs. condicionar

La distinción entre fijar y condicionar insumos es central para distinguir los verdaderos efectos causales de efectos causales espurios. En un estudio importante, Haavelmo (1943) hizo esta distinción para los modelos de ecuaciones lineales. La distinción de Haavelmo es la base del libro de Pearl (2000) sobre causalidad donde generaliza el análisis de Haavelmo a estructuras no lineales. Pearl define un operador “hacer” para representar el acto mental de fijar una variable, para distinguirlo del acto de condicionarla, que constituye una operación estadística. Si el conjunto de condicionamiento es lo suficientemente rico, fijar y condicionar resulta lo mismo en un modelo de todas las causas *ex post*. Pearl sugiere un mecanismo físico específico para fijar variables y hacer operativa la causalidad, pero ello no es central a su definición, o cualquier otra, de causalidad.

La distinción entre fijar y condicionar se ve más fácilmente en el modelo de regresión lineal analizado por Haavelmo (1943). Que sea  $y = x\beta + u$ . Mientras que  $y$  y  $u$  son escalares,  $x$  puede ser un vector. La ecuación lineal mapea cada pareja  $(x, u)$  hacia un escalar  $y \in \mathbb{R}$ . Supongamos que el soporte en los datos de la variable aleatoria  $(X, U)$  es el mismo que el dominio de los  $(x, u)$  que están fijos en el experimento mental hipotético, y que los  $(x, u)$  están libres de variación (es decir, pueden variar en forma independiente coordenada por coordenada). Hagamos abstracción del problema de soporte limitado analizado en la sección anterior. Omitiendo la notación “ $w$ ” para variables aleatorias, queda:

$$Y = X\beta + U.$$

Se supone que  $U$  posee un promedio finito. Entonces la *naturaleza* o *el mundo real* escoge  $(X, U)$  para determinar  $Y$ . El analista observa  $X$  pero no observa  $U$ , y los  $(X, U)$  son variables aleatorias. Este es un modelo de todas las causas en que  $(X, U)$  determina  $Y$ . La variación generada por el modelo hipotético varía una coordenada de  $(X, U)$ , fijando todas las demás coordenadas, para producir el efecto de variación sobre el resultado  $Y$ . La naturaleza (en oposición al modelo) quizás no permita tal variación.

Formalmente, este modelo definido a nivel poblacional, puede expresarse como la expectativa condicional,

$$E(Y|X = x, U = u) = x\beta + u.$$

Dado que la condición es tanto sobre  $X$  como sobre  $U$ , no existen fuentes adicionales de variación en  $Y$ . Este es un modelo determinístico que coincide con el modelo de todas las causas. Así en el tema del soporte, que también se supone como dominio de definición de la función, este modelo es igual al modelo determinístico hipotético  $y = x\beta + u$ . Fijar  $X$  en distintos valores corresponde con realizar distintos experimentos mentales con las  $X$ . En este caso, fijar y condicionar equivalen a lo mismo.

Si sólo se condicionan las  $X$ , sin embargo, se obtiene

$$(15) \quad E(Y|X = x) = x\beta + E(U|X = x).$$

Esta relación no genera relaciones  $(X, Y)$  con  $U$  constantes. Sólo genera una relación con  $X$  constante. Salvo que se condicionen todas las causas (las variables del lado derecho), la relación empírica (15) no identifica un efecto causal de  $X$  sobre  $Y$ . La variación en  $X$  también modifica el promedio condicional de  $U$ , dada  $X$ . Este análisis puede generalizarse a un modelo no lineal  $y = g(q, x, u)$  (ver Pearl, 2000). Puede generalizarse para tomar en cuenta la resolución temporal de la incertidumbre si se incluye  $v(s, \omega)$  como un argumento en el modelo causal *ex post*. Los resultados pueden incluir tanto resultados objetivos  $Y(s, \omega)$  como subjetivos  $R(Y(s, \omega), \omega)$ .

En paralelo a los modelos causales para resultados existen los modelos causales para la elección del tratamiento (ver Heckman y Vytlačil, 2007a). Explicar la incertidumbre y las valoraciones subjetivas de los resultados (por ejemplo, dolor y sufrimiento en el caso de un tratamiento médico) es la contribución muy importante del enfoque econométrico (ver, por ejemplo, Carneiro,

Hansen y Heckman, 2003; Chan y Hamilton, 2006; Cunha, Heckman y Navarro, 2005, 2006; Cunha y Heckman, 2007; y Heckman y Navarro, 2007). Los factores que llevan a un agente a participar en el tratamiento  $s$  pueden depender de los factores que afectan los resultados. Modelar esta dependencia es una fuente principal de información utilizada por el enfoque econométrico para la construcción de contrafácticos a partir de datos reales. Se puede formular un análisis paralelo para el caso en donde el tomador de decisiones no es el mismo agente cuyos resultados son objeto de evaluación.

### 3. El modelo econométrico vs. el modelo Neyman-Rubin

Muchos estadísticos y científicos sociales emplean un modelo de contrafácticos y causalidad atribuido a Donald Rubin por Paul Holland (1986). La estadística de este marco fue desarrollada por Neyman (1923), Cox (1958) y otros. Se desarrollaron marcos paralelos en forma independiente en la psicometría (Thurstone, 1927) y la economía (Haavelmo, 1943; Quandt, 1958, 1972; Roy, 1951). La literatura estadística sobre efectos de tratamiento se origina en el diseño de experimentos. Se nutre de experimentos hipotéticos para definir la causalidad, y de esa manera crea la impresión entre muchos de sus usuarios que la asignación aleatoria es el método más convincente para identificar modelos causales. Algunos dirían que es la única manera de identificar modelos causales.

Neyman y Rubin postulan contrafácticos  $\{Y(s, \omega)\}_{s \in S}$  sin modelar los factores que determinan los  $Y(s, \omega)$  como se hace en el enfoque econométrico (ver ecuaciones (10) – (13)). Rubin y Neyman no ofrecen un modelo de elección donde se seleccione el resultado. No se especifica explícitamente ningún modelo de todas las causas, en minúscula, en este enfoque, ni existe discusión de la ciencia social o la teoría que produce los resultados estudiados.

Con esta notación, Rubin (1986) invoca versiones de los supuestos tradicionales de invarianza econométrica conocidos como SUTVA (*Stable Unit Treatment Value Assumption*). En vista de que no desarrolla ecuaciones de elección ni evaluaciones subjetivas, no considera las condiciones de invarianza más generales para las evaluaciones tanto objetivas como subjetivas tratadas en la sección III-2. Los supuestos de invarianza fueron desarrollados mediante procedimientos econométricos de la *Cowles Commission* y formalizados por Hurwicz (1962). Se revisan en Heckman y Vytlačil, (2007a).

El modelo de Rubin postula:

- (R-1)  $\{Y(s, \omega)\}_{s \in S}$ , un conjunto de contrafácticos definido por resultados *ex post*. No analiza las valoraciones de los resultados por parte de los agentes ni especifica de manera explícita las reglas de selección, salvo contrastando la aleatorización con la no aleatorización;
- (R-2) Invarianza de contrafácticos para resultados objetivos con respecto al mecanismo de asignación dentro de un régimen de políticas;
- (R-3) Ausencia de interacciones sociales o efectos de equilibrio general para los resultados objetivos; y
- (R-4) Inexistencia de simultaneidad de efectos causales, es decir, los resultados no pueden causarse recíprocamente.

Dos supuestos adicionales que están implícitos en la aplicación del modelo son que PI es el único problema de evaluación que se aborda, y que efectos causales promedio son los únicos temas de interés.

El enfoque econométrico considera una colección más amplia de problemas de políticas que el enfoque estadístico de efectos de tratamiento. Sus características más representativas son:

- i) El desarrollo de un marco explícito para resultados  $Y(s, \omega)$ ,  $s \in S$ , mediciones y elección de resultados donde se desarrolla explícitamente el papel de las no observables (variables ausentes) en la creación de problemas de selección y la justificación de estimadores.
- ii) El análisis de las evaluaciones subjetivas de los resultados  $R(s, \omega)$ ,  $s \in S$  y el uso de datos de elección y cumplimiento para su inferencia.
- iii) El análisis *ex ante* y *ex post* de las realizaciones y evaluaciones de los tratamientos. Este análisis permite a los analistas modelar e identificar arrepentimientos y expectativas de los agentes. Los ítems (ii) y (iii) introducen la toma de decisiones por parte de los agentes en la literatura sobre efectos de tratamiento.
- iv) El desarrollo de modelos para identificar y evaluar distribuciones enteras de efectos de tratamiento (*ex ante* y *ex post*) en lugar de solo parámetros promedio tradicionales. Estas distribuciones permitan a los analistas determinar la proporción de personas que se benefician del tratamiento, un parámetro

causal que no considera la literatura estadística sobre efectos de tratamiento.

- v) Los modelos de causalidad simultánea.
- vi) La definición de parámetros sin necesidad de apelar a manipulaciones experimentales hipotéticas.
- vii) La clarificación de la necesidad de invarianza paramétrica en relación con las diferentes clases de manipulaciones para responder a diferentes clases de interrogantes.

A continuación amplío estos puntos.

Los modelos de selección definidos para resultados potenciales con mecanismos de asignación de tratamientos explícitos fueron desarrollados por Gronau (1974) y Heckman (1974, 1976, 1978, 1979) en la literatura económica antes de popularizarse el modelo Neyman-Rubin en la estadística. La literatura econométrica sobre elección discreta (McFadden, 1974, 1981) emplea utilitarios contrafácticos o evaluaciones subjetivas tal como lo hizo su literatura de inspiración en psicología matemática (Thurstone, 1927, 1959). El modelo esbozado en la sección III-7 considera tanto elecciones como los resultados de las elecciones. A diferencia del modelo Neyman-Rubin, estos modelos no parten con el experimento como un ideal sino que comienzan con modelos bien planteados y claramente articulados para resultados y elección de tratamientos, donde se hacen explícitas las no observables que subyacen el problema de selección y evaluación. Las manipulaciones hipotéticas definen los parámetros causales del modelo. La aleatorización es un metáfora, y no un ideal o *patrón oro*.

En contraste al modelo econométrico, la definición de efectos causales de Holland (1986)-Rubin (1978) se basa en la aleatorización. El análisis en los trabajos de Rubin de 1976 y 1978 es una dicotomía entre la aleatorización (ignorancia) y la no aleatorización, y no un tratamiento explícito de mecanismos de selección particulares para el caso no aleatorio como se desarrolló en la literatura econométrica. No hay una discusión explícita de reglas de selección de tratamiento como en la ecuación (8) y su relación con los resultados obtenidos. Aun en condiciones ideales, la aleatorización no puede dar respuesta a ciertas preguntas muy básicas, como qué proporción de la población se beneficia de un programa (Heckman, 1992). Véase Carneiro, Hansen y Heckman (2001, 2003) en donde esta proporción se identifica utilizando datos de elección o mediciones *proxy* complementarias. Véase también Cunha, Heckman y Navarro

(2005, 2006) y Cunha y Heckman (2007). Abbring y Heckman (2007) analizan este trabajo. En la práctica, los efectos de contaminación y de traspaso hacen de la aleatorización una solución poco confiable aun para construir  $E(Y_1 - Y_0)$ . Véase las evidencias de sesgo de interrupción y sesgo de contaminación que surgen en pruebas aleatorizadas presentadas en Heckman, LaLonde y Smith (1999) y en Heckman, Hohmann, Smith y Khoo (2000).

Muchos analistas de causalidad influyentes mezclan los tres puntos del cuadro 1. El análisis de Holland (1986) ilustra este aspecto, así como el papel central de la prueba aleatorizada en el análisis de Holland-Rubin. Después de explicar el *modelo de Rubin*, Holland ofrece una ilustración muy reveladora que combina las dos primeras tareas del cuadro 1. Afirma que no puede haber un efecto causal del género sobre los ingresos, porque los analistas no pueden asignar el género al azar. Esta aseveración confunde el acto de definir un efecto causal (un acto puramente causal) con las dificultades empíricas de estimarlo. Estas constituyen las tareas 1 y 2 del cuadro 1.

Como otro ejemplo del mismo punto, Rubin (1978, p. 39) niega que sea posible definir un efecto causal del género sobre la inteligencia dado que *en principio* no se puede llevar a cabo una aleatorización. “A falta de definiciones de tratamiento que especifiquen las acciones a realizarse sobre las unidades experimentales, no puedo hablar sin ambigüedad acerca de los efectos causales de tratamientos” (Rubin, 1978, p. 39). En este y muchos otros pasajes de la literatura estadística se define un efecto causal con base en una aleatorización. Se confunden los temas de definición e identificación. Esta confusión continúa floreciendo en la literatura de estadísticas aplicadas. Por ejemplo, Berk, Li y Hickman (2005) hacen eco de Rubin y Holland al insistir en que si no se puede en principio llevar a cabo un experimento, entonces no puede definirse un efecto causal. El parámetro LATE (*local average treatment effect*) de Imbens y Angrist (1994) se define por un instrumento y mezcla las tareas 1 y 2 (definición e identificación). Imbens y Angrist (1994) utilizan variables instrumentales como sustitutos de la aleatorización. Heckman y Vytlačil (1999, 2005), y Heckman, Urzua y Vytlačil (2006) definen de forma abstracta al parámetro LATE y separan los temas de la definición e identificación de parámetros.

El acto de definir es por lógica distinto de los actos de identificar e inferir. Un acto puramente mental puede definir un efecto causal del género. Esta es una tarea diferente a la de identificar el efecto causal. La afirmación de que la causalidad



sólo puede determinarse mediante la aleatorización equivale a materializar este concepto como el *patrón oro* de la inferencia causal.

El enfoque econométrico de la inferencia causal es más comprensivo que el modelo Neyman-Rubin de contrafácticos. Analiza modelos de la elección de contrafácticos  $\{D(s, \omega)\}$ ,  $s \in S$ , así como la relación entre las ecuaciones de elección y los contrafácticos. Los  $D(s, \omega)$  se modelan explícitamente como generados por la colección de variables aleatorias  $(Q(s, \omega), C(s, \omega), Y(s, \omega) \mid I_\omega)$ ,  $s \in S$  donde  $Q(s, \omega)$  es la característica del tratamiento  $s$  para el agente  $\omega$ ,  $C(s, \omega)$  son los costos,  $\{Y(s, \omega)\}_{s \in S}$  son los resultados y el “ $\mid$ ” denota que estas variables se definen condicionadas a  $I_\omega$ , (el conjunto de información del agente). (Recuérdese la discusión de la sección III-7.) Si otros agentes toman decisiones de asignación de tratamientos, entonces los determinantes de  $D(s, \omega)$  se modifican según lo que contenga su conjunto de información. Las variables que determinan las elecciones son analizadas.

El modelar los  $Y(s, \omega)$  en términos de las características de los tratamientos y de los tratados facilita la comparación de contrafácticos y los efectos causales derivados a través de estudios donde pueden variar la composición de los programas y los integrantes del grupo de tratamiento. También facilita la construcción de contrafácticos para poblaciones y políticas nuevas. El marco de Neyman-Rubin enfoca exclusivamente en los *efectos causales* promedio a nivel poblacional o efectos de tratamiento para las políticas implementadas en la realidad, sin ofrecer un marco para la extrapolación de hallazgos a nuevos entornos o el pronóstico de nuevas políticas (problemas P2 y P3).

#### 4. Causalidad simultánea

Una de las limitaciones principales del modelo Neyman-Rubin es que es recursivo. No modela los efectos causales de resultados que ocurren simultáneamente. Desde Haavelmo (1943, 1944) los econométricos han empleado la teoría de ecuaciones simultáneas para definir la causalidad en modelos no recursivos donde las causas son simultáneas e interdependientes. Heckman (2005), y Heckman y Vytlačil (2007a) presentan un extenso análisis de la causalidad simultánea.

Considérese el siguiente sistema de ecuaciones simultáneas no lineales donde la identificación de los efectos causales puede definirse mediante la exclusión de variables. Sean  $(Y_1, Y_2)$  una pareja de variables (internas) determinadas conjuntamente, con

las variables  $X_1, X_2, U_1, U_2$  especificadas externamente. Las variables especificadas externamente son aquellas que se especifican independientemente del sistema Según análisis. El sistema que determina las variables internas puede representarse como:

$$(16) \quad Y_1 = g_1(Y_2, X_1, X_2, U_1)$$

$$(17) \quad Y_2 = g_2(Y_1, X_1, X_2, U_2)$$

$\left. \frac{\partial g_1}{\partial Y_2} \right|_{Y_2=y_2, X_1=x_1, X_2=x_2, U_1=u_1}$  es el efecto causal de  $Y_2$  sobre  $Y_1$  man-

teniendo fijos a  $X_1, X_2, y U_1$ . Esta es la definición de Haavelmo del efecto causal aplicado a un sistema de ecuaciones simultáneas. Si se supone la existencia de soluciones locales, se pueden resolver estas ecuaciones para obtener las variables internas en términos de las variables externas  $Y_1 = \varphi_1(X_1, X_2, U_1, U_2)$  y  $Y_2 = \varphi_2(X_1, X_2, U_1, U_2)$ . Estas funciones pueden ser determinadas a partir de los datos con condiciones estándar de exogeneidad para  $X$  (ver, por ejemplo, Amemiya, 1985). Mediante la regla de la cadena se puede definir el efecto causal de  $Y_2$  sobre  $Y_1$  empleando la exclusión  $\left( \frac{\partial g_1}{\partial Y_2} = 0, \text{ para todo } (Y_2, X_1, X_2, U_1) \right)$  e identificar el efecto causal de  $Y_2$  sobre  $Y_1$  por:

$$\frac{\partial g_1}{\partial Y_2} = \frac{\partial Y_1}{\partial X_1} / \frac{\partial Y_2}{\partial X_1} = \frac{\partial \varphi_1}{\partial X_1} / \frac{\partial \varphi_2}{\partial X_1}$$

Se pueden identificar los efectos causales de  $Y_1$  sobre  $Y_2$  empleando parciales con respecto a  $X_2$  si existe exclusión en relación con  $X_2$  en la ecuación (17).

Un ejemplo simple está dado por el modelo de ecuaciones simultáneas de Haavelmo (1944):

$$(18) \quad Y_1 = \gamma_{12}Y_2 + b_{11}X_1 + b_{12}X_2 + U_1$$

$$(19) \quad Y_2 = \gamma_{21}Y_1 + b_{21}X_1 + b_{22}X_2 + U_2,$$

Supóngase que  $(U_1, U_2)$  es independiente de  $(X_1, X_2)$ . Según la condición de que:

$$\begin{bmatrix} 1 & -\gamma_{12} \\ -\gamma_{21} & 1 \end{bmatrix},$$

sea de rango completo, se puede resolver para las variables internas ( $Y_1, Y_2$ ) como función de las variables externas ( $X_1, X_2, V_1, V_2$ ):

$$(20) \quad Y_1 = \pi_{11}X_1 + \pi_{12}X_2 + V_1$$

$$(21) \quad Y_2 = \pi_{21}Y_1 + \pi_{22}X_2 + V_2,$$

Donde se supone que  $(X_1, X_2)$  es estadísticamente independiente de  $(V_1, V_2)$ . Se pueden estimar los  $\pi_{ij}$ ,  $i, j = 1, 2$  por mínimos cuadrados ordinarios. Si  $b_{11} = 0$ , entonces  $X_1$  sólo afecta a  $Y_1$  a través de su efecto sobre  $Y_2$  (ver ecuación (18)). De (20) y (21) obtenemos:

$$\frac{\partial Y_1}{\partial X_1} = \pi_{11}, \quad \frac{\partial Y_2}{\partial X_1} = \pi_{21}.$$

De la definición de  $\pi_{11}$ ,  $\pi_{21}$  como coeficientes de la solución de las variables externas en términos de las variables internas, se obtiene que:

$$\frac{\pi_{11}}{\pi_{21}} = \gamma_{12}.$$

De esta forma se puede definir e identificar el efecto causal de  $Y_2$  sobre  $Y_1$ . Este es el método desarrollado por Tinbergen (1930), y que es conocido como mínimos cuadrados indirectos cuando se aplica como método de estimación.

La intuición es simple. Dado que  $X_1$  se excluye de (18) (o (16)), puede utilizarse para desplazar  $Y_2$  en dicha ecuación manteniendo fijos  $X_2$  y  $U_1$ . Esto le permite al analista determinar el efecto causal de  $Y_2$  sobre  $Y_1$ , manteniendo fijos los demás determinantes de  $Y_1$ .

Esta definición de efectos causales en un sistema interdependiente generaliza las definiciones recursivas de causalidad que figuran en la literatura estadística de efectos de tratamiento (Pearl, 2000). La clave de la definición econométrica más general es la manipulación de insumos externos y la exclusión, no un método particular como la aleatorización, el cotejamiento o las variables instrumentales. Se puede usar el modelo de ecuaciones simultáneas de población para definir el tipo de variaciones admisibles y atender los problemas de definiciones (tarea 1 del cuadro 1). Si, para un modelo, dado las funciones (16) o (17) se desplazan al manipular las variables externas, o si las variables externas no pueden manipularse de forma independiente,

los efectos causales de una variable interna sobre otra no pueden definirse dentro de dicho modelo.

**V. LA MÁXIMA DE MARSCHAK Y LA RELACIÓN ENTRE  
LA LITERATURA ECONOMETRICA Y LA LITERATURA  
ESTADÍSTICA SOBRE EFECTOS DE TRATAMIENTO:  
UNA SÍNTESIS**

La ausencia de modelos explícitos de resultados y elecciones es una característica prominente de la literatura estadística sobre efectos de tratamiento. Los modelos bien planteados desde el punto de vista científico hacen explícitos los supuestos del analista en cuanto a preferencias, tecnología, la información disponible para los agentes, las restricciones bajo las cuales actúan, las reglas de interacción entre agentes en los mercados y los entornos sociales y las fuentes de variabilidad entre los agentes. Estas características explícitas hacen que los modelos econométricos sean vehículos útiles para *i)* interpretar la evidencia empírica aplicando la teoría; *ii)* recopilar y sintetizar evidencias a través de estudios utilizando la teoría económica; *iii)* medir los distintos efectos de las políticas; y *iv)* pronosticar el bienestar y los efectos directos de políticas previamente aplicadas en nuevos entornos y los efectos de políticas nuevas.

Estas características están ausentes de la literatura moderna sobre efectos de tratamiento. Al mismo tiempo dicha literatura hace menos supuestos estadísticos que la literatura estándar de estimación estructural en la econometría. Estas son las características atractivas de dicho enfoque.

Sin embargo, la literatura econométrica ha hecho grandes avances en años recientes en términos de crear una versión robusta de su producto. Los progresos importantes resumidos en Powell (1994), Heckman y Vytlacil (2007b), y Matzkin (2007) han logrado relajar los supuestos paramétricos fuertes que caracterizaron la temprana literatura econométrica.

Al conciliar estas dos literaturas me retrotraigo a un trabajo importante, aunque olvidado, de Marschak (1953). Marschak hizo notar que para muchas cuestiones específicas de análisis de políticas no es necesario identificar modelos plenamente especificados que sean invariantes a tipos de modificaciones de política. Puede que todo lo que haga falta para analizar políticas sea combinaciones de subconjuntos de los parámetros estructurales que correspondan con parámetros que son necesarios para pronosticar modificaciones específicas de política, que suelen ser

mucho más fáciles de identificar (es decir, requieren un menor y más débil número de supuestos).

A este principio lo denominó la Máxima de Marschak en honor a su perspicacia. La literatura estadística moderna sobre efectos de tratamiento emplea la Máxima de Marschak cuando las políticas analizadas son los tratamientos disponibles bajo un régimen de política particular, y la meta del análisis de políticas está limitada a evaluar las políticas que ya están en ejecución (tarea 1 del cuadro 1) y no a pronosticar los efectos de nuevas políticas o de políticas existentes sobre entornos nuevos. Lo que suele faltar en la literatura sobre efectos de tratamiento es una clara discusión acerca del aspecto de la política que está siendo atendido por el efecto de tratamiento particular que está siendo identificado, y por qué interesa.

Los promedios de los parámetros de tratamiento de población suelen identificarse bajo condiciones más débiles de las que tradicionalmente supone el análisis econométrico estructural. Así, para identificar el efecto de tratamiento promedio dado  $X$  para  $s$  y  $s'$  sólo se requiere  $E(Y(s, \omega) | X = x) - E(Y(s', \omega) | X = x)$ . Bajo condiciones de invarianza sobre las ecuaciones de resultado, este parámetro da respuesta a la interrogante de política en cuanto a determinar el efecto promedio sobre los resultados de pasar un agente de  $s$  a  $s'$  cuando no hay interacción social ni efectos de contagio. El parámetro no está diseñado para evaluar una multitud de políticas adicionales. No hace falta conocer la forma funcional de las funciones generadoras  $g_s$ , ni  $X$  tiene que ser exógeno. No hace falta invocar condiciones fuertes de invarianza de las ecuaciones de elección. Sin embargo, si se pretende identificar  $E(Y(s, \omega) | X = x, D(s, \omega) = 1) - E(Y(s', \omega) | X = x, D(s, \omega) = 1)$ , se debe invocar la invarianza de las ecuaciones de elección reconociendo el condicionamiento sobre una variable de elección. No se requiere condicionar una elección al definir efectos de tratamiento promedio.

Los efectos de tratamiento son efectos causales para políticas específicas que pasan a los agentes de  $s \in S$  a  $s' \in S$ ,  $s \neq s'$ , manteniendo sin cambio todas las demás características del agente y del entorno. Estos efectos están diseñados para abordar el problema de políticas P1. Los efectos de tratamiento y los modelos causales se pueden generar a partir de modelos económicos explícitos, y se interpretan con mayor felicidad. Modelos económicos invariantes formulados explícitamente son útiles para abordar los problemas de políticas P2 y P3: la extrapolación y la predicción de efectos de nuevas políticas, respectivamente.

Si la meta del análisis es predecir resultados, y el entorno es

estable, entonces se pueden hacer predicciones precisas sin parámetros causales o estructurales. Recuérdese el análisis de Haavelmo en la sección IV-2 acerca de fijar vs. condicionar. Él analizó el modelo de regresión lineal  $Y = X\beta + U$  y definió el efecto causal de  $X$  sobre  $Y$  como el efecto constante de  $U$  de las variaciones en  $X$ . Si el objetivo del análisis es predecir el efecto de  $X$  sobre  $Y$ , y si el entorno es estable de tal forma que los datos históricos tienen la misma distribución que los datos de la muestra del pronóstico, las proyecciones de mínimos cuadrados resultan ser predictores óptimos según el criterio de error cuadrático medio (ver, por ejemplo, Goldberger, 1964). No hace falta extraer el efecto causal de  $X$  sobre  $Y$ ,  $\beta$ , del efecto de  $X$  sobre las no observables que opera por medio de  $E(U|X)$ .

La Máxima de Marschak exhorta a los analistas a formular claramente el problema que se aborda, y utilizar el mínimo número de ingredientes para resolverlo. La literatura de efectos de tratamiento atiende el problema de comparar tratamientos según un régimen particular de políticas dentro de un entorno específico. Los pioneros econométricos originales consideraban tratamientos según distintos regímenes de políticas y en diferentes entornos. A medida que los analistas formulan preguntas más difíciles, es necesario especificar más características de los modelos utilizados para atender dichas preguntas.

La Máxima de Marschak es una aplicación de la Navaja de Occam a la evaluación de políticas. Para ciertas clases de intervenciones de política, diseñadas para resolver el problema P1, el enfoque del efecto de tratamiento puede ser muy poderoso y más convincente que los modelos explícitamente formulados, ya que conlleva menos supuestos. Sin embargo, como ya se ha notado, se han logrado progresos importantes en relajar la estructura paramétrica postulada para los primeros modelos económicos. A medida que la literatura de efectos de tratamiento se va extendiendo hasta atender el conjunto más general de problemas de pronóstico en materia de políticas que contiene la literatura económica, incluyendo la evaluación de resultados subjetivos, y se van relajando los supuestos paramétricos del enfoque econométrico original, ambas literaturas convergerán hacia una síntesis creativa.

## REFERENCIAS

Abadie, A., J. D. Angrist y G. Imbens (2002), "Instrumental

- variables estimates of the effect of subsidized training on the quantiles of trainee earnings”, *Econometrica*, vol. 70, pp. 91–117.
- Abbring, J. H., y G. J. Van den Berg (2003), “The nonparametric identification of treatment effects in duration models”, *Econometrica*, vol. 71, pp. 1491–517.
- Abbring, J. H., y J. J. Heckman (2007), “Econometric evaluation of social programs, part III: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation”, en J. Heckman y E. Leamer (eds.), *Handbook of Econometrics*, vol. 6B, Elsevier, Amsterdam, pp. 5145–303.
- Abbring, J. H., y J. J. Heckman (2008), “Dynamic policy analysis”, en L. Matyas y P. Sevestre (eds.), *The Econometrics of Panel Data*, Springer, Dordrecht, 3rd edn., próxima publicación.
- Amemiya, T. (1985), *Advanced Econometrics*, Harvard University Press, Cambridge, Ma.
- Angrist, J. D., G. W. Imbens y D. Rubin (1996), “Identification of causal effects using instrumental variables”, *Journal of the American Statistical Association*, vol. 91, pp. 444–55.
- Angrist, J. D., y G. W. Imbens (1995), “Two-stage least squares estimation of average causal effects in models with variable treatment intensity”, *Journal of the American Statistical Association*, vol. 90, pp. 431–42.
- Berk, R., A. Li y L. J. Hickman (2005), “Statistical difficulties in determining the role of race in capital cases: A re-analysis of data from the state of Maryland”, *Journal of Quantitative Criminology*, vol. 21, pp. 365–90.
- Björklund, A., y R. Moffitt (1987), “The estimation of wage gains and welfare gains in self selection”, *Review of Economics and Statistics*, vol. 69, pp. 42–49.
- Brock, W. A., y S. N. Durlauf (2001), “Interactions-based models”, en J. J. Heckman y E. Leamer (eds.), *Handbook of Econometrics*, vol. 5, Elsevier, Amsterdam, pp. 3463–568.
- Carneiro, P., J. J. Heckman y E. J. Vytlačil (2006), “Estimating marginal and average returns to education”, *American Economic Review*, under revision.
- Carneiro, P., K. Hansen y J. J. Heckman (2001), “Removing the veil of ignorance in assessing the distributional impacts of social policies”, *Swedish Economic Policy Review*, vol. 8, pp. 273–301.
- Carneiro, P., K. Hansen y J. J. Heckman (2003), “Estimating distributions of treatment effects with an application to the

- returns to schooling and measurement of the effects of uncertainty on college choice”, *International Economic Review*, vol. 44, pp. 361–422.
- Cartwright, N. (2004), “Causation: One word many things”, *Philosophy of Science*, vol. 71, pp. 805–19.
- Chan, T. Y., y B. H. Hamilton (2006), “Learning, private information and the economic evaluation of randomized experiments”, *Journal of Political Economy*, vol. 114, pp. 997–1040.
- Cox, D. R. (1958), *Planning of Experiments*, Wiley, Nueva York.
- Cunha, F., J. J. Heckman y S. Navarro (2005), “Separating uncertainty from heterogeneity in life cycle earnings, The 2004 Hicks Lecture”, *Oxford Economic Papers*, vol. 57, pp. 191–261.
- Cunha, F., J. J. Heckman y S. Navarro (2006), “Counterfactual analysis of inequality and social mobility”, en S. L. Morgan, D. B. Grusky y G. S. Fields (eds.), *Mobility and Inequality: Frontiers of Research in Sociology and Economics*, cap. 4, Stanford University Press, Stanford, Ca., pp. 290–348.
- Cunha, F., y J. J. Heckman (2007), *The evolution of uncertainty in labor earnings in the US economy*, texto mimeografiado, University of Chicago.
- Dawid, A. (2000), “Causal inference without counterfactuals”, *Journal of the American Statistical Association*, vol. 95, pp. 407–424.
- Gill, R. D., y J. M. Robins (2001), “Causal inference for complex longitudinal data: The continuous case”, *Annals of Statistics*, vol. 29, pp. 1785–811.
- Goldberger, A. S. (1964), *Econometric Theory*, Wiley, Nueva York.
- Granger, C. W. J. (1969), “Investigating causal relations by econometric models and crossspectral methods”, *Econometrica*, vol. 37, pp. 424–38.
- Gronau, R. (1974), “Wage comparisons – a selectivity bias”, *Journal of Political Economy*, vol. 82, pp. 1119–43.
- Haavelmo, T. (1943), “The statistical implications of a system of simultaneous equations”, *Econometrica*, vol. 11, pp. 1–12.
- Haavelmo, T. (1944), “The probability approach in econometrics”, *Econometrica*, vol. 12, pp. iii–vi y 1–115.
- Heckman, J. J. (1974), “Shadow prices, market wages, and labor supply”, *Econometrica*, vol. 42, pp. 679–94.
- Heckman, J. J. (1976), “The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models”, *Annals of Economic and Social Measurement*, vol. 5, pp. 475–92.



- Heckman, J. J. (1978), "Dummy endogenous variables in a simultaneous equation system", *Econometrica*, vol. 46, pp. 931–959.
- Heckman, J. J. (1979), "Sample selection bias as a specification error", *Econometrica*, vol. 47, pp. 153–62.
- Heckman, J. J. (1990), "Varieties of selection bias", *American Economic Review*, vol. 80, pp. 313–18.
- Heckman, J. J. (1992), "Randomization and social policy evaluation", en C. Manski y I. Garfinkel (eds.), *Evaluating Welfare and Training Programs*, Harvard University Press, Cambridge, Ma., pp. 201–30.
- Heckman, J. J. (2001), "Micro data, heterogeneity, and the evaluation of public policy: Nobel Lecture", *Journal of Political Economy*, vol. 109, pp. 673–748.
- Heckman, J. J. (2005), "The scientific model of causality", *Sociological Methodology*, vol. 35, pp. 1–97.
- Heckman, J. J. y B. E. Honor'e (1990), "The empirical content of the Roy model", *Econometrica*, vol. 58, pp. 1121–49.
- Heckman, J. J., J. A. Smith y N. Clements (1997), "Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts", *Review of Economic Studies*, vol. 64, pp. 487–536.
- Heckman, J. J., L. J. Lochner y C. Taber (1998a), "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents", *Review of Economic Dynamics*, vol. 1, pp. 1–58.
- Heckman, J. J., L. J. Lochner y C. Taber (1998b), "General-equilibrium treatment effects: A study of tuition policy", *American Economic Review*, vol. 88, pp. 381–86.
- Heckman, J. J., L. J. Lochner y C. Taber (1998c), "Tax policy and human-capital formation", *American Economic Review*, vol. 88, pp. 293–97.
- Heckman, J. J., N. Hohmann, J. Smith y M. Khoo (2000), "Substitution and dropout bias in social experiments: A study of an influential social experiment", *Quarterly Journal of Economics*, vol. 115, pp. 651–94.
- Heckman, J. J., R. J. LaLonde y J. A. Smith (1999), "The economics and econometrics of active labor market programs", en O. Ashenfelter y D. Card (eds.), *Handbook of Labor Economics*, vol. 3A, cap. 31, North-Holland, Nueva York, pp. 1865–2097.
- Heckman, J. J., S. Urzua y E. J. Vytlačil (2006), "Understanding instrumental variables in models with essential heterogeneity", *Review of Economics and Statistics*, vol. 88, pp. 389–432.

- Heckman, J. J., y E. J. Vytlačil (1999), "Local instrumental variables and latent variable models for identifying and bounding treatment effects", *Proceedings of the National Academy of Sciences*, vol. 96, pp. 4730-34.
- Heckman, J. J., y E. J. Vytlačil (2005), "Structural equations, treatment effects and econometric policy evaluation", *Econometrica*, vol. 73, pp. 669-738.
- Heckman, J. J., y E. J. Vytlačil (2007a), "Econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation", en J. Heckman y E. Leamer (eds.), *Handbook of Econometrics*, vol. 6B, Elsevier, Amsterdam, pp. 4779-874.
- Heckman, J. J., y E. J. Vytlačil (2007b), "Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative economic estimators to evaluate social programs and to forecast their effects in new environments", en J. Heckman y E. Leamer (eds.), *Handbook of Econometrics*, vol. 6B, Elsevier, Amsterdam, pp. 4875-5144.
- Heckman, J. J., y J. A. Smith (1998), "Evaluating the welfare state", en S. Strom (ed.), *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, Cambridge University Press, Nueva York, pp. 241-318.
- Heckman, J. J., y S. Navarro (2007), "Dynamic discrete choice and dynamic treatment effects", *Journal of Econometrics*, vol. 136, pp. 341-96.
- Hensher, D., J. Louviere y J. Swait (1999), "Combining sources of preference data", *Journal of Econometrics*, vol. 89, pp. 197-221.
- Holland, P. W. (1986), "Statistics and causal inference", *Journal of the American Statistical Association*, vol. 81, pp. 945-60.
- Hurwicz, L. (1962), "On the structural form of interdependent systems", en E. Nagel, P. Suppes y A. Tarski, *Logic (eds.), Methodology and Philosophy of Science*, Stanford University Press, pp. 232-239.
- Imbens, G. W., y J. D. Angrist (1994), "Identification and estimation of local average treatment effects", *Econometrica*, vol. 62, pp. 467-75.
- Katz, D., A. Gutek, R. Kahn y E. Barton (1975), *Bureaucratic Encounters: A Pilot Study in the Evaluation of Government Services*, Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor.
- Lewis, H. G. (1974), "Comments on selectivity biases in wage comparisons", *Journal of Political Economy*, vol., 82, pp. 1145-1155.

- Marschak, J. (1953), "Economic measurements for policy and prediction", en W. Hood y T. Koopmans(eds.), *Studies in Econometric Method*, Wiley, Nueva York, pp. 1–26.
- Marshall, A. (1890), *Principles of Economics*, Macmillan and Company, Nueva York.
- Matzkin, R. L. (2007), "Nonparametric identification", en J. Heckman y E. Leamer (eds.), *Handbook of Econometrics*, vol. 6B, Elsevier, Amsterdam.
- McFadden, D. (1974), "Conditional logit analysis of qualitative choice behavior", en P. Zarembka (ed.), *Frontiers in Econometrics*, Academic Press, Nueva York.
- McFadden, D. (1981), "Econometric models of probabilistic choice", en C. Manski y D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, Ma.
- Neyman, J. (1923), "Statistical problems in agricultural experiments", *Journal of the Royal Statistical Society*, vol. II (Supplemento), pp. 107–80.
- Pearl, J. (2000), *Causality*, Cambridge, Cambridge University Press, England.
- Powell, J. L. (1994), "Estimation of semi-parametric models", en R. Engle y D. McFadden (eds.), *Handbook of Econometrics*, vol. 4, Elsevier, Amsterdam, pp. 2443–521.
- Quandt, R. E. (1958), "The estimation of the parameters of a linear regression system obeying two separate regimes", *Journal of the American Statistical Association*, vol. 53, pp. 873–80.
- Quandt, R. E. (1972), "A new approach to estimating switching regressions", *Journal of the American Statistical Association*, vol. 67, pp. 306–10.
- Roy, A. (1951), "Some thoughts on the distribution of earnings", *Oxford Economic Papers*, vol. 3, pp. 135–46.
- Rubin, D. B. (1976), "Inference and missing data", *Biometrika*, vol. 63, pp. 581–92.
- Rubin, D. B. (1978), "Bayesian inference for causal effects: The role of randomization", *Annals of Statistics*, vol. 6, nº 1, pp. 34–58.
- Rubin, D. B. (1986), "Statistics and causal inference: Comment: Which ifs have causal answers", *Journal of the American Statistical Association*, vol. 81, pp. 961–62.
- Shadish, W. R., y T. D. Cook (2007), *Experimental and Quasi-Experimental Designs for Field Research*, Lawrence Erlbaum and Associates, Hillsdale, NJ, próxima publicación.
- Sims, C. A. (1972), "Money, income, and causality", *American Economic Review*, vol. 62, pp. 540–52.

- Tamer, E. (2003), "Incomplete simultaneous discrete response model with multiple equilibria", *Review of Economic Studies*, vol. 70, pp. 147–65.
- Thurstone, L. L. (1927), "A law of comparative judgement", *Psychological Review*, vol. 34, pp. 273–86.
- Thurstone, L. L. (1959), *The Measurement of Values*, University of Chicago Press, Chicago.
- Tinbergen, J. (1930), „Bestimmung und deutung von angebotskurven“, *Zeitschrift für National ökonomie*, vol. 1, pp. 669–79.
- Tukey, J. W. (1986), "Comments on alternative methods for solving the problem of selection bias in evaluating the impact of treatments on outcomes", en H. Wainer (ed.), *Drawing Inferences from Self-Selected Samples*, Springer-Verlag, Nueva York, pp. 108–10; reimpresso en 2000, Lawrence Erlbaum Associates, Mahwah, NJ.
- Van der Laan, M. J., y J. M. Robins (2003), *Unified Methods for Censored Longitudinal Data and Causality*, Springer-Verlag, Nueva York.

*Daniel Barrález*  
*Carolina Pagliacci*

# Identificación de segmentos de precios en el mercado de fondos *overnight* usando modelos ocultos de Markov

## I. INTRODUCCIÓN

Se ha observado que las tasas de interés para las transacciones en el mercado de fondos *overnight* se pueden agrupar en torno a niveles de tasas de interés significativamente diferentes a lo largo de un típico día de negociación. Podemos pensar en este fenómeno como la existencia de distintos segmentos de precios que se alternan en el tiempo en tanto las condiciones del mercado varían. Para demostrar esta afirmación, vamos a explicar a continuación dos maneras mediante las cuales podría ocurrir la segmentación de precios y definiremos estadísticamente un segmento de precios.

El análisis de los mercados financieros radica en la premisa de que los precios de transacción difieren de los valores fundamentales de los activos debido a la existencia de fricciones,

*Traduce y publica el CEMLA, con la debida autorización, el artículo de D. Barrález (Profesor de la Universidad Central de Venezuela, Investigador invitado al Banco Central de Venezuela) y C. Pagliacci (Investigadora económica en el Banco Central de Venezuela), presentado con el título Identification of Price Segments in the Overnight Fund Market using Hidden Markov Models, en la XII Reunión de la Red de Investigadores de Bancos Centrales del Continente Americano, realizada en Madrid, del 5 al 7 de noviembre de 2007, con el auspicio del Banco de España.*

definidas como el grado de dificultad con el cual un activo es negociado (Stoll, 2000). Estas desviaciones de los precios de transacción están asociadas con los costos de transacción de diversa índole, tales como costos de selección adversa, costos de inventario o costos de procesamiento de órdenes. Gran parte de la literatura tiene como objetivo la medición y comprensión de estos costos de transacción, sin que se estudie explícitamente los valores fundamentales de los activos o precios eficientes. Esto es así porque los precios eficientes no son observables directamente y responden a cambios en las condiciones del mercado. En un intento para descomponer los precios de transacción, Hasbrouck (1993) asume que el precio eficiente, aunque no es observable, sigue una caminata aleatoria y por lo tanto se mueve más lentamente que el precio de transacción observado.

En el mercado de fondos *overnight*, la tasa de interés de transacción puede ser considerada también como la suma de dos componentes, el valor fundamental o eficiente de los fondos y los costos de transacción. En condiciones normales, el valor fundamental debería cambiar de vez en cuando para reflejar el efecto de los choques asociados a la llegada de nueva información al mercado. En cambio, los costos de transacción presionan las tasas de interés de transacción de modo que se aparten de su valor fundamental, de modo que, las transacciones iniciadas por los compradores de fondos terminan por encima del valor fundamental (operaciones en el precio de venta), mientras que las transacciones iniciadas por los oferentes de fondos terminan por debajo del valor fundamental (operaciones en el precio de compra). Según estas circunstancias, cuando un choque lo suficientemente grande y permanente golpea al mercado, los precios de transacción empiezan a agruparse alrededor de un nivel de tasa de interés significativamente diferente, mostrando una especie de segmentación de precios o de estacionariedad por tramos.

Como se señaló en un trabajo anterior (Pagliacci, 2006), el mercado de fondos *overnight* venezolano posee un sistema de negociación electrónico similar a un mercado de órdenes limitadas, pero tiene una característica peculiar que permite a cada uno de sus participantes asignar líneas de crédito discrecionales al resto de los participantes. Esto implica que el monto de negociaciones potenciales en el mercado presumiblemente se reduce, dado que el intercambio de fondos solo ocurre cuando la tasa de interés y las líneas de crédito asignadas por los oferentes de fondos convergen con las de los compradores.<sup>1</sup> Sin embargo,

<sup>1</sup> Podría darse el caso de que para una orden de compra entrante con la

una consecuencia más importante es que las líneas de crédito podrían promover indirectamente la discriminación de precios por parte de los oferentes de fondos, debido a la combinación de dos circunstancias: primero, la oferta de fondos usualmente se concentra en pocos bancos; segundo, el tamaño reducido del mercado y la interacción repetida entre los participantes permite a los oferentes identificar una demanda potencialmente elevada por parte de algunos compradores. En otras palabras, la discriminación de precios surge porque los choques de oferta y demanda se distribuyen de manera asimétrica en un mercado imperfectamente anónimo, y las restricciones del crédito proporcionan el mecanismo para implementarlo.<sup>2</sup> Si la discriminación de precios tiene lugar, los precios de transacción para los participantes de crédito restringido terminan siendo más altos que los precios para los no restringidos, mostrando también una forma de segmentación de precios. En este caso, las tasas de interés de transacción se agrupan simultáneamente alrededor de dos valores no observables diferentes: el valor fundamental del activo por un lado, y el valor fundamental del activo más una prima de crédito por el otro lado. Esto se traduce en dos segmentos diferentes de tasas de interés que se alternan frecuentemente en pequeños intervalos de tiempo.

A partir de la explicación anterior, resulta claro que la observación de segmentos de precios diferentes en el mercado *overnight* venezolano podría racionalizarse ya sea como: *i*) un cambio en el valor fundamental de los fondos; o *ii*) la existencia de discriminación de precios a través de las líneas de crédito. No obstante, independientemente de la razón que causa la segmentación de tasas de interés, en primer lugar es necesario reconocer cuáles son los segmentos de precios diferentes. Esta tarea no puede realizarse mediante observación y requiere el uso de una adecuada herramienta estadística.

---

tasa de interés idéntica a una orden de venta en espera en el sistema, no ocurra coincidencia alguna, debido a las condiciones del crédito impuestas. Esto explica que en un determinado punto en el tiempo, la mejor cotización de venta puede ser menor en relación con la mejor cotización de compra. Otra manera de pensar sobre esta anomalía es imaginando al sistema (*order book*) con el cronograma de ventas intersectando al de compras.

<sup>2</sup> Algunos analistas del mercado argumentan que la discriminación de precios ocurre porque algunos bancos son percibidos como más riesgosos que otros. No obstante, el riesgo de incumplimiento soportado por los oferentes puede ser minimizado a través de la solicitud de una garantía equivalente al monto de los fondos negociados. Esta operación es perfectamente posible dentro del sistema electrónico venezolano, lo cual razonablemente ignora este argumento como una causa posible para la discriminación de precios.

Estadísticamente hablando, la observación de diferentes segmentos de precios es equivalente a identificar diferentes procesos estocásticos que generan precios de transacción. De modo más específico, una secuencia de tasas de interés puede considerarse como perteneciente a un segmento único, si todas las realizaciones en él provienen del mismo proceso de generación de datos. Un marco natural para identificar procesos estocásticos mixtos lo proporcionan los Modelos Ocultos de Markov (*Hidden Markov Models*, HMM), que asumen que cada proceso estocástico es producido por un estado no observable (oculto). En este estudio, usamos HMM para clasificar las transacciones en cuatro segmentos diferentes de tasas de interés, asociando los segmentos a los estados ocultos. Cada uno de estos estados ocultos puede ser racionalizado como condiciones de mercado distintivas que afectan el valor del activo subyacente alrededor del cual las tasas de interés de transacción se agrupan.

Entre las aplicaciones diferentes de los HMM, la más popular se relaciona con el reconocimiento de voz o del habla, que se basa en el trabajo seminal de Rabiner (1989 y 1993). También en 1989, el trabajo de investigación de Hamilton abrió un campo rico de aplicaciones para la macroeconomía. Aunque su modelación asume la existencia de información oculta, tiene una estructura diferente que la de los clásicos HMM. En finanzas, los HMM han sido utilizados principalmente para analizar los rendimientos diarios de índice SP&500 de Estados Unidos y otros precios de acciones. En particular, Rydén, Teräsvirta y Asbrink (1998) modelaron la tasa de fondos *overnight* y mostraron que las predicciones tienen un buen poder predictivo y satisfacen algunas características de los datos reales. En Weigend y Shi (2000), la predicción del SP&500 se mejora permitiendo que los procesos estocásticos involucren también una función de una variable exógena. En el 2004, un libro de Bhar y Hamori muestra otras aplicaciones financieras.

En este trabajo, aplicamos HMM, no para predecir la tasa de fondos *overnight*, sino para comprender y caracterizar el mercado, identificando los segmentos de precios diferentes. En particular, estamos interesados en la segmentación de precios porque proporciona información valiosa a los analistas acerca de cómo se desenvuelve el mercado a diario en términos de sus fundamentos, y cuando choques significativos golpean al mercado. Adicionalmente, como señalaron Bahr y Hamori (2004), la utilización de HMM permite detectar cambios bruscos en la volatilidad del mercado que normalmente no son detectados



por otros modelos estadísticos como ARCH y GARCH (usados en Weigend y Shi, 2000).

Para completar la caracterización de la segmentación de precios, mostramos que la identificación de segmentos de precios contabiliza una gran variación de la tasa de interés.

Describimos el patrón de alternancia (o intercambio) de los segmentos de precios a través de la construcción de un índice. Los valores del índice ayudan a inferir el tipo de segmentación que ocurre en el mercado en un determinado período de tiempo.

La estructura del estudio es la siguiente. Primero, explicamos la relación entre segmentos de precios y HMM. Entonces, presentamos el método de estimación y los resultados. Tercero, identificamos los segmentos de precios, y al final, los usamos para describir su poder explicativo y caracterizar los tipos diferentes de segmentación.

## II. MODELOS OCULTOS DE MARKOV Y SEGMENTOS DE PRECIOS

Los HMM se basan en las Cadenas de Markov clásicas, que definimos a continuación. Considere un proceso estocástico  $(Q_t)$  descrito en el período  $t$  por un estado obtenido a partir de un conjunto de  $N$  estados,  $S_1, S_2, \dots, S_N$ . El proceso  $(Q_t)$  es una cadena de Markov si, para realizar una predicción de lo que va a suceder en el período  $t$ , solamente la información en el período  $t-1$  es necesaria, mientras que el resto de la información es desechada. Esto es:

$$(1) \quad P(q_t = s_j / q_{t-1} = s_i, q_{t-2} = s_k, \dots, q_1 = s_m) = P(q_t = s_j / q_{t-1} = s_i)$$

Esta cadena de Markov es homogénea, debido a que el lado derecho de la expresión (denominada, la probabilidad de transición del estado  $q_{t-1}$  al estado  $q_t$ ) es independiente del tiempo. Para describir la evolución en el tiempo del proceso, una matriz probabilidades de transición de estados  $A = \{a_{ij}\}$  y la distribución inicial  $\pi$  son definidas como:

$$(2) \quad a_{ij} = P(q_t = s_j | q_{t-1} = s_i) \quad 1 \leq i, j \leq N$$

$$(3) \quad \pi_i = P(q_1 = s_i) \quad 1 \leq i \leq N$$

La anterior cadena de Markov se denomina observable porque los resultados del proceso son los estados en sí mismos.

Suponga ahora que las observaciones de series de tiempo no estacionarias, por decir  $Y_t$ , tienen funciones de densidad que dependen del estado aleatorio de una cadena de Markov, pero dicho estado no es observable directamente. Este doble proceso estocástico es conocido como HMM, el cual se caracteriza por:

- $N$  estados ocultos  $\{S_1, S_2, \dots, S_N\}$  y un estado de la distribución de probabilidades de transición;

$$(4) \quad \begin{aligned} a_{ij} &= P(q_t = s_j | q_{t-1} = s_i) \\ \sum_{j=1}^N a_{ij} &= 1 \end{aligned} \quad 1 \leq i, j \leq N$$

- la distribución inicial de los estados ocultos;

$$(5) \quad \pi_i = P(q_1 = s_i) \quad 1 \leq i \leq N$$

- y, el conjunto de densidades de observación  $B = \{b_j\}$ , denominadas densidades de emisión, donde  $b_j$  es la densidad de probabilidad cuando el proceso está en el estado  $j$ . En particular, para la observación  $y_t$ .

$$(6) \quad b_j(y_t) = P(y_t | q_t = s_j)$$

Un HMM proporciona el mecanismo para un sistema aleatorio que funciona de la siguiente manera. En el período  $t=1$ , el estado inicial  $q_1 = S_i$  se escoge aleatoriamente, de acuerdo con la distribución de probabilidad inicial  $\pi$ . Dado este estado  $S_i$ , una señal  $y_1$  será observada de acuerdo con la densidad de emisión  $b_i$ . En el período  $t=2$ , el proceso cambia al estado  $S_j$  con una probabilidad  $a_{ij}$ , y una realización  $y_2$  se obtiene de acuerdo con  $b_j$ , y así sucesivamente. Observe que una especificación completa de un HMM está dada por la especificación de las medidas de probabilidad  $A$ ,  $B$  y  $\pi$ , donde  $\lambda = \{A, B, \pi\}$  denota el conjunto completo de parámetros.

Con la finalidad de identificar segmentos diferentes de precios en el mercado *overnight*, asumimos que las densidades de emisión (procesos estocásticos) que generan tasas de interés de transacción dependen de los estados ocultos de un HMM. En particular, una realización de la tasa de interés dada ( $y_t$ ) pertenece al segmento  $i$ , si el estado oculto  $s_i$  ocurre y la tasa de interés observada fue generada de acuerdo con  $b_i$ . En este caso, el estado oculto  $s_i$  puede ser interpretado como las condiciones del mercado particulares que hacen que las tasas de interés de

transacción se agrupan alrededor de la media del proceso estocástico descrito por  $b_i$  (el valor subyacente del activo).

La identificación del segmento de precios se cumple en dos etapas. Primero, se estima un HMM. En esta etapa, también denominada de entrenamiento, encontramos los parámetros del modelo  $(A, B, \pi)$  que mejor explican el conjunto de datos. En la segunda etapa, llamada de clasificación o reconocimiento, la secuencia de las tasas de interés de transacción está etiquetada de acuerdo con el estado oculto más probable que la generó. Esta tarea se realiza usando el algoritmo de Viterbi. Estas dos etapas se abordarán en detalle en las próximas secciones.

### III. ESTIMACIÓN DEL MODELO

El conjunto de datos consiste en las tasas de interés de transacciones individuales para 173 días de negociación de abril a diciembre del 2005, sumando un total de 26,398 transacciones.

El conjunto de datos se divide en dos grupos o clases de transacciones: negociaciones iniciadas por vendedores y negociaciones iniciadas por compradores de fondos. Estimamos dos HMM separados, cada uno para cada grupo de datos.<sup>3</sup> Estadísticamente, este enfoque se justifica por la diferencia en el promedio de la muestra y la varianza de las transacciones por clase.<sup>4</sup> Económicamente, permitimos dos estimaciones independientes para capturar discrepancias potencialmente significativas en el comportamiento entre oferentes y demandantes de fondos. En circunstancias normales, dichas diferencias podrían estar asociadas con grandes márgenes y las características particulares de los cronogramas de venta y compra. En situaciones de discriminación de precios, las diferencias pueden surgir del hecho de que restricciones obligatorias de crédito son asignadas discrecionalmente por los oferentes de fondos.

Las tasas de interés de transacciones fluctúan discretamente desde 0.2% a 15% de la siguiente manera:<sup>5</sup>

<sup>3</sup> Las observaciones para negociaciones iniciadas por vendedores son 11,636, mientras que las observaciones para negociaciones iniciadas por compradores son 14,762.

<sup>4</sup> Para transacciones iniciadas por vendedores, el promedio muestral y la desviación estándar son 2.28% y 2.85% respectivamente, mientras que para las transacciones iniciadas por compradores, el promedio muestral y la desviación estándar son 2.81% y 3.24%.

<sup>5</sup> Si bien el sistema de negociación electrónico permite ofertas de tasas de interés continuas, observamos que las tasas de interés de transacción terminan

$$(7) \quad y_t = \begin{cases} 0.2, \dots, 0.8 & \text{tamaño de la variación mínima} = 0.05 \\ 1, \dots, 2 & \text{tamaño de la variación mínima} = 0.25 \\ 2.5, \dots, 6 & \text{tamaño de la variación mínima} = 0.50 \\ 7, \dots, 15 & \text{tamaño de la variación mínima} = 1.00 \end{cases}$$

Estos 35 valores diferentes para las tasas de interés forman el soporte de las probabilidades de emisión a ser estimadas para cada uno de los estados ocultos.<sup>6</sup> Dado que cada estado oculto representa condiciones particulares de mercado que afectan las probabilidades de emisión, seleccionamos cuatro estados ocultos ( $N=4$ ) para representar situaciones en que los niveles de tasas de interés serían extremadamente bajos, medio bajos, medio altos y extremadamente altos.

Para estimar los parámetros  $\lambda$  para cada modelo, usamos el algoritmo Baum-Welch. Este algoritmo maximiza iterativamente  $P(Y/\lambda)$ , es decir, la probabilidad de la secuencia de observación dada por los parámetros del modelo. Esta maximización es alcanzada por los parámetros reestimados  $\hat{\lambda}$  tal que  $P(Y|\hat{\lambda}) \geq P(Y|\lambda)$ , donde  $\lambda$  son los valores de los parámetros estimados en la iteración previa. Después de un cierto número de iteraciones, este algoritmo alcanza un máximo local. Operativamente, la probabilidad  $P(Y/\lambda)$  es calculada mediante la desintegración de la secuencia de observaciones de las tasas de interés entera ( $Y$ ) en subsecuencias más pequeñas  $Y^d$ , y se calcula  $P(Y/\lambda) = \prod_d P(Y^d/\lambda)$ .<sup>7</sup> El algoritmo procede a lo largo de las siguientes etapas:

1. *Inicialización del modelo.* Los parámetros iniciales (previos)  $\lambda_0=(A_0, B_0, \pi_0)$  son proporcionados al modelo para empezar la maximización de  $P(Y/\lambda)$ . Imponemos que el HMM estimado es ergódico, lo que significa que todas las transiciones de un estado a otro son permitidas, y por lo tanto, todos los elementos de  $A_0$  son no nulos. Las densidades de emisión iniciales ( $B_0$ ) para cada estado oculto son construidas arbitrariamente dividiendo su soporte en cuatro secciones. Varios parámetros iniciales son

---

mostrando incrementos discretos (variación mínima del precio) cercanos a los descritos. Cualquier tasa de interés diferente a estos valores, es simplemente redondeada al entero más cercano.

<sup>6</sup> Para propósitos de estimación, los valores de soporte también son estandarizados a enteros que van de 1 a 35.

<sup>7</sup> Escogemos una longitud de 14 observaciones para formar dichas subsecuencias.

proporcionados y se selecciona el óptimo.<sup>8</sup> La distribución inicial de los estados ocultos se asume uniforme. Una iteración única es completada una vez que la secuencia entera de observaciones de tasas de interés ha sido procesada.

2. *Etapas de reestimación.* El algoritmo reestima  $\hat{\lambda} = (\hat{A}, \hat{B}, \hat{\pi})$  como sigue:

- Probabilidades iniciales, es decir, la frecuencia esperada en el estado  $s_i$  en el período  $t=1$ , se calculan como:

$$(8) \quad \hat{\pi}_i = \gamma_1(i) \quad 1 \leq i \leq 4$$

$$(9) \quad \gamma_t(i) = P(q_t = s_i | Y, \lambda),$$

donde  $\gamma_t(i)$  es la probabilidad *a posteriori* de estar en el estado  $s_i$  en el período  $t$ .

- Probabilidades de transición, es decir, el número esperado de transiciones desde el estado  $s_i$  al estado  $s_j$ , se establecen como:

$$(10) \quad \hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \varepsilon_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \dots\dots\dots 1 \leq i, j \leq 4$$

$$(11) \quad \varepsilon_t(i, j) = P(q_t = s_i, q_{t+1} = s_j | Y, \lambda),$$

donde  $\varepsilon_t(i, j)$  es la probabilidad *a posteriori* de estar en el estado  $s_i$  en el período  $t$  y en el estado  $s_j$  en el período  $t+1$ .

- Probabilidades de emisión de estado están dadas por:

$$(12) \quad \hat{b}_j(a_k) = \frac{\text{Núm. esperado de veces en estado } j \text{ y observando } a_k}{\text{Núm. esperado de veces en estado } j} \\ = \frac{\sum_{t=1 \text{ and } y_t = a_k}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)},$$

donde  $a_k$  es un valor de soporte de la probabilidad de emisión.

3. *Regla de detención.* El procedimiento de la estimación se detiene

<sup>8</sup> Más detalles sobre este procedimiento se darán en la sección de resultados.

cuando la probabilidad de la secuencia de la observación se estabiliza, esto es cuando:

$$(13) \quad \left| \frac{(Y | \hat{\lambda}) - P(Y | \lambda)}{P(Y | \lambda)} \right| < c,$$

donde  $c$  es usualmente establecido en  $10^{-3}$ . Alternativamente, puede fijarse un número máximo de iteraciones.

#### IV. RESULTADOS DE LA ESTIMACIÓN

Los resultados de la estimación usando el algoritmo Baum-Welch dependen en gran medida de los parámetros iniciales asignados a las densidades de emisión ( $B_0$ ). Para evitar una selección *ad hoc* de estos parámetros iniciales, los escogemos óptimamente usando el algoritmo Baum-Welch. Esto es, estimamos los parámetros del modelo ( $\lambda$ ) para varios parámetros iniciales potenciales y mantenemos los parámetros iniciales y los parámetros estimados que maximizan la probabilidad de las observaciones de las tasas de interés. Los parámetros iniciales buscados son definidos arbitrariamente dividiendo el soporte de  $Y$  en cuatro secciones, cada uno de ellos lleva la probabilidad más alta para el mismo estado oculto, por ejemplo, para el estado  $S_3$ , la sección tercera del soporte siempre tiene la mayor probabilidad asociada.

Los resultados de la estimación se presentan por clase o grupo de transacciones. Ver cuadros 1 y 3 para las negociaciones iniciales de vendedores y las gráficas II y IV para las negociaciones iniciadas por los compradores.

Los parámetros iniciales para la densidad de emisión que generaron la mayor probabilidad de observaciones son los mismos en ambos grupos. Las probabilidades de emisión estimadas difieren en su media y varianza. Para las negociaciones iniciadas por compradores, en todos los estados ocultos se espera el primero, la media de las distribuciones de probabilidad es más grande, pero paradójicamente, la desviación estándar es más pequeña. Esto implica que las realizaciones de las tasas de interés de transacción para las negociaciones iniciadas por compradores tenderán a agruparse de modo más compacto alrededor de niveles más altos.

Es interesante notar que existen secciones del soporte de la tasa de interés donde las probabilidades de emisión se superponen con probabilidades diferentes. Esto es particularmente cierto

CUADRO 1. PROBABILIDADES DE EMISIÓN PARA NEGOCIACIONES INICIADAS POR VENDEDORES

Soporte	Probabilidades iniciales				Probabilidades estimadas			
	S1	S2	S3	S4	S1	S2	S3	S4
0.20 - 0.65	0.60	0.30	0.08	0.08	0.91	0.08	0.02	0.00
0.70 - 3.00	0.30	0.60	0.08	0.08	0.09	0.92	0.55	0.02
3.50 - 6.00	0.04	0.04	0.64	0.21	0.00	0.00	0.42	0.28
7.00 - 15.00	0.06	0.06	0.21	0.64	0.00	0.00	0.01	0.70
	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
			Promedio		0.50	1.17	3.33	8.74
			Desviación estándar		0.31	0.50	1.40	3.05

CUADRO 2. PROBABILIDADES DE EMISIÓN PARA NEGOCIACIONES INICIADAS POR COMPRADORES

Soporte	Probabilidades iniciales				Probabilidades estimadas			
	S1	S2	S3	S4	S1	S2	S3	S4
0.20 - 0.65	0.60	0.30	0.08	0.08	0.88	0.06	0.01	0.00
0.70 - 3.00	0.30	0.60	0.08	0.08	0.11	0.93	0.50	0.02
3.50 - 6.00	0.04	0.04	0.64	0.21	0.00	0.00	0.49	0.23
7.00 - 15.00	0.06	0.06	0.21	0.64	0.00	0.00	0.01	0.75
	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
			Promedio		0.49	1.24	3.47	8.95
			Desviación estándar		0.23	0.49	1.26	2.95

para los estados  $S_3$  y  $S_2$ , los cuales intersecan en la segunda sección del soporte, y para los estados  $S_3$  y  $S_4$ , que intersecan en la tercera sección del soporte. Dichas intersecciones pueden posibilitar que *ex post*, ciertas tasas de interés puedan ser clasificadas como pertenecientes a cualquiera de estos estados.

**CUADRO 3.** MATRIZ DE TRANSICIÓN PARA LA NEGOCIACIONES INICIADAS POR VENDEDORES

	$S1$	$S2$	$S3$	$S4$
$S1$	0.96	0.04	0.00	0.00
$S2$	0.04	0.93	0.02	0.00
$S3$	0.00	0.03	0.96	0.01
$S4$	0.00	0.00	0.02	0.98

La estimación de probabilidades de transición muestra que, para ambos modelos, las probabilidades de permanecer en un estado después de visitarlo son muy grandes, y cercanas a 0.93. Una vez en los estados 2 y 3, existen probabilidades positivas de visitar los dos estados vecinos, pero una vez en los estados 1 y 4, solo un estado vecino puede ser alcanzado. Esto implica que la mayoría de transacciones de las negociaciones diarias serán asociadas con dos segmentos de precios, solo algunas veces con tres, pero nunca con cuatro segmentos.

**CUADRO 4.** MATRIZ DE TRANSICIÓN PARA NEGOCIACIONES INICIADAS POR COMPRADORES

	$S1$	$S2$	$S3$	$S4$
$S1$	0.97	0.03	0.00	0.00
$S2$	0.02	0.96	0.02	0.00
$S3$	0.00	0.02	0.96	0.01
$S4$	0.00	0.00	0.02	0.98

En términos de las diferencias entre los dos modelos, no existen muchas. Podemos solo afirmar que, una vez en el estado 2 y 3, existen ligeramente más oportunidades de visitar los estados más bajos para las negociaciones iniciadas de vendedores que para las negociaciones iniciadas de los compradores.

A partir de las probabilidades iniciales estimadas,  $\hat{\pi}^{\text{vendedores}} = \{0.33, 0.34, 0.21, 0.12\}$  y  $\hat{\pi}^{\text{compradores}} = \{0.27, 0.33, 0.23, 0.16\}$ , podemos deducir que las oportunidades de iniciar el día de negociaciones en los estados 3 y 4 son mayores para las negociaciones iniciadas por compradores que por vendedores.



## V. CLASIFICACIÓN DE LAS TASAS DE INTERÉS DE TRANSACCIÓN

Después de que la estimación de los parámetros se completa para ambos modelos, usamos el algoritmo de Viterbi para determinar a cuál segmento pertenece cada tasa de interés de transacción. Esto es equivalente a establecer, cuál estado oculto más probablemente generó una determinada observación de tasa de interés.

El algoritmo de Viterbi se basa en una rutina de programación dinámica (Rabiner, 1989) que encuentra la mejor secuencia de estados ocultos (en el sentido de probabilidad máxima) para una secuencia de observaciones determinada. De manera más específica, recursivamente calcula  $q_i^*$  de modo que:

$$(14) \quad Q^* = \operatorname{argmax} P(Q, Y / \lambda),$$

donde  $P(Q, Y / \lambda)$  es la probabilidad conjunta de observación de la secuencia de estados ocultos y la secuencia de observación.

Este algoritmo se implementa a través de los siguientes pasos:

1. *Inicio*. Las condiciones iniciales son definidas dados  $\hat{\lambda} = (\hat{A}, \hat{B}, \hat{\pi})$ .

$$(15) \quad \delta_1(i) = \pi_1 b_i(y_1)$$

$$(16) \quad \psi_1(i) = 0 \dots\dots\dots 1 \leq i \leq 4,$$

donde  $\delta_1(i)$  representa la probabilidad de estar en el estado  $s_i$  y observando  $y_1$ ; y  $\psi_1(i)$  es, en  $t=1$ , el estado más probable que precede el estado  $s_i$ .

2. *Recursión*. Dada una secuencia de observación de longitud  $T$ , las siguientes variables se calculan y guardan para cada  $t$ :

$$(17) \quad \delta_t(j) = \max\{\delta_{t-1}(i) a_{ij} b_j(y_t)\} \quad 2 \leq t \leq T$$

$$(18) \quad \psi_t(j) = \operatorname{argmax}_i [\delta_{t-1}(i) a_{ij}] \quad 1 \leq j \leq 4,$$

donde  $\delta_t(j)$  representa la probabilidad máxima (acumulada) de estar en el estado  $s_j$  y observando  $y_t$ ; y  $\psi_t(j)$  es, en el período  $t$ , el estado más probable  $s_i$  que precede al estado  $s_j$ .

3. *Finalización*. La secuencia óptima de estados ocultos es calculada, usando un procedimiento de "retroceso".

$$(19) \quad P^* = \max\{\delta_T(i)\}$$

$$(20) \quad q_t^* = \operatorname{argmax}[\delta_T(i)] \quad 1 \leq i \leq 4$$

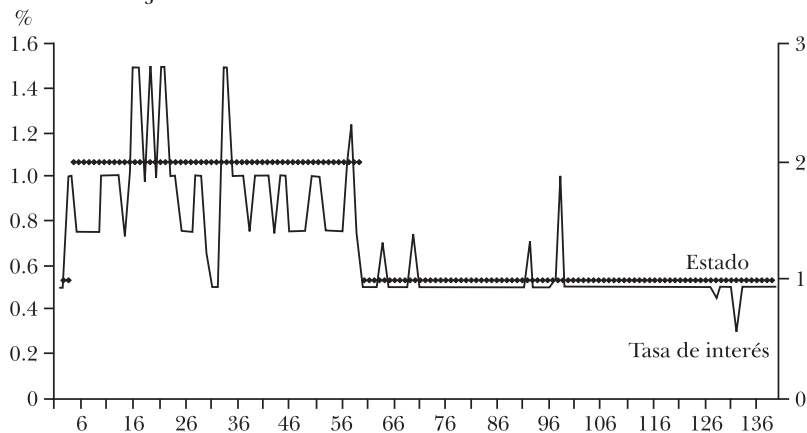
$$(21) \quad q_{t-1}^* = \psi_t(q_t^*) \quad q_{t-1}^* = \psi_t(q_{t-1}^*) \quad 2 \leq t \leq T,$$

donde  $P^*$  es la probabilidad global (máxima) de la secuencia de observación total.

En este estudio, la clasificación de las tasas de interés de transacción según el estado oculto asociado se efectúa de manera separada para cada uno de los dos modelos (negociaciones iniciadas de vendedores y compradores). En cada modelo, el algoritmo de Viterbi se aplica a las secuencias diarias de tasas de interés de transacciones, para evitar secuencias de observaciones excesivamente largas.<sup>9</sup> Entonces, las transacciones clasificadas se integran en series únicas según el momento de su llegada al sistema. De esta manera, podemos observar continuamente los segmentos de tasas de interés como en tiempo real. A continuación, proporcionamos dos ejemplos de identificación de segmentos para las transacciones de tasas de interés para dos días de negociación diferentes (gráficas I y II).

En la gráfica I, observamos que las tasas de interés pertenecen al segmento 1 (cuando las transacciones son valoradas en

**GRÁFICA I.** EJEMPLO DE IDENTIFICACIÓN DE SEGMENTOS

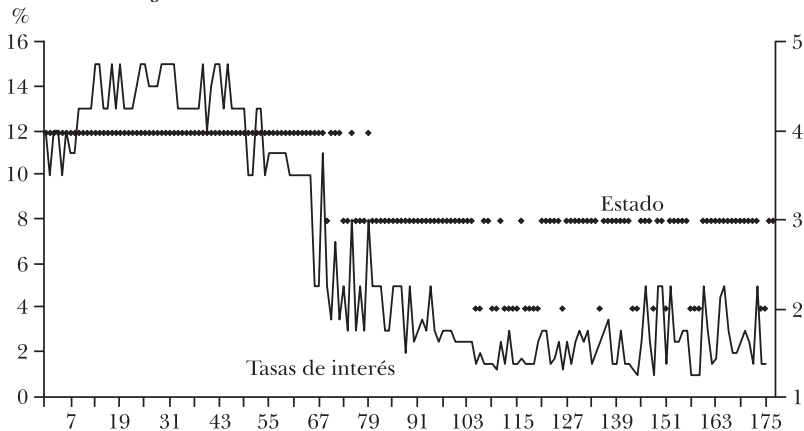


<sup>9</sup> Recuerde que en la medida en que las secuencias de observación se hacen demasiado largas, la probabilidad de quedarse en un único estado tiende rápidamente a cero, aumentando abruptamente la tasa de cambio entre los estados ocultos (Rabiner, 1989). Este es un problema a considerar, especialmente si se usa el algoritmo de Viterbi.

alrededor del 0.51%) y al segmento 2 (cuando las transacciones son valoradas en alrededor del 0.95%). Además, la manera en que se identifican los segmentos (en dos corridas separadas), permite inferir que una reducción en el valor fundamental de los fondos ocurre en la mitad del día de negociación. Esto puede asociarse con un cambio en las condiciones del mercado y con un ajuste en el comportamiento de los participantes. Es interesante notar que si bien la clasificación de las transacciones se efectúa en forma separada, las transacciones iniciadas tanto por los compradores como por los vendedores responden de manera similar a las condiciones del mercado.

En la gráfica II, si bien una reducción en el nivel de la tasa de interés puede observarse para la segunda parte del día, las transacciones se clasifican en dos segmentos de precios diferentes (3 y 2), los cuales se alternan frecuentemente en corridas pequeñas. Esto puede interpretarse como la existencia simultánea de dos valores subyacentes de fondos alrededor de los cuales se agrupan transacciones similares. Como se explicó anteriormente, este comportamiento de segmentos puede ser el resultado de prácticas discriminatorias de precios en asociación con restricciones crediticias.

**GRÁFICA II.** EJEMPLO DE IDENTIFICACIÓN DE SEGMENTOS



En este punto, efectuamos dos preguntas adicionales. Primero, ¿cuán informativa puede ser la identificación de segmentos de precios? Segundo, ¿puede el cambio de la caracterización de segmentos de precios decirnos algo acerca de las condiciones del mercado? Estas preguntas se responden a continuación.

## VI. PODER EXPLICATIVO DE LOS SEGMENTOS DE PRECIOS

La estimación de una densidad de emisión para cada estado oculto implica que el promedio y la varianza de dichas distribuciones son diferentes. Estimamos una regresión lineal simple para determinar cuál porción de la variación de las tasas de interés de transacción se explica por esta información.

Para contabilizar las diferencias entre estados, construimos variables *dummies* que identifican cada uno de los estados ocultos, y se usan como regresores. También incluimos dos términos de interacción contruidos con una variable flujo de negociación, la cual toma el valor de 1 cuando la transacción es iniciada por un comprador, y 0 de otra manera. Estos términos de interacción tratan de contabilizar la diferencia promedio en el margen negociado, cuando ambos, vendedores y compradores, inician las transacciones dentro del mismo segmento de precios.<sup>10</sup>

Una estimación simple MCO muestra que el  $R^2$  ajustado es igual a 0.81, demostrando que gran parte de la variación de las tasas de interés se explica de forma simple mediante la identificación del segmento al cual pertenecen los precios. Sin embargo, los resultados de la regresión indican que los errores y los errores al cuadrado tienen correlaciones significativas con las transacciones vecinas. Para controlar los problemas de correlación y permitir una interpretación correcta de los coeficientes, estimamos un modelo con componentes de errores autoregresivos en el promedio, una estructura GARCH(1,1) en la varianza y variables *dummies* de estado como variables explicativas exógenas para la varianza. Los principales resultados se muestran en el cuadro 5 y el resultado de la regresión se encuentra en el anexo 1.

**CUADRO 5. DISTRIBUCIÓN EMPÍRICA DE LOS SEGMENTOS DE PRECIOS**

<i>Estado</i>	<i>Promedio (pp)</i>	<i>Desviación Estándar</i>	<i>Margen Negociado (pp)</i>
1	0.673691	0.015133	—
2	1.173847	0.901713	0.078745
3	2.906688	1.164009	0.282493
4	6.074066	3.023467	—

NOTA:  $R^2$  ajustado es igual a 0.8684.

<sup>10</sup> El margen negociado se define como la diferencia promedio de tasas de interés entre negociaciones a la venta (iniciadas por el comprador) y las negociaciones a la compra (iniciadas por el vendedor).

Según estos resultados, el promedio empírico y la varianza de los segmentos de precios difieren en algún grado de los teóricos. No obstante, las principales diferencias entre estados se preservan. Los resultados también muestran que existe un mejoramiento en el  $R^2$  ajustado cercano al 5%, probablemente debido a la incorporación de componentes autorregresivos en el promedio y la varianza. El significado de estos componentes autorregresivos evidencia la identificación de segmentos de precios que no pueden ser contabilizados por todas las dinámicas de tiempo de las tasas de interés.

## VII. CARACTERIZACIÓN DE LOS SEGMENTOS DE TASAS DE INTERÉS

A partir de los ejemplos en la sección V, resulta claro que la alternancia o el comportamiento de cambio de los segmentos de precios conduce a información relevante acerca de las condiciones según las cuales ocurren las transacciones.

Definamos una corrida como una secuencia continua de transacciones que pertenecen a un segmento de precios idéntico. Proponemos un índice para capturar las diferencias en los patrones de alternancia de los segmentos, es decir, en el cambio de corridas de los segmentos. Este índice de cambio ( $SI$ ) se define para cualquier intervalo de tiempo  $t$  a ser estudiado, es decir, sobre las transacciones ocurridas en dicho intervalo:

$$(22) \quad SI_t = 1 - \frac{1}{2} \left[ \frac{l_t^m}{T} + \left( \frac{l_t^a}{T} \right)^{\frac{h-1}{2}} \right],$$

$$(23) \quad l_t^a = \frac{\sum_h \left( \sum_{y \in i} C(y^i) / ni \right)}{h},$$

donde  $l^m$  es la longitud máxima de corrida;  $l^a$  es la longitud promedio de corrida;  $T$  es el número de transacciones durante el intervalo de tiempo  $t$ ;  $h$  es el número de segmentos observado;  $C(y^i)$  es una función indicador que toma el valor de 1 cuando la tasa de interés de transacción pertenece al segmento  $s_i$ ; y  $ni$  es el número de corridas del tipo  $s_i$ .

El índice de cambio se calcula para intervalos diarios y su objetivo es discriminar entre tipos diferentes de segmentación de

precios. Este índice toma el valor de cero cuando las tasas de interés observadas se clasifican como pertenecientes a un segmento único. Alternativamente, toma un valor cercano a uno, cuando el cambio ocurre muy a menudo, en el límite, cuando cada tasa de interés es seguida y precedida por tasas de interés que pertenecen a segmentos de precios diferentes. En los ejemplos elegidos (gráficas I y II), el índice toma el valor de 0.42 y 0.77, respectivamente.

La distribución de probabilidad empírica del índice se muestra en el cuadro 6. Observando los resultados del índice de cambio y los patrones correspondientes de los segmentos, escogemos un *valor umbral* igual a 0.5 para discriminar dichos patrones. Valores del índice por encima de 0.5 muestran una segmentación muy cercana a la ocurrencia de prácticas de discriminación de precios, adicionalmente a cambios en el valor fundamental de los fondos. Valores del índice por debajo de 0.5 tienden a ilustrar una forma de segmentación donde la discriminación de precios parece no ocurrir, u ocurre para transacciones aisladas. De un total de 173 días de negociación, encontramos que el 49% de estos casos presentan segmentación de precios asociada a prácticas de discriminación de precios. Del 41% de días restantes, el 23% exhibe valores del índice igual a 0. Para estos días, si bien la tasa de interés fluctúa, todas las transacciones se clasifican como pertenecientes a un segmento único. Esto es interesante considerando que las clasificaciones se toman con dos modelos diferentes. Esto también sugiere que los márgenes negociados ocurren, pero necesariamente dentro de un segmento de precios único.

**CUADRO 6. ÍNDICE DE CAMBIO**

<i>SI</i>	<i>Número de observaciones</i>	<i>F(SI)</i>
0.1	40	0.23
0.2	46	0.27
0.3	52	0.30
0.4	56	0.32
0.5	71	0.41
0.6	98	0.57
0.7	123	0.71
0.8	147	0.85
0.9	168	0.97
1.0	173	1.00

## VIII. CONCLUSIONES

La aplicación de HMM en datos financieros no es nueva en la literatura. A menudo, el énfasis de estos modelos ha sido el de predecir series que son típicamente no estacionarias. En este estudio, usamos HMM para comprender el comportamiento de las tasas de interés en un mercado que tiene una característica particular: la asignación discrecional de líneas de crédito. Con la clasificación de tasas de interés de transacción en cuatro segmentos diferentes, podemos entender si las fluctuaciones de la tasa de interés ocurren dentro de un proceso estocástico único o son asociadas a varios procesos estocásticos (cambios en las condiciones de mercado). Además, somos capaces de inferir si dichas fluctuaciones están asociadas con restricciones de crédito potenciales. *Ex post*, esta información facilita a los analistas del mercado calificar las fluctuaciones de las tasas de interés e identificar las condiciones del mercado que pueden haber provocado un patrón particular de alternancia de segmentos.

Una ampliación natural de este trabajo debería surgir con una nueva definición del margen efectivo usando información de segmentos. Para el caso de mercados sin líneas de crédito, esto debería ser directo, pero en el caso del mercado venezolano, todavía es un problema pendiente.

*Anexo I*

Variable Dependiente: TASAS DE INTERÉS

Método: ML - ARCH (Marquardt) – Distribución del Error Generalizado (GED)

Observaciones incluidas: 26,393 después de ajustes

Convergencia alcanzada después de 16 iteraciones

	<i>Coefficiente</i>	<i>Error estándar</i>	<i>Estadístico z</i>	<i>Probabilidad</i>
DUM_1	0.673691	6.95E-05	9,688.086	0.0000
DUM_2	1.173847	6.41E-05	18,302.54	0.0000
DUM_3	2.906688	0.003174	915.8599	0.0000
DUM_4	6.074066	0.013589	446.9822	0.0000
TFLOW*DUM_3	0.078745	0.001731	45.50159	0.0000
TFLOW*DUM_4	0.282493	0.003111	90.81112	0.0000
AR(1)	0.244575	1.11E-06	219,711.1	0.0000
AR(2)	0.257011	9.80E-08	2,623,437	0.0000
AR(3)	0.149792	1.88E-07	796,640.7	0.0000
AR(4)	0.147277	1.41E-06	104,525.4	0.0000
AR(5)	0.201345	2.01E-07	1,001,262	0.0000
MA(1)	0.451789	8.39E-06	53,829.16	0.0000

(sigue)

Anexo (concluye)

	Coefficiente	Error estándar	Estadístico z	Probabilidad
MA(2)	0.269038	9.34E-06	28,804.07	0.0000
MA(3)	0.197264	7.42E-06	26,600.22	0.0000
MA(4)	0.095479	2.09E-06	45,756.30	0.0000
MA(5)	-0.061138	2.30E-06	- 26,536.58	0.0000
Ecuación de la Varianza				
C	0.813086	4.96E-09	1.64E+08	0.0000
RESID(-1)^2	26.77259	1.577800	16.96830	0.0000
GARCH(-1)	0.169478	0.002044	82.90214	0.0000
DUM_1	-0.812857	5.55E-09	-1.47E+08	0.0000
DUM_3	0.541831	0.166123	3.261619	0.0011
DUM_4	8.328266	1.427554	5.833941	0.0000
PARÁMETRO GED	0.251736	0.001269	198.3927	0.0000
R <sup>2</sup>	0.868517	Promedio de la variable de- pendiente		2.585321
R <sup>2</sup> ajustado	0.868407	Desviación estándar de la va- riable dependiente		3.087337
Error estándar de la re- gresión	1.119953	Criterio Akaike		0.410175
Suma de los cuadrados de los residuos	33,075.76	Criterio Schwarz		0.417304
Probabilidad (log like- lihood)	-5,389.874	Criterio Hannan-Quinn		0.412477

## REFERENCIAS

- Bhar, R., y S. Hamori (2004), *Hidden Markov Models: Applications to Financial Economics*, Kluwer Academic Publishers, Boston.
- Bolívar, W. (2005), *Una introducción a la teoría clásica de los modelos ocultos de Markov y modelos híbridos neuro-markovianos para problemas de clasificación*, Universidad Central de Venezuela (Mathematics degree paper).
- Hamilton, J. D. (1989), "A New Approach to the Economic Analysis of Non-Stationary Time Series and Business Cycle", *Econometrica*, vol. 57, pp. 357-84.
- Hasbrouck, J. (1991), "Measuring the Information Content of Stock Trades", *The Journal of Finance*, vol. 16, n° 1, pp. 179-207.
- Hasbrouck, J. (1993), "Assessing the Quality of a Security Market: A New Approach to Transaction-Cost Measurement". *The Review of Financial Studies*, vol. 6, n° 1, pp. 191-212.



- Pagliacci, C. (1999), *Discriminating among Decision Rules: the Persistence Criterion*, próxima publicación, Texas A&M University (Working Paper).
- Pagliacci, C. (2006), *The Venezuelan Overnight Fund Market: Understanding a Credit Constrained Limit Order Market*, próxima publicación, Banco Central de Venezuela (Working Paper).
- Rabiner, L., y B.-H. Juang (1993), *Fundamentals of Speech Recognition*, Prentice-Hall Inc., New Jersey.
- Rabiner, L. (1989), "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", *Proceedings of the IEEE*, vol. 77, pp. 257-85.
- Rydén, T., T. Teräsvirta y S. Asbrink (1998), "Stylized Facts on Daily Return Series and the Hidden Markov Model", *Journal of Applied Econometrics*, vol. 13, n° 3, pp. 217-44.
- Stoll, H. (2000), "Friction", *The Journal of Finance*, vol. 55, n° 4, pp. 1479-1514.
- Weigend, A., y S. Shi (2000), "Predicting Daily Probability Distributions of S&P500 Returns", *Journal of Forecasting*, vol. 19, pp. 375-92.
- Welch, L. (2003), "Hidden Markov Models and the Baum-Welch Algorithm", *IEEE Information Theory Society Newsletter*, vol. 53, n° 4, diciembre, pp. 1 y 10-13.



*Manfred Esquivel Monge*

# Histéresis en dolarización: evidencias de la economía costarricense

## I. INTRODUCCIÓN

En las últimas dos décadas la economía de Costa Rica, tal y como ha sucedido con otras economías pequeñas y abiertas en América Latina, se ha experimentado la dolarización no oficial (sustitución de su moneda por el dólar estadounidense, USD). Las autoridades responsables por la economía muestran su preocupación no sólo al ver que los dólares son la moneda preferida para efectos de ahorro, sino también como medio de intercambio. Una de las manifestaciones de esta preferencia es la manera en que los USD se prefieren cada vez más sobre la moneda local (el Colón, CRC) al momento de extender nuevos créditos. La preocupación surge ya que la mayoría de los agentes que aceptan dichos créditos no generan su ingreso en USD sino en CRC. A pesar de que el tipo de cambio ha estado bajo control,

*Traduce y publica el CEMLA, con la debida autorización, el artículo de M. Esquivel Monge, funcionario del Banco Central de Costa Rica. Una versión modificada de este trabajo fue presentada por el autor en septiembre de 2005 como disertación para aplicar por el grado MSc en Economía en la University of Essex, en el Reino Unido. Este estudio se presentó en la XII Reunión de la Red de Investigadores de Bancos Centrales del Continente Americano, que se realizó del 5 al 7 de noviembre de 2007, en Madrid, España, con el auspicio del Banco de España.*

en el contexto de una economía pequeña y muy abierta, el riesgo de una fuerte devaluación siempre está latente. Entonces, una situación en la que una proporción cada vez mayor de los créditos está denominada en moneda extranjera representa un escenario de alto riesgo para el sistema bancario nacional.

Esta situación de riesgo es sólo uno de los problemas que surgen al enfrentarse a una economía altamente dolarizada. El proceso también impone restricciones sobre la gestión de la política monetaria ya que altera la evolución de las medidas del dinero en circulación así como los mecanismos que gobiernan la transmisión de acciones emprendidas por los formuladores de políticas. Además, una economía con alta dolarización limita la función desempeñada por el banco central como prestamista de última instancia disminuyendo la demanda de moneda local, y reduciendo las ganancias por señoreaje.

Una de las características de la dolarización que se ha mencionado ampliamente en la literatura es que en muchos casos parece un proceso irreversible o al menos, que será muy difícil inducir a los agentes a desdolarizar. La irreversibilidad se puede analizar estudiando si algunas de las variables que explican el proceso de dolarización producen histéresis sobre ésta. Es precisamente en esta área de análisis en la que se enfoca este estudio.

El objetivo principal de este trabajo es analizar el proceso de dolarización de la economía costarricense durante los últimos 18 años para encontrar evidencia que muestre si algunas de las variables que determinan su evolución producen histéresis. Dicha evidencia podría ofrecer soporte a modelos que predican que una vez iniciado un proceso de dolarización, lo más probable es que sea irreversible.

Para alcanzar este objetivo general, el estudio está organizado como sigue: en la sección II presentaré una definición y las causas, así como las implicaciones políticas de la dolarización con un breve panorama respecto a los enfoques teóricos más populares. También se hará una breve mención a algunos trabajos empíricos. Después de esto, en la sección III presentaré una breve descripción del proceso de dolarización en Costa Rica. Algunas medidas de dolarización serán explicadas y calculadas para la economía costarricense. La sección también mencionará algunos cambios importantes en la legislación que podrían haber afectado la evolución del proceso de dolarización en el país. También, se intentará realizar una primera aproximación para encontrar evidencia de histéresis con base en las propiedades de series de tiempo de las principales variables. La sección

IV presenta un modelo que será la base de las estimaciones econométricas. Es un modelo simple basado en el comportamiento de los prestamistas y deudores que producen expresiones apropiadas para la estimación. La sección V tiene que ver con elementos de metodología; las variables dependientes e independientes serán claramente definidas; el método de estimación será explicado y, finalmente se expondrán algunas limitaciones del estudio. La sección VI se dedica a los resultados econométricos más importantes. Finalmente, con base en los resultados econométricos, las conclusiones respecto a posible evidencia de histéresis en el proceso de dolarización se alcanzarán y analizarán en la sección VII, junto con algunos comentarios finales y posibles extensiones futuras de este estudio.<sup>1</sup>

## II. DOLARIZACIÓN: CONCEPTOS PRELIMINARES

### 1. El concepto de dolarización

Como definición general, es posible indicar que es el proceso por medio del cual una moneda extranjera (ya sea el dólar estadounidense o cualquier otra) se utiliza ampliamente para satisfacer cualquiera de las tres funciones clásicas del dinero: depósito de valor, unidad de cuenta y medio de cambio.

Cuando persisten altas tasas de inflación, la moneda doméstica falla en cumplir con tales funciones. Como consecuencia, comienza a ser reemplazada por monedas que tienen amplio reconocimiento de mantener su poder adquisitivo a lo largo del tiempo. El dólar estadounidense ha sido típicamente la moneda que reemplaza el dinero local en América Latina.

Se ha observado que el proceso de dolarización comienza con el uso del dólar como depósito de valor cuando los agentes intentan proteger su patrimonio de la inflación. Eventualmente, la moneda extranjera comienza a ser utilizada como unidad de cuenta, y los precios de activos de gran valor son denominados en dólares. Finalmente hasta las transacciones se llevan a cabo con dólares como un medio de cambio. La mayoría de la literatura hace referencia a esta última etapa como *sustitución monetaria*. Tal y como refieren Calvo y Végh (1992), “la sustitución

<sup>1</sup> La versión original de este documento abarca información hasta julio de 2005. En el Apéndice se muestran resultados econométricos actualizados que es importante mencionar, no alteran las conclusiones principales a las que se llegó originalmente.

monetaria es normalmente la última etapa del proceso de dolarización”.

A pesar de que en la literatura los conceptos de dolarización y sustitución monetaria son diferentes, los utilizaré como sinónimos a menos que causaran confusión.

También es importante distinguir entre dolarización oficial y no oficial. La primera se refiere al cambio por medio del cual, el gobierno de un país (Ecuador o El Salvador por ejemplo) reemplaza legalmente la moneda local por el dólar estadounidense. El proceso no oficial se refiere a un proceso en el cual los agentes guardan dinero, hacen depósitos en el banco o denominan cualquier tipo de activo tanto en moneda local como en moneda extranjera.

## 2. Causas de la dolarización

Voy a distinguir, tal y como hacen Kikut y Méndez (2003) los factores económicos e institucionales que pueden comenzar potencialmente el proceso de dolarización.

### *a) Factores económicos*

El principal factor económico es la inflación. La razón es que la inflación erosiona el poder adquisitivo de la moneda nacional. Pero se asume que los agentes no sólo toman en consideración la inflación doméstica, sino que consideran el diferencial entre las tasas de inflación nacional y extranjera.<sup>2</sup>

Otro factor, probablemente relacionado con el diferencial entre tasas de inflación, es el diferencial entre las tasas de interés. Si consideramos una medida de dolarización como la porción de créditos denominados en dólares, los agentes tendrán un incentivo para preferir créditos denominados en dólares cuando su tasa de inflación real neta<sup>3</sup> difiera favorablemente de la tasa de inflación real en créditos denominados en moneda local. Además, si existen expectativas de una futura devaluación, los agentes también tenderán a dolarizar sus activos. Kikut y Méndez (2003) enfatizan el hecho de que el déficit de cuenta corriente

<sup>2</sup> Este es el mismo enfoque de Peiers y Wrase (1997). Utilizaremos el término *tasa de inflación extranjera* refiriéndonos a la tasa de inflación de Estados Unidos. Esto se debe a que el dólar estadounidense es la divisa más utilizada como sustituto del Colón, la moneda nacional costarricense.

<sup>3</sup> Entendemos como tasa de inflación real neta en créditos denominados en moneda extranjera el interés que resulta después de realizar el ajuste por la variación en el tipo de cambio nominal.

y la disminución de las reservas internacionales del banco central pueden inducir dichas expectativas.

Son factores que pueden iniciar el proceso, pero se ha notado que la dolarización persiste aún cuando esos factores dejan de impulsar a los agentes a cambiar monedas. Autores como Uribe (1997), Guidotti y Rodríguez (1992), Peiers y Wrase (1997) han mencionado factores que provocan la persistencia del proceso, explicaremos sus argumentos más adelante.

#### *b) Factores institucionales*

Los principales factores institucionales que Kikut y Méndez (2003) consideran son:

- Grado de liberalización en el mercado de divisas. Cuanto mayor es la liberalización del mercado, menores son los costos por intercambiar moneda doméstica por moneda extranjera.
- Qué tan abierta es la economía. Cuanto más importante sea el comercio internacional como parte del PIB, mayor será la necesidad por la moneda extranjera.
- Qué tan desarrollado está el mercado financiero doméstico. Si no hay suficiente variedad de instrumentos financieros denominados en moneda doméstica los agentes tendrán un incentivo para cambiar a instrumentos denominados en dólares.

### **3. Consecuencias de la dolarización**

Fomentar la dolarización o no, depende de sus consecuencias. A pesar de que los autores han propuesto argumentos a favor y en contra del proceso, no hay aún un acuerdo general. Aquí hago un resumen de las consecuencias que se han mencionado.

Entre los argumentos que favorecen la dolarización se afirma que, tanto como sea posible adquirir activos denominados en dólares y en moneda nacional, será posible para los inversionistas obtener una cartera con riesgo bien equilibrado.<sup>4</sup>

La reducción de costos en transacciones internacionales también favorece la dolarización. Se supone que una economía

<sup>4</sup> Ver Gomis-Poqueras, Serrano, y Somuano (2000) y Baliño, Bennette y Borensztein (1999).

dolarizada se integrará más fácilmente con el resto del mundo comercial y financiero.

Rostovsky (1992) presenta un argumento interesante. Su idea básica es que imponer costos a los agentes que mantienen moneda extranjera priva al sistema de medios de pago más útiles, con el consiguiente impacto negativo en el comercio y en el producto, reduciendo directamente la riqueza potencial de la economía.

Por otro lado, tenemos argumentos en contra de la dolarización. El fenómeno plantea preguntas muy difíciles, no sólo para propósitos analíticos sino también políticos. Desde el punto de vista del banco central, la dolarización se ha convertido en un desafío que ha introducido una dimensión adicional a la ya de por sí complicada tarea de ejecutar la política monetaria.

La principal consecuencia que el proceso puede tener es afectar la efectividad del banco central en su función de estabilizar la inflación y el producto. Los desafíos surgen principalmente porque la dolarización reduce la porción de la oferta de dinero sobre la que la autoridad tiene control. Entonces afecta la manera en que los formuladores de políticas pueden establecer objetivos basados en agregados monetarios. Además, el proceso puede afectar el mecanismo de transmisión de la política monetaria que cambia las acciones óptimas que de otro modo conllevarían a los resultados esperados.

La reducción de los ingresos del gobierno por la creación del dinero también se afirma como argumento en contra de la dolarización. Calvo (1992) describe un proceso en el que el déficit público inicial conlleva a alta inflación y luego a dolarización, lo que reduce los ingresos provenientes del señoreaje y provoca un mayor desequilibrio fiscal, conduciendo a un aumento en la inflación, por lo que el proceso se alimenta a sí mismo.

Además, cuanto mayor es la dolarización de la economía, más limitado es el papel del banco central como prestamista de última instancia, lo que puede aumentar la vulnerabilidad del sistema financiero cuando afronta una fuerte devaluación.

A pesar de que no hay acuerdo alguno respecto a si debe fomentarse la dolarización o no, se aceptan algunas recomendaciones generales cuando una economía se enfrenta al aumento en la dolarización. La vulnerabilidad inherente que se genera en el sistema financiero se puede atenuar al imponer grandes requerimientos de depósito y períodos de vencimiento mínimos para depósitos en dólares. Kikut y Méndez (2003) resumen la postura de expertos del FMI respecto al tema, ya que expertos del Fondo no encuentran diferencias significativas (respecto al



desempeño económico) entre economías dolarizadas y no dolarizadas. Consideran que la dolarización es un síntoma con el que los países deben de vivir en lugar de atacar.

#### 4. Panorama respecto a enfoques teóricos y empíricos

La dolarización es un tema ampliamente estudiado, especialmente en América Latina donde la crisis de la deuda y los episodios de alta inflación que la acompañaron a principios de la década de los ochenta provocaron un intensivo proceso de dolarización. La literatura en el campo es extensiva tanto en el estudio teórico como empírico del fenómeno. Aquí resumo algunos enfoques donde los resultados pueden ser más útiles en el análisis que se desarrolla posteriormente.

Calvo y Végh (1992) realizaron un trabajo seminal en el campo. Ofrecieron una corta revisión de los problemas analíticos más importantes durante el tratamiento del proceso de dolarización en el contexto de países en vías de desarrollo. Ellos trataron temas como los siguientes, si la sustitución monetaria debería fomentarse o no, la manera en que este fenómeno afecta la elección de anclas nominales en los programas para la reducción de la inflación; los efectos que tienen los cambios en el tasa de crecimiento de la moneda doméstica sobre el tipo de cambio real; y la interacción entre las finanzas inflacionarias y la sustitución monetaria. Ya que el objetivo aquí es más bien empírico, dichos temas no se tratarán a fondo en este documento, pero hacer referencia a ellos puede ser una herramienta útil al intentar comprender la naturaleza y muchas de las implicaciones del proceso de dolarización.

La mayor parte de la literatura teórica respecto a la sustitución monetaria se puede dividir en dos grupos principales dependiendo de la manera en que los autores tratan la presencia de dos monedas en la derivación de la demanda de dinero.

Por una parte, tenemos lo que se denomina modelos de *efectivo por adelantado*, los cuales derivan la demanda de moneda extranjera y moneda local con base en un problema estándar de maximización sujeto a la restricción de obtener efectivo antes de la liquidación de las transacciones. Algunas de las características típicas de estos modelos incluyen la sustitución perfecta de los activos y las monedas en las que se encuentran denominados. Ambas monedas se utilizan como medio de intercambio y no hay costos adicionales para adquirir la moneda extranjera.

Por otra parte, tenemos modelos basados en costos de transacciones. Derivan demandas por saldos de dinero dentro de

un contexto donde los agentes maximizan una función de utilidad afrontando restricciones que incluyen los costos de adquirir efectivo. Dichos costos aumentan en el nivel de consumo y disminuyen el nivel de saldos monetarios.

Un buen ejemplo de un modelo efectivo por adelantado se describe en el documento de Uribe (1997). El construye un modelo donde es caro comprar bienes con moneda extranjera y los costos disminuyen en el nivel de experiencia que la economía acumula al realizar transacciones con ese dinero. El modelo entonces muestra histéresis en la velocidad del dinero. Como resultado importante, para que los agentes dejen de usar la moneda extranjera, este dinero no necesariamente tiene que estar dominado por el dinero local en tasa de rendimientos. Al final, el autor concluye que las tasas de inflación que comenzaron el proceso no necesariamente implican un estado estacionario en el que el dinero local cae totalmente en desuso.

Lebre de Freitas (2004) analiza la relación entre el dinero y la inflación en el contexto de una economía pequeña y abierta donde la moneda doméstica y extranjera son sustitutos perfectos como medio de pago. Este modelo concluye que si la ruta de crecimiento en la oferta de moneda doméstica es tal que es óptimo que los agentes cambien la moneda, habrá un período de ajuste en donde la inflación doméstica tenderá a igualar la inflación extranjera. Además, el modelo captura (sin asumir la existencia del costo de la dolarización) que un aumento temporal en el tasa de inflación puede producir un efecto permanente en el uso de moneda extranjera, es decir, la inflación produce histéresis en el proceso de dolarización.

En la literatura muy a menudo se menciona la publicación de Guidotti y Rodríguez (1992). Presenta un modelo en el cual la dolarización evoluciona de acuerdo con la existencia de costos asociados a cambios en la moneda que se utiliza para el pago de las transacciones. La diferencia entre las tasas de inflación nacional y extranjera se asume como la principal causa del principio de este proceso. Entonces los costos definen una banda en el diferencial de inflación sobre el que existe un incentivo para cambiar de moneda local a internacional y por debajo del valor mínimo de la banda existe el incentivo para desdolarizar.

La posibilidad de un proceso inverso a la dolarización (desdolarización) es perfectamente plausible y de hecho generada en los modelos mencionados. Pero lo que se ha observado en estudios empíricos es que la dolarización parece tener un elemento intrínseco de irreversibilidad.

La irreversibilidad se genera en Lebre de Freitas (2004)

porque para iniciar un proceso de desdolarización, la moneda local requiere volverse más atractiva que la moneda extranjera, lo que es posible sólo si la inflación doméstica se presenta persistentemente menor a la inflación externa. Con base en lo planteado por el autor, esto es muy poco probable al menos en las economías latinoamericanas.

Tandon y Wang (2003) formalizan un enfoque muy interesante respecto a la persistencia de la dolarización. Incluyen el aspecto tan importante de la confianza de los agentes en la moneda doméstica, aunada a la confianza implícita en la política monetaria, como una variable determinante en el proceso. Como resultado del comportamiento de la variable de confianza es posible generar la irreversibilidad o la histéresis. Los autores presentan un modelo para analizar la sustitución monetaria en un entorno donde el dinero doméstico podría sufrir una falta de confianza. Consideran dos variantes: en primer lugar, la situación en la que el nivel de confianza es exógeno y constante; entonces obtienen un estado estacionario donde los agentes realizan transacciones con una proporción estable de dinero local y moneda extranjera.

Por otra parte, cuando introducen la confianza de los agentes como una variable endógena, la dinámica de la sustitución monetaria se ve gobernada por la evolución de las creencias de los agentes. La economía finalmente termina en un estado estacionario de moneda dual que puede tener un alto o bajo nivel de dolarización. El factor relevante para nuestros propósitos es que las creencias de los agentes exhiben persistencia (los agentes tienden a tener memoria de largo plazo) por lo que es posible generar la histéresis en la dolarización.

Mencionemos ahora algunos trabajos empíricos en el área. La mayoría de las publicaciones se han enfocado en el caso de países latinoamericanos donde, debido a la liberalización financiera que tuvo lugar durante la década de los años ochenta e incluso en los noventa, y las crisis inflacionarias que afectaron la región, los dólares se convirtieron rápidamente en un sustituto de la moneda doméstica.

El estudio de Peiers y Wrase (1997) se encuentra en los que se basan en el costo de las transacciones. Los autores exponen un modelo en el cual una red de externalidades, que aparece como resultado del uso creciente de la moneda extranjera, produce eficiencias que determinan la persistencia observada en el proceso de dolarización. Intentando poner a prueba las implicaciones del modelo, utilizan datos de créditos denominados en dólares de una región particular de Bolivia y realizan una

estimación *probit* de la probabilidad de la emisión de un nuevo crédito denominado en dólares y encuentran, por ejemplo, que algunas variables utilizadas como sustitutos para la credibilidad de la política económica resultaron significativas al explicar el proceso de dolarización.<sup>5</sup> Además las variables de trinquete basadas en la inflación y la variabilidad de la inflación también se encontraron como significativas, ofreciendo evidencias de histéresis en el proceso. Finalmente concluyeron que una política de estabilización creíble no es suficiente para alcanzar la desdolarización.

Kikut y Méndez (2003) describen el proceso de dolarización de la economía costarricense en términos de la evolución de las regulaciones jurídicas; también desarrollan un enfoque económico en donde se explican diversas medidas de la dolarización en términos de rendimiento de activos denominados en moneda extranjera, la devaluación del tipo de cambio esperada y el diferencial de las tasas de interés. Todas esas variables se encontraron significativas al momento de explicar la evolución de las medidas de dolarización tomadas en consideración por los autores. Además, encontraron significativas las variables de trinquete, proporcionando evidencia de la histéresis en la dolarización al menos cuando se medía en términos de coeficientes de depósitos bancarios respecto a liquidez total.

### III. DESCRIPCIÓN DEL PROCESO DE DOLARIZACIÓN EN COSTA RICA

#### 1. Regulación

Distinto a la mayoría de los países latinoamericanos, en la última mitad del siglo XX, Costa Rica no ha sufrido ningún episodio de confiscación de depósitos en moneda extranjera ni un proceso forzado de desdolarización.<sup>6</sup>

Desde el principio de la existencia del banco central en 1950 la legislación respecto a la creación de depósitos en moneda extranjera ha sido bastante abierta. Desde entonces, los bancos tuvieron la autorización de recibir depósitos en cualquier moneda, aunque no estaba permitido pagar interés sobre ellos.

<sup>5</sup> Este resultado está implícito en el enfoque teórico de Tandon y Wang (2003).

<sup>6</sup> Ejemplos de dichos episodios son Bolivia (1982), México (1982) y Perú (1985) donde los depósitos denominados en dólares fueron convertidos a moneda local en una acción *de facto* por parte del gobierno.

Se introdujeron reformas importantes en el campo en 1992 y 1996. Resumo a continuación las que consideré más importantes:

- En 1992, se eliminó el monopolio del Banco Central para comprar y vender divisas al permitirse también hacerlo a los bancos comerciales.
- La reserva legal de los depósitos en dólares que los bancos mantienen en el Banco Central se redujo del 100% al 5% entre 1992 y 1996.
- Se alcanzó mayor protección a la creación de depósitos en divisas en 1995 cuando se estableció que todas las instituciones financieras reguladas debían forzosamente pagar cualquier obligación en la moneda original acordada.
- Antes de 1996 sólo se permitía a bancos comerciales públicos ofrecer depósitos en cuenta corriente en divisas. El cambio abrió la opción a intermediarios privados. También se liberalizaron las tasas de interés pagadas por depósitos en dólares que anteriormente fijaba el Banco Central.

Además de esos cambios, algunos argumentos existen respecto al régimen del tipo de cambio en Costa Rica. Hasta 1983 lo normal era un tipo de cambio fijo con ajustes muy poco comunes. Después de la crisis de la deuda a principios de la década de los ochenta, el sistema cambió hacia uno con muy pequeños pero periódicos ajustes al tipo de cambio. Este sistema, conocido como *minidevaluación*, es el que permanece al momento de este estudio.<sup>7</sup> La idea básica es ajustar el tipo de cambio de acuerdo con la evolución de precios en Costa Rica y Estados Unidos de modo que el poder adquisitivo de la paridad de la moneda nacional permanece constante.

## 2. Magnitud de la dolarización en Costa Rica

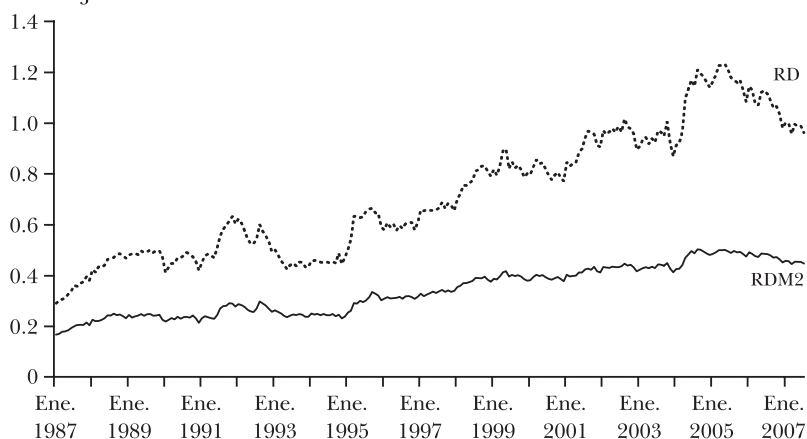
Los enfoques empíricos al fenómeno de la dolarización y principalmente en su última etapa de sustitución monetaria tuvieron un problema fundamental: no existe información exacta acerca de la cantidad real de moneda extranjera que se utiliza como medio de cambio. Entonces, cualquier estudio descriptivo o econométrico tendría que lidiar con una variable básicamente no observada.

<sup>7</sup> Se manejó con poco éxito la flotación dirigida en 1992.

La práctica ha sido aproximar la cantidad de dolarización utilizando algunos coeficientes de activos denominados en dólares para medir el total de la liquidez, la cual es usualmente aproximada por  $M2$ .<sup>8</sup> El fenómeno también puede medirse utilizando coeficientes de créditos denominados en dólares con respecto a los créditos totales en la economía.

Para efectos descriptivos, hablaré más en detalle de Kikut y Méndez (2003). Utilizaron tres tipos de coeficientes para describir diferentes fases del proceso de dolarización. La sustitución de activos, la sustitución monetaria y la dolarización en general. Voy a agregar a la sustitución monetaria indicadores de la participación de créditos denominados en dólares en el total de los créditos,<sup>9</sup> que se utilizará más adelante como variable dependiente para realizar las regresiones.

**GRÁFICA A.I.** COSTA RICA: COEFICIENTES DE DOLARIZACIÓN, ENERO DE 1987 A JULIO DE 2007



*a) Indicadores de sustitución de activos*

Coefficiente de depósitos a plazo:

$$RTD = \frac{\text{Depósitos a plazo denominados en dólares}}{\text{Depósitos a plazo denominados en colones}}$$

<sup>8</sup> En este estudio  $M2$  incluirá la suma de la base monetaria ( $M1$ , o dinero de alto poder) más todos los depósitos de ahorro con un período de vencimiento de 30 días o más.

<sup>9</sup> Los indicadores presentados aquí se construyeron con datos mensuales obtenidos en septiembre de 2005 de la página Web del Banco Central de Costa Rica ([www.bccr.fi.cr](http://www.bccr.fi.cr)). Para créditos denominados en dólares las cantidades fueron convertidas al tipo de cambio promedio al final de cada período.

Coefficiente de depósitos de ahorro:

$$RSD = \frac{\text{Depósitos en cuenta de ahorro en dólares}}{\text{Depósitos en cuenta de ahorro en colones}}$$

La gráfica I muestra la evolución de estos dos indicadores en los últimos 18 años. Es posible observar una tendencia creciente en ambos coeficientes a pesar de que este comportamiento no ha sido el mismo a lo largo de toda la muestra. Específicamente *RTD* muestra períodos de tendencia decreciente en 1993 y 2003. Ambas series comienzan cerca del 0.5, alcanzan niveles de hasta 1.33 en *RTD* y 2.02 en *RSD*, de modo que aumentaron más del 100% y 3 veces.

b) *Indicadores de sustitución monetaria*

Coefficiente de depósitos en cuenta corriente:

$$RCA = \frac{\text{Depósitos en cuenta corriente en dólares}}{\text{Depósitos en cuenta corriente en colones}}$$

Coefficiente de créditos:

$$RL = \frac{\text{Créditos denominados en dólares}}{\text{Total de créditos}}$$

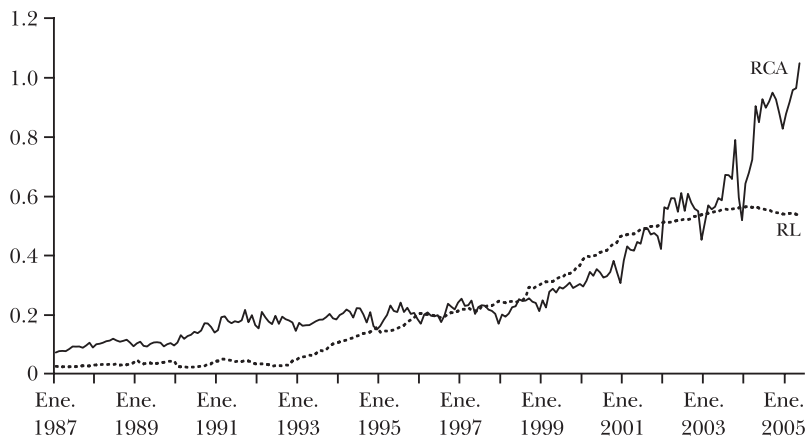
La gráfica II ilustra su evolución. Una vez más se observa una clara tendencia de crecimiento. Es importante observar que en los últimos 18 años *RCA* aumentó 14 veces, de sólo 0.07 a un valor máximo de 1.05 alcanzado al final de la muestra. Este coeficiente comenzó a acelerar su tendencia de crecimiento desde el año 2000 y no se ha detenido. Este comportamiento, junto con el hecho de que *RL* parece haber acelerado desde 1999, podría indicar que el proceso de dolarización comenzó la etapa de la sustitución monetaria al final de la década de los años noventa. Observe, sin embargo, que *RL* muestra un ligero retroceso desde abril de 2004.

c) *Indicadores de la dolarización en general*

Coefficiente de depósitos a M2:

$$RDM2 = \frac{\text{Total de depósitos en dólares}}{\text{Liquidez total en colones (M2)}}$$

**GRÁFICA II. COSTA RICA: COEFICIENTES DE SUSTITUCIÓN MONETARIA, ENERO DE 1987 A MAYO DE 2005**

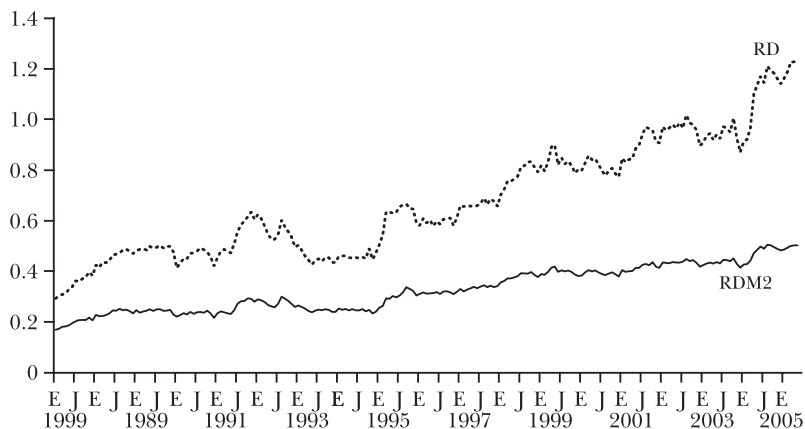


Coefficiente de depósitos:

$$RD = \frac{\text{Total de depósitos en dólares}}{\text{Total de depósitos en colones}}$$

La gráfica III muestra cómo han evolucionado los dos coeficientes durante los últimos 18 años. En general, podemos observar la misma tendencia de crecimiento. Sin embargo, en el caso del *RDM2* el crecimiento se distribuye de manera más suave con casi total ausencia de retrocesos fuertes. Mientras que *RD* comienza en 0.29 y termina con un nivel máximo de 1.23, en mayo de 2005 (un crecimiento de más de 3 veces), *RDM2*

**GRÁFICA III. COSTA RICA: COEFICIENTES DE DOLARIZACIÓN, ENERO DE 1987 A MAYO DE 2005**





muestra un crecimiento de casi 2 veces, comenzando en 0.17 y terminando en un nivel de 0.5 al final de la muestra.

El cuadro 1 resume lo que reflejan las gráficas descritas anteriormente. El cuadro se divide en cuatro columnas principales representando diferentes períodos de tiempo. De izquierda a derecha hay tres subperíodos y en la extrema derecha tenemos estadísticas de la muestra completa. La razón para tal división es analizar principalmente dos aspectos: en primer lugar, el impacto de los cambios realizadas a la legislación en 1996 sobre la evolución de los coeficientes. En segundo lugar, observar lo que sucedió con nuestros indicadores después de diciembre de 2002 que fue el final de la muestra que utilizaron Kikut y Méndez (2003).

**CUADRO 1.** COSTA RICA: COEFICIENTES DEL PROMEDIO SIMPLE DE DOLARIZACIÓN Y SUS COEFICIENTES DE VARIANZA, ENERO DE 1987 A MAYO DE 2005

Coefi- cientes	Ene 1987 - Jun 1996		Ago 1996 - Dic 2002		Ene 2003 - May 2005		Ene 1987 - May 2005	
	Prom	V. Coeff. (%)	Prom	V. Coeff. (%)	Prom	V. Coeff. (%)	Prom	V. Coeff. (%)
RTD	0.66	17.2	1.06	16.9	1.10	7.2	0.86	28.8
RSD	0.76	21.4	0.99	10.7	1.63	14.6	0.96	33.6
RCA	0.16	28.7	0.34	37.6	0.77	21.5	0.30	74.4
RL	0.07	78.4	0.36	32.5	0.55	1.7	0.24	83.5
RDM2	0.25	13.5	0.38	9.4	0.45	6.7	0.32	26.5
RD	0.49	17.1	0.81	14.2	1.05	12.1	0.68	34.0

FUENTE: Cálculo propio con datos del Banco Central de Costa Rica.

NOTA: V. Coeff. (%) representa cuánto es la desviación estándar como porcentaje del promedio.

Observe que el promedio de todos los coeficientes aumenta a medida que ordenamos los períodos cronológicamente. A pesar de que estamos añadiendo treinta y seis observaciones al principio de la muestra y treinta al final, es la misma tendencia general que se encuentra en Kikut y Méndez (2003).

Lo que es más interesante señalar tiene que ver con las medidas que incluimos como indicadores de volatilidad, por ejemplo V. Coeff. Todos excepto uno de los coeficientes reduce su volatilidad entre el primero y segundo período y lo mismo sucede entre el segundo y el tercero. Este hecho, aunado al aumento sostenible de los coeficientes nos habla de un proceso de dolarización que está creciendo pero gradualmente a un ritmo menos dinámico. Este aspecto no se revela claramente en el

trabajo de Kikut y Méndez (2003). Ellos informan ligeros aumentos en volatilidad entre los períodos 1990-1996 y 1997-2002. Una vez que extendemos la muestra de 1987 a mayo de 2005 es posible ver dicha desaceleración.

Resulta evidente que hay un proceso de dolarización actualmente en Costa Rica. Comenzó aún incluso antes de la primera observación de nuestra muestra (observe que los coeficientes de sustitución de activos estaban por encima de 0.4 en enero de 1987). Pero ¿qué tan dolarizada está la economía costarricense comparada con los estándares? Es posible usar como referencia un estudio del FMI escrito por Baliño, Bennette, y Borensztein (1999). Consideran una economía como altamente dolarizada cuando el coeficiente de depósitos denominados en dólares frente a una definición general del dinero (en nuestro caso *RDM2*) excede 0.3. Siguiendo nuestro cálculo, Costa Rica ha caído en esta categoría desde mediados de 1995 y nunca la re-vertió (*RDM2* alcanzó 0.5 en mayo de 2005).

### 3. Algunas evidencias iniciales de histéresis

El principal objetivo de este estudio es encontrar evidencia de histéresis en el proceso de dolarización en general, enfocando en particular en una medida de sustitución monetaria (la participación de créditos totales denominados en moneda extranjera). Antes de presentar el modelo que ofrece microfundamentos y el análisis econométrico subsiguiente, aquí presento algunos análisis iniciales, casi descriptivos, que podrían ayudarnos en alcanzar nuestro objetivo.

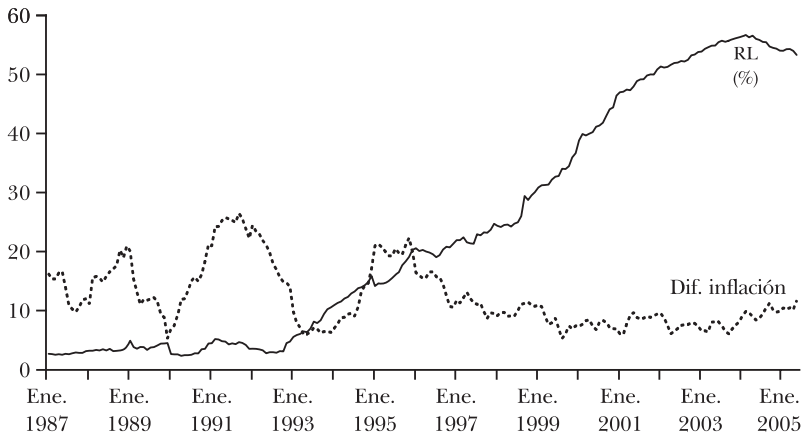
Como expliqué en el apartado 2, la literatura tradicionalmente utilizó la inflación y las diferencias entre las tasas de interés como las principales causas de la dolarización. Sin embargo, evidencias en muchos países muestran que el proceso persiste aún cuando disminuye la inflación.<sup>10</sup>

La gráfica IV muestra la evolución durante los últimos 18 años del coeficiente de créditos denominados en dólares con respecto al total de créditos en términos porcentuales (*RL%*) y la diferencia entre la tasa de inflación de Estados Unidos y de Costa Rica (*Inf. diff*).

Observe que después de 1996, el diferencial de inflación no sólo disminuyó en promedio sino que incluso se tornó relativamente estable cerca del 10%. Por otra parte, nuestra variable *proxy* de dolarización fue muy estable hasta inicios de 1993,

<sup>10</sup> Ver por ejemplo Guidotti y Rodríguez (1992).

**GRÁFICA IV.** COSTA RICA: DOLARIZACIÓN Y DIFERENCIA ENTRE LA TASA DE INFLACIÓN DOMÉSTICA Y LA ESTADOUNIDENSE, ENERO DE 1987 A MAYO DE 2005



después comenzó a crecer incluso cuando el diferencial de inflación se volvió menor y estable. El caso de Costa Rica, similar a lo que se informó para Bolivia en Guidotti y Rodríguez (1992) es un ejemplo donde la dolarización tuvo un fuerte aumento en períodos de inflación estable e incluso descendente.

Examinemos más en detalle estas dos series para establecer de manera más técnica sus propiedades de series de tiempo. El cuadro 2 muestra resultados a partir de pruebas de raíz unitaria para el diferencial de inflación (Diff. Inf) y la variable *proxy* de dolarización (*RL%*). Se realizaron las pruebas incluyendo un intercepto y una tendencia de tiempo para ambas variables.

**CUADRO 2.** PRUEBA DICKEY-FULLER AUMENTADO PARA RAÍZ UNITARIA PARA LA DIFERENCIA EN LA INFLACIÓN Y LA VARIABLE PROXY DE DOLARIZACIÓN

	Diff. Inf.	RL (%)	Valor crítico 1% <sup>a</sup>	Valor crítico 5% <sup>a</sup>	Valor crítico 10% <sup>a</sup>
Prueba DF Aumentado por niveles	-2.710	-2.015	-4.003	-3.432	-3.139
Prueba DF Aumentado para primeras diferencias	-7.281	-5.931	-4.003	-3.432	-3.139

FUENTE: Banco Central de Costa Rica.

<sup>a</sup> Valor crítico de MacKinnon, valor para rechazar hipótesis nula de raíz unitaria.

Observe que la prueba es incapaz de rechazar la hipótesis nula de una raíz unitaria en niveles para ambas series en todos los niveles de significancia. Por otra parte cuando se aplica la prueba

a la primera diferencia de las series, lo cual se muestra en la segunda fila del cuadro, la hipótesis nula de una raíz unitaria se rechaza. Entonces tenemos evidencia de que ambas series tienen una raíz unitaria en niveles pero no en diferencias, es decir, son integradas de orden uno.

Dado que ambas series son no estacionarias es posible aplicar la prueba de Johansen para cointegración. El cuadro 3 resume lo que se encontró después de tal procedimiento. La prueba, que especifica a  $RL\%$  como variable dependiente, incluye también un intercepto y tendencia de tiempo en la ecuación de cointegración (CE).

**CUADRO 3. PRUEBA DE COINTEGRACIÓN DE JOHANSEN**

	<i>Razón de Verosimilitud</i>	<i>Valor Crítico 5%</i>	<i>Valor Crítico 1%</i>	<i>No hipotético de CE(s)</i>
Eigenvalue				ninguno
0.080901	22.38	25.32	30.45	

FUENTE: Banco Central de Costa Rica.

NOTAS: La razón de verosimilitud rechaza cualquier cointegración al nivel de significancia de 5%. Los valores críticos producidos por E-views son los de Osterwald-Lenum (1992), no los tabulados en el trabajo de Johansen, S. y Juselius, K. (1990).

La prueba es para la hipótesis nula de no cointegración; por lo tanto fallamos en rechazarla en los dos niveles de significancia, lo cual produce evidencia de que las dos series no están cointegradas.

Surgen dos puntos clave del análisis de los cuadros 2 y 3. En primer lugar, dado que las variables no están cointegradas, es posible decir que la dolarización no se explica completamente por medio de la evolución del diferencial de inflación; deben incluirse otros elementos para mayor especificación del fenómeno. En segundo lugar, debido a que la variable *proxy* de la dolarización tiene raíz unitaria en sus valores por niveles, los choques que la afectan tienen efectos permanentes. Esto constituye una señal de que algún elemento de histéresis debe estar involucrado en la evolución de la dolarización.

Los especialistas en el área<sup>11</sup> han argumentado que el desarrollo de nuevos instrumentos financieros como defensa en contra de la alta inflación desempeña una función clave al generar histéresis. La posibilidad de hacer depósitos denominados en moneda extranjera, como una manifestación de esta

<sup>11</sup> Ver por ejemplo a Dornbusch, Sturzenegger y Wolf (1990).

adaptación financiera, disminuye la demanda de dinero doméstico para una tasa de interés dada. Entonces, debido a que este proceso de creación de nuevos instrumentos requiere aprendizaje y obviamente no está libre de costo, después de realizar la *inversión*, los agentes se sujetarán a estos nuevos instrumentos aún cuando la inflación disminuya y permanezca bajo control.

La falta de confianza en la política económica doméstica podría también ser considerada para la persistencia que se observa en la dolarización. Ejemplos como Argentina, Perú y Bolivia,<sup>12</sup> constituyen evidencia de que una política creíble es difícil de alcanzar y que la credibilidad por sí misma no es suficiente para comenzar la desdolarización. En esos países la hiperinflación en la década de los años ochenta se controló y las medidas de dolarización no revirtieron su tendencia al alza.

En el caso de Costa Rica, a pesar de que la inflación está bajo control después de altos niveles a finales de los ochenta y principios de los noventa, un problema fundamental de déficit público persistente hace imposible que las autoridades monetarias alcancen el objetivo de tener inflación de un dígito. Como resultado, la confianza en las políticas económicas no es precisamente el sentimiento general de los agentes costarricenses hoy en día. Veremos si nuestras estimaciones muestran que esta falta de confianza tiene un impacto significativo al explicar lo que elegimos como variable *proxy* de la dolarización.

El modelo presentado en la siguiente sección fue seleccionado debido a que se ajusta a la evidencia anterior. Nos proporciona microfundamentos que justifican la inclusión de elementos adicionales (aparte del diferencial de inflación) para explicar la evolución del proceso de dolarización y permitimos utilizar la prueba para histéresis utilizando variables de tipo *ratchet* en la especificación.

#### IV. DOLARIZACIÓN: UN MODELO DE COSTO DE TRANSACCIÓN

Este modelo presentado aquí seguirá de cerca el enfoque desarrollado por Peiers y Wrase (1997). La idea básica del modelo es intentar explicar la manera en que se toman las decisiones de extender y tomar créditos en moneda doméstica o extranjera. Sus implicaciones nos llevarán entonces a expresiones adecuadas

<sup>12</sup> Fenske y Melvin (1992) describen el caso de la economía boliviana en donde se mantuvo una alta inflación, incluso aumentó después de un exitoso plan de estabilización.

con fines de estimación y por lo tanto, nos permitirán probar la presencia de la histéresis y la significancia de otras variables al explicar el proceso de dolarización.

Consideraremos una economía con las siguientes características:

- Existe un único bien y dos monedas para realizar las transacciones, extranjera y doméstica, a saber USD (\$) y CRC (¢).
- Existen muchos hogares idénticos, cada uno compuesto por un vendedor de bienes, un comprador y un intermediario financiero.
- En cada período de tiempo los compradores adquieren bienes del vendedor utilizando CRC o USD que obtienen de un préstamo del intermediario financiero.

Vamos a considerar por separado el comportamiento de cada uno de estos agentes:<sup>13</sup>

*El vendedor:* este agente recibirá una dotación inicial total de  $Y_i^T$  del bien a asignar como sigue:

$$(1) \quad Y_i^T = Y_i + Y_i^*,$$

donde  $Y_i(Y_i^*)$  se vende en el mercado doméstico utilizando ¢ (\$) a un precio de  $P_i(P_i^*)$ .<sup>14</sup>

*El comprador:* antes de comprar, este agente debe tomar prestado  $b_i(b_i^*)$  unidades de ¢ (\$) del intermediario financiero. Entonces los saldos de monedas domésticas (extranjeras) se utilizan para comprar  $C_i(C_i^*)$  unidades del bien a un precio de  $P_i(P_i^*)$ .

Se asumirá que el vendedor tiene algún costo real cuando realiza transacciones, estos costos se especificarán como sigue:

$$\Omega^C \left( C_i, \frac{b_i}{P_i}, T_i^d, POL_i \right),$$

cuando se denominen las transacciones en colones;

$$\Omega^{C^*} \left( C_i^*, \frac{b_i^*}{P_i^*}, T_i^d, POL_i \right),$$

<sup>13</sup> A continuación un \* indicará una variable como denominada en USD. Utilizaremos caracteres en minúscula y mayúscula para especificar respectivamente, las cantidades demandadas y ofrecidas.

<sup>14</sup> Para mayor simplicidad, estamos omitiendo la posibilidad de comercio internacional.

cuando se denominen las transacciones en dólares; donde:<sup>15</sup>

- $\Omega_{c_t}^C > 0$ ,  $\Omega_{c_t^*}^{C^*} > 0$ , es decir, el costo del comprador aumenta en el volumen de las transacciones.
- $\Omega_{b_t/P_t}^C < 0$ ,  $\Omega_{b_t^*/P_t^*}^{C^*} < 0$ , es decir, el costo del comprador disminuye en saldos reales utilizados para realizar las transacciones.
- $T_t^d$  es un índice de factores que reduce (aumenta) el costo marginal de realizar negocios en \$ ( $\mathcal{E}$ ). Entonces, el supuesto es que los costos del comprador dependen de la red de externalidades del uso del dólar que está representado aquí por medio de  $T_t^d$ . Entonces tendremos  $\Omega_{b_t/P_t, T_t^d}^C \geq 0$  y  $\Omega_{b_t^*/P_t^*, T_t^d}^{C^*} < 0$ . En palabras, a medida que aumenta  $T_t^d$ , los agentes ganan experiencia al tratar con divisas, un hecho que va a reducir (aumentar) el costo marginal de tomar préstamos y gastar en \$ ( $\mathcal{E}$ ). Observe que los factores involucrados en  $T_t^d$  pueden incluir variables con inercia, que entonces podrían causar histéresis en el uso de la moneda extranjera.
- $POL_t$  representa un índice de credibilidad en la política económica que afecta los costos de transacción independientemente de los efectos de la red de externalidades. Se supone que el costo marginal de pedir préstamos en  $\mathcal{E}$  es mayor cuando la credibilidad en la política económica es baja, y lo opuesto sucede para el costo los préstamos en \$. Entonces, considerando un aumento en  $POL_t$  como una declinación en la credibilidad, tendremos  $\Omega_{b_t/P_t, POL_t}^C > 0$  y  $\Omega_{b_t^*/P_t^*, POL_t}^{C^*} \leq 0$ .

*El intermediario financiero:* el intermediario distribuirá la riqueza del hogar inicial siguiendo esta expresión:

$$(2) \quad M_t + \rho_t M_t^* = B_t + B_t^*,$$

donde  $\rho_t$  es el tipo de cambio nominal en el período  $t$ ;  $M_t$  y  $M_t^*$  representan los saldos monetarios iniciales en moneda doméstica y extranjera;  $B_t$  y  $B_t^*$  son los créditos en moneda doméstica y extranjera que los intermediarios ponen a disponibilidad de los compradores en el período  $t$ . Dichos créditos cobran un tipo de interés de  $R_t$  y  $R_t^*$ , respectivamente.

También se asume que, tal y como sucede en el caso del comprador, el intermediario incurre en costos de transacción reales cuando otorga un crédito. Estos costos se tomarán como sigue:

<sup>15</sup> Notación derivada parcial: debemos entender  $A_\alpha$  como la derivada parcial de  $A$  respecto a  $\alpha$ .

$$\Omega^L \left( \frac{B_t}{P_t}, T_t^s, POL_t \right),$$

cuando el crédito se exprese en  $\mathcal{C}$ .

$$\Omega^{L^*} \left( \frac{B_t^*}{P_t^*}, T_t^s, POL_t \right),$$

cuando el crédito se exprese en  $\mathcal{S}$ .

Donde se supone que  $\Omega^L_{B_t/P_t} > 0$  y  $\Omega^{L^*}_{B_t^*/P_t^*} > 0$ , es decir, la función del costo es creciente en el volumen de la oferta de créditos reales;  $T_t^s$  representa variables que pueden reducir (aumentar) el costo marginal de la transacción al realizar un crédito adicional denominado en  $\mathcal{S}$  ( $\mathcal{C}$ ), entonces tenemos que  $\Omega^{L^*}_{B_t^*/P_t^*, T_t^s} \leq 0$  y  $\Omega^L_{B_t/P_t, T_t^s} \geq 0$ , se asume que los factores que afectan los costos marginales surgen de una red de externalidades que se crea cuando los agentes hacen tratos adicionales con moneda extranjera, reduciendo costos intrínsecos por el uso de dicha moneda;  $POL_t$  captura, como en el caso del comprador, un índice de credibilidad en la política económica, suponemos  $\Omega^{L^*}_{B_t^*/P_t^*, POL_t} < 0$  y  $\Omega^L_{B_t/P_t, POL_t} > 0$ , es decir, cuando declina la credibilidad (que consideramos como un aumento en  $POL_t$ ), el costo de un crédito adicional en  $\mathcal{S}$  disminuye, y sucede lo contrario en el caso de un crédito denominado en  $\mathcal{C}$ .

Después de cada período los agentes llevan la siguiente cantidad de dinero al hogar:<sup>16</sup>

	<i>Moneda doméstica</i>	<i>Moneda extranjera</i>
Vendedor	$P_t Y_t$	$P_t^* Y_t^*$
Comprador	$b_t - P_t [C_t + \Omega^C(t)] - (1 + R_t) b_t$	$b_t^* - P_t^* [C_t^* + \Omega^{C^*}(t)] - (1 + R_t^*) b_t^*$
Intermediario	$(1 + R_t) B_t - P_t \Omega^L(t)$	$(1 + R_t^*) B_t^* - P_t^* \Omega^{L^*}(t)$

Agregando las posiciones de efectivo de los miembros se obtiene el saldo monetario del hogar:

$$(3) \quad M_{t+1} = (1 + R_t) B_t - P_t \Omega^L(t) + P_t Y_t + b_t - P_t [C_t + \Omega^C(t)] - (1 + R_t) b_t$$

<sup>16</sup> Observe que las cantidades corresponden a los rendimientos de los créditos netos de los costos de transacción.



$$(4) \quad M_{t+1}^* = (1 + R_t^*)B_t^* - P_t^* \Omega^{L^*}(t) + P_t^* Y_t^* + b_t^* - P_t^* [C_t^* + \Omega^{C^*}(t)] - (1 + R_t^*)b_t^*$$

Las ecuaciones (1) a (4) representan restricciones que los hogares deben satisfacer al maximizar una función de utilidad que

se asume generalmente expresada como  $U = E_t \left[ \sum_0^\infty \phi^i \lambda(C_t, C_t^*) \right]$ ,

donde  $E_t$  es el operador de expectativa condicionado a la información disponible hasta  $t$ . Entonces los hogares eligen patrones de consumo, suministro de cantidades para el vendedor y saldos de créditos para el intermediario tomando como dado los saldos de dinero iniciales ( $M_t$  y  $M_t^*$ ), precios, tipos de interés y factores  $T_t^s, T_t^d, POL_t$ .

Condiciones de optimalidad para cada hogar implican lo siguiente:

$$(5) \quad R_t = \Omega_{b_t/P_t}^C \left( C_t, \frac{b_t}{P_t}, T_t^d, POL_t \right)$$

$$(6) \quad R_t^* = \Omega_{b_t^*/P_t^*}^{C^*} \left( C_t^*, \frac{b_t^*}{P_t^*}, T_t^d, POL_t \right)$$

$$(7) \quad E_t(\rho_t/\rho_{t+1}) = \frac{\left[ (1 + R_t^*) - \Omega_{B_t^*/P_t^*}^{L^*} \left( \frac{B_t^*}{P_t^*}, T_t^s, POL_t \right) \right]}{\left[ (1 + R_t) - \Omega_{B_t/P_t}^L \left( \frac{B_t}{P_t}, T_t^s, POL_t \right) \right]}$$

Observe que las condiciones (5) y (6) son casi intuitivas; dicen que el comprador tomará préstamos hasta el punto en que el valor de las transacciones marginales de cada moneda igualen el costo respectivo de la tasa de interés. Por otro lado, la condición (7) especifica paridad en las tasas de interés cuando existen costos de transacción y se asocia con la oferta de créditos del intermediario.

Ahora necesitamos especificar una forma funcional de los costos de transacción para obtener una relación que nos permita hacer cálculos. Entonces, las formas supuestas de los costos son:

$$(8) \quad \Omega^C = (C_t)^{\alpha_1} (b_t/P_t)^{-\alpha_2} (T_t^d)^{-\alpha_3} (POL_t)^{-\alpha_4}$$

$$(9) \quad \Omega^{C^*} = (C_t^*)^{\beta_1} (b_t^*/P_t^*)^{-\beta_2} (T_t^d)^{\beta_3} (POL_t)^{\beta_4},$$

con:

$$\alpha_i > 0 \text{ para } i=1,2,4 \text{ y } \alpha_3 \geq 0$$

$$\beta_i > 0 \text{ para } i=1,2,3 \text{ y } \beta_4 \geq 0$$

Ahora podemos tomar las derivadas correspondientes de (8) y (9) y conectarlos con las condiciones (5) y (6), algunos reacomodos nos llevan a las demandas de los créditos:

$$(10) \quad (b_t/P_t) = \left[ \frac{\alpha_2 (C_t)^{\alpha_1}}{R_t (T_t^d)^{\alpha_3} (POL_t)^{\alpha_4}} \right]^{\frac{1}{\alpha_2+1}}$$

$$(11) \quad (b_t^*/P_t^*) = \left[ \frac{\beta_2 (C_t^*)^{\beta_1} (T_t^d)^{\beta_3} (POL_t)}{R_t^*} \right]^{\frac{1}{\beta_2+1}}$$

Por lo tanto, la demanda de créditos denominados en moneda doméstica (extranjera) está disminuyendo en el costo directo del préstamo  $R_t(R_t^*)$ , y disminuye (aumenta) en el índice de factores  $T_t^d$ .

También es posible especificar una forma funcional para los costos de transacciones a los que se enfrenta el intermediario.

$$(12) \quad \Omega^L = (B_t/P_t)^{\gamma_1} (T_t^s)^{\gamma_2} (POL_t)^{\gamma_3}$$

$$(13) \quad \Omega^{L^*} = (B_t^*/P_t^*)^{\delta_1} (T_t^s)^{-\delta_2} (POL_t)^{-\delta_3},$$

con:

$$\gamma_i > 0 \text{ para } i=1,3 \text{ y } \gamma_2 \geq 0$$

$$\delta_i > 0 \text{ para } i=1,2 \text{ y } \delta_3 \geq 0$$

Al tomar las derivadas de (12) y (13), conectando el resultado en (7) y reacomodando, llegamos a la expresión de la oferta de créditos en dólares:

$$(14) \quad \frac{B_t^*}{P_t^*} = \left\{ \left[ \frac{(T_t^s)^{\delta_2} (POL_t)^{\delta_3}}{\delta_1} \right] \left[ (1+R_t^*) - E_t \left( \frac{\rho_t}{\rho_{t+1}} \right) \left( 1+R_t - \delta_1 \left( \frac{b_t}{P_t} \right)^{\delta_1-1} (T_t^s)^{\delta_2} (POL_t) \right) \right] \right\}^{\frac{1}{\delta_1-1}}$$

La expresión (14) implica, dado  $b_t/P_t$ , que la oferta de créditos denominada en \$ está aumentando en la tasa de interés extranjera ( $R_t^*$ ) y disminuyendo en la tasa de interés doméstica ( $R_t$ ). También aumenta en la tasa esperada de depreciación de la moneda doméstica, la credibilidad de la política ( $POL_t$ ) y el índice de oferta de elementos externos  $T_t^s$ .

Ahora, es posible obtener una expresión estimable empíricamente si hacemos lineal a (14), imponiendo equilibrio en el mercado de créditos. Este equilibrio se obtiene al sustituir ( $b_t/P_t$ ) de la expresión (10) por ( $B_t/P_t$ ) en (14). Estas condiciones producen relaciones de la forma:<sup>17</sup>

$$\begin{aligned} (b_t^*/P_t^*) &= F \left[ \begin{array}{ccccccc} R_t^* & R_t & E_t(\rho_{t+1}/\rho_t) & T_t^d & T_t^s & C_t & POL_t \\ + & - & + & + & - & + & + \end{array} \right] \\ (B_t^*/P_t^*) &= G \left( \begin{array}{cccc} R_t^* & T_t^d & C_t^* & POL_t \\ - & + & + & + \end{array} \right) \end{aligned}$$

Para realizar una regresión, la oferta y demanda de créditos denominados en moneda extranjera se pueden resumir, respectivamente, en las siguientes expresiones lineales:

$$\begin{aligned} (B_t^*/P_t^*) &= a_{0B} - a_{1B}R_t^* + a_{2B}T_t^d + a_{3B}C_t^* + a_{aB}POL_t \\ (b_t^*/P_t^*) &= a_{0b} + a_{1b}R_t^* - a_{2b}R_t + a_{3b}E_t(\rho_{t+1}/\rho_t) + a_{4b}T_t^d + a_{5b}T_t^s - \\ &\quad - a_{6b}C_t + a_{7b}POL_t \end{aligned}$$

Finalmente, combinando estas dos ecuaciones de regresión últimas obtenemos la forma reducida siguiente:

$$(15) \frac{b_t^*}{P_t^*} = \frac{B_t^*}{P_t^*} = \mu_0 - \mu_1 R_t - \mu_2 E_t(\rho_{t+1}/\rho_t) + \mu_3 T_t^d + \mu_4 T_t^s + \mu_5 POL_t - \mu_6 C_t + \mu_7 C_t^*$$

Para las que:

$$\begin{aligned} \mu_0 &= \eta(a_{0b}a_{1B} + a_{1b}a_{0B}); \mu_1 = \eta(a_{2b}a_{1B}); \mu_2 = \eta(a_{3b}a_{1B}); \\ \mu_3 &= \eta(a_{1b}a_{2B} + a_{4b}a_{1B}); \mu_4 = \eta(a_{5b}a_{1B}); \mu_5 = \eta(a_{1b}a_{4B} + a_{7b}a_{1B}); \\ \mu_6 &= \eta(a_{6b}a_{1B}); \mu_7 = \eta(a_{1b}a_{3B}); \text{ y } \eta = (a_{1b} + a_{1B}). \end{aligned}$$

La expresión (15) ofrece la base para la estimación econométrica.

<sup>17</sup> Observe que los signos de derivadas parciales se muestran debajo de cada variable.

De acuerdo con (15), los créditos denominados en moneda extranjera aumentarán cuando la tasa de interés de la moneda doméstica disminuya, la tasa de depreciación esperada del dinero doméstico sea menor; aumenten los factores de red del dólar y cuando decaiga la credibilidad en la política económica.<sup>18</sup>

## V. ASPECTOS METODOLÓGICOS

A pesar de que (15) es la base de nuestro análisis econométrico, se incluirá una variable *dummy* para representar los cambios en la legislación que tuvieron lugar en 1996.<sup>19</sup> Entonces finalmente estimamos las ecuaciones como sigue:

$$Dol = C_0 + C_1R + C_2ExD + C_3T^d + C_4T^s + C_5POL + C_6Dum$$

Vamos a definir específicamente lo que se usa en cada caso:

- *Dol*: Nuestra variable dependiente estará definida como el coeficiente de créditos denominados en dólares respecto al total de créditos (en las estimaciones econométricas, el coeficiente se expresará como un porcentaje). Ambos componentes del coeficiente corresponden a créditos otorgados al sector privado.
- *R*: La variable utilizada fue *Rdom*, que corresponde a la tasa de interés promedio que se cobra por créditos denominados en moneda doméstica en el sistema bancario costarricense.
- *ExD*: Porcentaje esperado por variación en el tipo de cambio. Tal y como sucede en Kikut y Méndez (2003), asumimos expectativas adaptativas. La manera en que implementamos dicho supuesto es como sigue: La variación esperada en el período *t* se asumió como el promedio de variación de los últimos tres meses antes de *t* respecto al mismo período un año antes.<sup>20</sup> Los mismos autores aseguran que para el caso de

<sup>18</sup> No tomaremos en consideración efectos de consumo ya que no es posible obtener una variable *proxy* con periodicidad mensual como se utilizó con el resto de las variables.

<sup>19</sup> En los trabajos de Peiers y Wrase (1997) y Kikut y Méndez (2003), también se utilizó una variable *dummy* para representar dichos cambios en la regulación.

<sup>20</sup> Este enfoque difiere ligeramente del de Kikut y Méndez (2003), utilizan la variación anualizada de los últimos tres meses respecto al mismo período un año antes. Nuestro enfoque produce series muy similares pero con ligeramente

Costa Rica, varios estudios empíricos apoyan el supuesto de expectativas adaptativas con base en valores pasados de la variable.

- $T^d$  : Factores de red que afectan la demanda de créditos. Utilizamos alternativamente algunos de los coeficientes ya mencionados en la sección III, inciso 2. Se incluyeron  $RD$  o  $RDM2$  (en términos porcentuales) con un rezago. Tal y como se mencionó anteriormente, la idea es que cuantos más agentes utilicen dólares, menor será el costo de cambiar de moneda doméstica a dólares. Además, incluimos estas variables como tipo *ratchet* ( $RchtRDM2$ ,  $RchtRD$ ) para probar la presencia de histéresis que podría surgir porque una vez que los agentes incurren en los costos de cambio de moneda, el incentivo para desdolarizar será muy débil.
- $T^s$  : Factores de red afectando la oferta de crédito. Llamamos  $SpdC$  al otorgamiento de créditos promedio sobre la tasa de solicitud de créditos denominados en moneda local.<sup>21</sup> Se asume que  $T^s$  afecta positivamente la proporción de préstamos totales denominados en moneda extranjera. Asumimos implícitamente que los prestatarios hallarán más fácil otorgar créditos en moneda extranjera (y no en moneda doméstica) cuando este margen sea amplio.
- $POL$  : Incluimos varios *proxies* para representar variaciones y persistencia (memoria a largo plazo) en la credibilidad.  $DInf$  corresponde a la diferencia entre las tasas de inflación entre Costa Rica y los Estados Unidos.  $Volinf$  es la volatilidad realizada en el diferencial de las tasas de inflación.<sup>22</sup> También intentamos aislar los efectos de histéresis asociados con cambios en la credibilidad de aquellos asociados con externalidades de red incluyendo variables *proxy* para las políticas de tipo *ratchet*:  $RatchDInf$  en el período  $t$  corresponde al valor máximo de  $DInf$  hasta  $t-1$ ;  $RatchVolinf$  es el equivalente pero para el valor máximo de  $Volinf$ . Observe que los incrementos en

---

menos variabilidad especialmente durante períodos de alta volatilidad.

<sup>21</sup> La tasa para créditos es la misma que la definida para  $R$ . Por otro lado, como indicador de la tasa de otorgamiento de créditos tomamos la tasa anualizada de depósitos a seis meses denominados en CRC.

<sup>22</sup> Como volatilidad mensual de una variable  $A$  en el período  $t$  tomamos:

$$Vol(A_t) = \frac{A_t - A_{t-1}}{Av(A)}$$

donde  $Av(A)$  es el valor promedio de la variable  $A$ .  $Volinf$  no se incluyó directamente en las estimaciones, sólo se utilizó la representación de tipo *ratchet*.

dichas *proxies* se asociarán con menor credibilidad de modo que el signo esperado del coeficiente sea positivo, es decir, menor credibilidad debe asociarse con mayores niveles de dolarización.

- *Dum*: Sólo se incluyó una variable *dummy* (*Dum96*). Está compuesta por ceros hasta agosto de 1996 y números uno después.<sup>23</sup> A diferencia de estudios en que las variables *dummies* se incluyeron para representar efectos de Pascua o Navidad (correcciones por temporadas),<sup>24</sup> dichas correcciones no fueron incluidas debido a la naturaleza de nuestra variable dependiente que es el coeficiente de préstamos que no se supone se vea afectada por cambios estacionales.

A pesar de que en la sección III, inciso 2, mostramos estadísticas descriptivas para un período de 18 años, el tamaño final de la muestra utilizada para efectos econométricos tuvo 149 observaciones (enero 1993 a mayo 2005). La razón es que algunas de las variables requeridas para realizar las regresiones no estaban disponibles en los sitios oficiales para una muestra tan grande.<sup>25</sup> Además, evitamos utilizar datos de 1992 porque, como ya mencionamos, se realizaron varios cambios a la legislación ese año. Dichos cambios afectaron seguramente la tendencia del proceso de dolarización, algo que no sería posible probar sin datos previos a 1992.

Se utilizó la técnica Mínimos Cuadrados Ordinarios (MCO) para realizar las regresiones y el software fue *Econometric Views* (E-Views 4.0).

Cuando se esperaba la multicolinealidad obvia de variables explicativas, algunas de ellas se excluyeron. Ya que también realizamos cálculos utilizando *RD* o *RDM2* de manera alternativa como *proxies* para el costo de red que afecta la demanda de préstamos, finalmente obtuvimos diez especificaciones diferentes que se dividieron en dos cuadros con cinco regresiones cada uno. Esto se muestra en los cuadros 4 y 5 en la sección de resultados. Por ejemplo, se espera que *Dinf* y *Rchinf* estén correlacionadas, por lo que no se incluyeron en la misma regresión.

Antes de continuar con los resultados, es importante señalar

<sup>23</sup> Introdujimos el cambio en las variables *dummies* dos meses después de los cambios en la legislación de modo que las modificaciones tuvieran tiempo de hacer efecto.

<sup>24</sup> Por ejemplo, Kikut y Méndez (2003).

<sup>25</sup> Específicamente, no fue posible obtener datos de tasas de interés activas denominadas en dólares antes de 1992.

algunos problemas que hacen de este trabajo algo distinto a lo que se ha realizado en otros análisis empíricos del proceso de dolarización en Costa Rica u otros países.

En primer lugar, a pesar de que utilizamos el mismo modelo teórico propuesto por Peiers y Wrase (1997), nuestro enfoque econométrico es diferente; mientras que utilizamos MCO para explicar el comportamiento del coeficiente de créditos denominados

**CUADRO 4.** SALIDAS DE LA REGRESIÓN USANDO RD COMO PROXY DE Td<sup>a</sup>

Regresiones	Variable dependiente: Dol				
	1.1	1.2	1.3	1.4	1.5
Constante	-33.51 (-6.096)	-32.552 (-6.229)	-12.368 (-2.065)	-36.185 (-8.186)	-17.006 (-3.487)
Rdom	-1.086 (-5.033)	-1.337 (-6.846)	-1.636 (-8.399)	-0.91 (-5.323)	-1.228 (-7.438)
ExD	0.065 (0.405)	-0.287 (-1.752)	-0.686 (-3.869)	-0.419 (-3.085)	-0.805 (-5.697)
SpdC	3.657 (10.07)	3.197 (12.447)	3.951 (813.609)	2.939 (10.448)	3.039 (12.112)
RD (-1)	0.567 (18.149)	0.51 (16.284)	0.47 (15.45)		
RchtRD				0.581 (20.833)	0.54 (20.786)
Dinf	0.067 (0.391)				
RchtDinf		0.716 (3.94)		0.647 (4.263)	
RchtVolinf			25.293 (6.129)		23.994 (7.254)
Dum96	5.572 (3.306)	1.104 (0.587)	-2.336 (-1.223)	0.687 (0.435)	-2.848 (-1.832)
R <sup>2</sup> ajustado	0.928	0.932	0.941	0.952	0.61
Estadístico F	1,370.1	1,520.7	1,736.3	2,159.6	2,628.5
Probabilidad (estadístico F)	0.000	0.000	0.000	0.000	0.000
Error estándar de la regresión	4.636	4.404	4.125	3.703	3.359
Criterio de informa- ción Akaike	5.952	5.849	5.718	5.502	5.307
Criterio Schwarz	6.093	5.99	5.859	5.643	5.448
Estadístico Durbin- Watson	0.27	0.292	0.324	0.205	0.275

NOTA: Estadístico t en paréntesis.

<sup>a</sup> No hay significancia estadística al 5%.

en dólares respecto al número total de créditos, utilizaron un modelo Probit para determinar la probabilidad de que un nuevo crédito sea denominado en dólares dadas algunas variables explicativas. Su técnica requiere una base de datos (panel) mucho más rica con información específica para cada uno de los créditos individuales.

Con respecto a Kikut y Méndez (2003), es preciso mencionar varias diferencias. En primer lugar, el enfoque econométrico parece estar basado razonablemente en la teoría, pero no formalizaron un modelo específico para proporcionar bases microeconómicas. En segundo lugar, como *proxies* de dolarización utilizaron coeficientes de depósitos bancarios respecto a diferentes agregados de liquidez, esas son sus variables dependientes. En este estudio el foco es en el comportamiento del crédito como un *proxy* de una etapa futura de la dolarización. Otra diferencia es el período de análisis: mientras que ellos utilizaron datos mensuales de 1990 a 2002, aquí la muestra empieza 2 años más tarde pero contiene 29 observaciones más.

También es importante tener en cuenta las limitaciones principales que podrían tener los resultados de los siguientes análisis. En primer lugar, ya se mencionó que la dolarización es inherentemente difícil de medir. Requerimos basarnos en *proxies* como los explicados anteriormente, de modo que el ámbito real en el que se usan los dólares en la economía básicamente se desconoce y la precisión de los resultados depende de la manera en que los *proxies* representan la variable subyacente.

Otra limitación es la naturaleza de los datos que utilicé, que no es muy rica. Se utilizaron variables como las tasas de interés promedio y la cantidad acumulada de créditos. Es posible que los resultados generales (signos de los coeficientes por ejemplo) no cambien si en su lugar se utiliza un panel con información individual específica para cada crédito, pero en ese caso seguramente se obtendrían resultados más precisos.

Finalmente, a pesar de que algunas de las variables se utilizaron con rezago para incluirse en las regresiones, la selección de cuántos rezagos incluir se realizó *ad hoc*. Ya que no es objetivo de este trabajo producir resultados adecuados para la proyección sino obtener dirección sobre los efectos, y dado que la mayoría de los signos fueron como se esperaba, no se realizó mayor análisis para determinar la combinación óptima de rezagos de las variables independientes.<sup>26</sup>

<sup>26</sup> Kikut y Méndez (2003) ofrecen un enfoque útil para determinar la combinación óptima de rezagos, incluso ofrecen un anexo con un programa



## VI. RESULTADOS

Los resultados obtenidos de las salidas de regresión se resumen más adelante en los cuadros 4 y 5. Cada cuadro contiene información acerca de cinco regresiones. La diferencia entre los

CUADRO 5. SALIDAS DE LA REGRESIÓN USANDO RDM2 COMO PROXY DE Td

Regresiones	Variable dependiente: Dol				
	2.1	2.2	2.3	2.4	2.5
Constante	-54.313 (-9.871)	-52.234 (-9.581)	32.251 (-5.081)	-59.941 (-11.887)	-39.97 (-6.977)
Rdom	-1.012 (-5.012)	-1.113 (-5.782)	-1.448 (-7.642)	-0.734 (-4.067)	-1.088 (-6.248)
ExD	-0.262 <sup>a</sup> (-1.773)	-0.42 <sup>a</sup> (-2.699)	-0.832 (-5.03)	-0.577 (-4.167)	-0.972 (-6.769)
SpdC	3.535 (10.402)	3.589 (11.617)	3.698 (13.186)	2.865 (9.823)	3.201 (11.650)
RD (-1)	1.888 (19.983)	1.775 (17.269)	1.619 (16.704)		
RchtRD				1.946 (20.062)	1.788 (19.98)
Dinf	0.207 (1.315)				
RchtDinf		0.508 (2.818)		0.496 (3.098)	
RchtVolinf			22.78 (5.731)		22.351 (6.459)
Dum96	1.928 (1.116)	-1.610 <sup>a</sup> (-0.882)	-5.246 <sup>a</sup> (-2.888)	1.474 <sup>a</sup> (0.37)	-5.158 <sup>a</sup> (-3.222)
R <sup>2</sup> ajustado	0.935	0.938	0.949	0.949	0.958
Estadístico F	1,576.6	1,645.7	1,922.3	2,040.3	2,476.8
Probabilidad (estadístico F)	0.000	0.000	0.000	0.000	0.000
Error estándar de la regresión	4.326	4.326	3.923	3.809	3.46
Criterio de infor- mación Akaike	5.813	5.771	5.617	5.558	5366
Criterio Schwarz	5.954	5.912	5.758	5.699	5.507
Estadístico Durbin- Watson	0.331	0.333	0.383	0.2	2.76

NOTA: Estadístico t en paréntesis.

<sup>a</sup> No hay significancia estadística al 5%.

en Eviews para realizar dicha tarea.

cuadros es la variable que se utilizó como *proxy* del costo de red afectando la demanda de créditos denominados en dólares ( $T^d$ ); el cuadro 4 utiliza *RD*, mientras que el cuadro 5 utiliza *RDM2*. Las diferencias entre regresiones dentro de cada cuadro corresponden a la combinación de variables remanentes que se incluyeron en cada caso.

Sólo cuatro variables son comunes a todas las regresiones: *Rdom*, *ExD*, *SprC* y *Dum96*. Observe que la primera columna en cada cuadro (regresión 1.1 y 2.1) muestra el resultado de una especificación que no incluye ningún *ratchet* y las otras cuatro columnas contienen al menos uno. El objetivo de esta distinción inicial es analizar si obtenemos un mejor ajuste cuando se incluyen los efectos de las *ratchet*, es decir, si incluir las variables que representan la histéresis mejora los resultados.

Por otra parte, para evitar la multicolinealidad, ninguna variable se utilizó en la misma ecuación con su representación *ratchet*, por eso cada cuadro de la columna segunda y tercera es diferente de la cuarta y quinta. Finalmente, también se evitó utilizar la representación *ratchet* de una variable junto con el *ratchet* de su volatilidad, que es la diferencia por ejemplo, entre la segunda y tercera columna.

Dicha división de los resultados nos permite determinar si los efectos de histéresis producidos por la diferencia de inflación y el costo de red son significativos a nivel estadístico cuando se incluyen en la misma ecuación (ver, por ejemplo, las regresiones 1.4, 1.5, 2.4 y 2.5), que serán indicador de que la persistencia en el proceso de dolarización es causada no sólo por la falta de confianza (memoria a largo plazo de episodios pasados de altos niveles de inflación)<sup>27</sup> sino también por un proceso en el que una vez que los agentes incurren en el costo inicial del cambio de moneda, no tendrán incentivo alguno para desdolarizar.<sup>28</sup>

Vamos a resumir los problemas más notables que surgieron de los resultados de la regresión. En primer lugar, lo que expusimos en la sección V, casi todas las direcciones de los efectos fueron como se esperaba. Sólo *Dum96* mostró signos inesperados en algunas de las regresiones, pero en la mayoría de los casos no fue significativo a nivel estadístico. *ExD* muestra un signo

<sup>27</sup> Tandon y Wang (2003) explican en su modelo la manera en que la falta de confianza puede generar histéresis. Del mismo modo, el modelo presentado por Lebre de Freitas (2004) genera la posibilidad de que un aumento temporal en el índice inflacionario pueda producir un efecto permanente en el uso de la moneda extranjera.

<sup>28</sup> El papel que tienen los costos por cambio de moneda sobre la histéresis se desarrolla en el documento escrito por Guidotti y Rodríguez (1992).

inesperado en la regresión 1.1, pero su valor no es significativo tampoco.

También es importante observar que *Dum96* es la variable que se encontró como no significativa de manera más consistente, lo que podría ser una indicación de que los cambios en la legislación realizados en 1996 no afectaron realmente el ya continuo proceso de dolarización de créditos en Costa Rica. Este resultado contrasta con lo que Kikut y Méndez (2003) encontraron; también incluyeron una variable *dummy* en 1996 como variable explicativa y la encontraron significativa de manera sistemática. Ya que esto podría parecer un resultado contradictorio, es necesario dar una explicación. Este estudio se enfoca en un indicador de una etapa más tardía en un proceso de dolarización, de modo que nuestra variable dependiente es un *proxy* para la sustitución monetaria, mientras que en Kikut y Méndez (2003) el interés es más general, utilizaron variables dependientes como coeficientes de depósitos a plazo respecto a la liquidez total (dolarización en etapa temprana). Entonces no es de sorprender que las enmiendas a la ley presentadas en 1996 tengan un impacto diferente dependiendo de los *proxies* de dolarización que se utilicen.

Observando los indicadores de bondad de ajuste, los resultados no se ven mal. No sólo la mayoría de las variables fueron significativas de manera individual, sino que también la prueba *F* fue incapaz de rechazar la hipótesis nula de ausencia de significancia conjunta. Además las *R* cuadradas fueron consistentemente altas.

Es interesante resaltar que los *proxies* para las externalidades de red (*RD* y *RDM2* para  $T^d$  y *SpdC* para  $T^s$ ) tienen el signo positivo esperado y fueron estadísticamente significativos. Esos resultados fueron robustos con la inclusión de los *ratchets* de diferencias en la inflación.

Por otra parte, los *proxies* de credibilidad en las políticas (*Dinf*, *RchtDinf*, *RchtVolinf*) tienen también el signo esperado y, excluyendo el caso de *Dinf* en las regresiones 1.1 y 2.1, fueron consistentemente significativos. Entonces existe evidencia de que la credibilidad en la política desempeña un papel en el proceso, tal y como predice cierta literatura.<sup>29</sup>

Analizando cuáles eran las combinaciones de resultados cuando se incluyeron las variables tipo *ratchet*, alcanzamos más hallazgos interesantes sobre la histéresis. En primer lugar, observe que al comparar con el resto de las respectivas regresiones

<sup>29</sup> Ver, por ejemplo, a Tandon y Wang (2003).

en su cuadro, las regresiones 1.1 y 2.1 (que corresponden a regresiones incluyendo variables *ratchet*) tienen el mayor número de variables no significativas, también muestran los valores estadísticos F más pequeños y el mayor criterio Akaike. Por lo tanto, el ajuste de estas dos regresiones es comparativamente peor que las otras cuatro. Entonces, la evidencia muestra que incluir variables tipo *ratchet* en la regresión realmente mejora los resultados, es decir, hay elementos con persistencia que afectan el proceso de dolarización.

Esta última conclusión se ve reafirmada cuando observamos las cuatro columnas de la derecha de cada cuadro. En todos los casos, cada variable *ratchet* incluida era significativa estadísticamente. Sin importar que sea una variable *ratchet* en la diferencia de inflación o en la variable *proxy* para los costos de red o si están incluidos juntos, las variables *ratchet* fueron siempre representativas.

Existe, en efecto, un hallazgo inesperado y notorio: cuando no se consideró en forma de *ratchet*, la diferencia en la inflación (*Dinf*), a pesar de tener un coeficiente con el signo positivo esperado, resultó ser estadísticamente no significativa. De hecho, en regresiones cuyos resultados no se muestran aquí, fue necesario rezagar *Dinf* al menos siete períodos para lograr un coeficiente significativo.

Estos últimos resultados podrían indicar no sólo que los agentes requieren de varios meses para incorporar cambios inflacionarios en sus decisiones acerca de los préstamos, sino también que el impacto que esos cambios muestran persistencia. Los agentes no olvidan fácilmente períodos de alta inflación y la falta de confianza que producen permanece durante mucho tiempo.

Al analizar los resultados 1.4, 1.5, 2.4 y 2.5 surge un asunto interesante. En cada una de esas regresiones se incluyeron dos efectos *ratchet*, uno para capturar la persistencia que la diferencia de inflación o su alta volatilidad tienen en la dolarización (*RchtDinf* y *RchtVolinf* respectivamente), y la otra tomar en cuenta los efectos que los costos de red tienen sobre el proceso (*RchtRM2* o *RchtRD*). Los resultados muestran que, al incluirse juntos, las variables de *ratchet* son significativas tanto de forma individual como conjunta.

Parece que, la histéresis que muestra el proceso de dolarización tiene al menos dos fuentes que pueden aislarse. La falta de confianza en la política que permanece durante períodos largos después de una diferencia de alta inflación (o un episodio de alta volatilidad) y la existencia de los costos iniciales del cambio de

monedas realizado en alguna ocasión, se tornan una barrera para la desdolarización. Esto constituye un verdadero problema para los formuladores de política que quieren revertir el proceso de dolarización. No sólo tienen el problema de la memoria de largo plazo de los agentes sino también deben hacer de la moneda local algo tan atractivo que los agentes consideren óptimo incurrir una vez más en el costo por el cambio de monedas. No hay duda de que es una tarea muy difícil, especialmente en una economía como la de Costa Rica, en donde déficit públicos altos y persistentes mantienen las tasas de inflación en dos dígitos desde hace mucho tiempo.

Como actualización a la versión original de este documento y antes de proceder con la sección de conclusiones, es importante mencionar que en octubre de 2006 el Banco Central de Costa Rica abandonó un viejo sistema de dos décadas de existencia de minidevaluaciones y cambió a un sistema de bandas cambiarias controladas. Además de reducir la inflación, inducir a los agentes a desdolarizar fue uno de los principales objetivos del cambio. El anexo muestra que la mayoría de nuestras conclusiones principales no se ven afectadas si actualizamos las series hasta julio de 2007 y continuamos con los mismos procedimientos econométricos.

## VII. CONCLUSIONES

Teniendo en mente nuestros objetivos iniciales y el análisis posterior de la evidencia empírica que se mostró en las secciones precedentes, los siguientes comentarios resumen las conclusiones de este estudio.

- La política abierta para la regulación en el uso de divisas ha hecho de la dolarización un elemento estándar en Costa Rica. Al comienzo del muestreo, utilicé para fines descriptivos (1987) la sustitución de CRC por dólares a medida que el depósito de valor se convirtió en un fenómeno constante. En 1987, la mayoría de los indicadores de dolarización colocaron a Costa Rica entre las economías altamente dolarizadas.
- Los indicadores de sustitución monetaria comenzaron a crecer rápidamente entre 1999 y 2000, sugiriendo esos años como el inicio de una nueva etapa en el proceso de dolarización; la gente comenzó a utilizar dólares no sólo como depósito de valor sino también como unidad de cuenta e incluso como medio de

intercambio. Este esquema de evolución en el proceso concuerda con el descrito en trabajos teóricos<sup>30</sup> y la experiencia de otros países.<sup>31</sup>

- A pesar de que mucha de la teoría otorga a los altos niveles de inflación un papel central al explicar la evolución del proceso, nuestra evidencia muestra que dicha alta inflación podría ser un punto inicial, pero existen otros elementos que explican la evolución posterior del proceso. Esto inicialmente es sugerido por el hecho de que ambas series (*proxy* de dolarización y la diferencia en inflación) tienen una raíz unitaria y no están cointegrados.

- El hecho de que nuestra serie de dolarización tiene una raíz unitaria tiene una fuerte implicación porque significa que los cambios que afectan la variable tienen efectos permanentes y algunos de esos cambios podrían producir histéresis en su evolución.

- El análisis econométrico confirma lo que sugiere el hecho de que las series no están cointegradas. Algunas otras variables son significativamente altas al explicar la evolución de nuestra variable *proxy* de dolarización. La tasa de interés doméstica, la tasa de depreciación esperada de la moneda doméstica, las *proxies* de credibilidad de la política y factores de creación de redes tuvieron el signo de coeficiente esperado y resultaron significativas estadísticamente al explicar el comportamiento de la variable dependiente.

- Todas las variables de tipo *ratchet* incluidas en las regresiones tienen el signo esperado y fueron significativas estadísticamente. Si añadimos el hecho de que nuestra variable *proxy* de dolarización tiene raíz unitaria, tenemos una fuerte evidencia de histéresis. Por lo tanto, parece que el proceso es irreversible de alguna manera, tal y como ha sucedido en otros países latinoamericanos, con la diferencia de que los indicadores de dolarización no han crecido de manera explosiva en el caso de Costa Rica.

- Es importante resaltar que las variables *ratchet* de la diferencia de inflación y en factores de creación de redes son significativas,

<sup>30</sup> Ver a Calvo y Végh (1992).

<sup>31</sup> Ver a Guidotti y Rodríguez (1992).

no sólo cuando se incluyen por separado sino también cuando se incluyen en forma conjunta. Este hecho implica que hay al menos dos efectos que generan histéresis que pueden aislarse. Por un lado tenemos que los agentes recuerdan episodios de alta inflación durante mucho tiempo y entonces la falta de confianza en la política económica del país se mantiene por largo tiempo, haciendo que la moneda extranjera sea más atractiva aún cuando la inflación está bajo control. Por otro lado, los costos del cambio inicial de moneda también mantiene a los agentes utilizando constantemente la moneda extranjera. El colón no se ha vuelto lo suficientemente atractivo para que los agentes cambien de moneda.

- El hecho de que ambos factores (tanto la falta de credibilidad en la política económica como los efectos de creación de redes) producen histéresis en el proceso de dolarización tiene una implicación clave; es casi imposible revertir el proceso al menos a mediano plazo utilizando la política económica estándar. No sólo es necesario que la moneda doméstica se vuelva más atractiva comparada con el dólar estadounidense, sino también que su atractivo sea duradero. Además, el incentivo debe ser lo suficientemente fuerte para superar el costo de cambio hacia la moneda doméstica. Esto requerirá un período largo en el que la tasa de inflación local se mantenga lo suficientemente baja para hacer menos atractivo el uso de dólares, lo que en una economía como la de Costa Rica, con un problema fundamental de déficit público, parece ser un objetivo inalcanzable.

- Las autoridades en la economía costarricense han intentado luchar contra la dolarización manteniendo tasas de interés locales a niveles que harían más atractivo el uso de CRC, pero esta política no ha dado tan buen resultado. Toda medida a la dolarización continúa creciendo. Es posible que sólo hayan restado velocidad al proceso. El problema fundamental que es el gran déficit público que causa altos niveles de inflación y deteriora la confianza de los agentes en la moneda doméstica, se mantiene alimentando el incentivo para dolarizar. Si tomamos en consideración que esta política de la tasa de interés podría tener un alto costo para el banco central y entonces agravar el problema fiscal, las autoridades podrían participar en un esquema del que sería muy difícil salir.

- Dada la dificultad de luchar en contra del proceso de dolarización y tomando en consideración la experiencia de otros

países donde era posible desdolarizar sólo por medio de esquemas de confiscación, nuestra recomendación va alineada con lo que recomiendan los expertos del FMI: la dolarización es un proceso con el que los países deben aprender a vivir en lugar de luchar contra él. Es fundamental evitar nuevos episodios de alta inflación y por supuesto, se debe resolver el problema fiscal o al menos, controlarse.

- Las modificaciones recientemente implantadas al régimen cambiario (octubre de 2006) no alteran las principales conclusiones de este documento, aunque la evidencia reciente sugiere que el nuevo sistema de bandas cambiarias controladas (y la intención explícita del banco central de migrar hacia un sistema más flexible) ha tenido un impacto en las *proxies* de dolarización. No diremos que se está induciendo un proceso de desdolarización pero al menos el cambio ha tenido un efecto estadístico importante. La dolarización, al menos, ha dejado de crecer. Dado que encontramos agentes que tienen buena memoria y que hay costos por el cambio de monedas, se requiere un período de tiempo más largo para saber si es posible inducir a los agentes a la desdolarización.

Para terminar, señalamos dos líneas de investigación que podrían complementar y contribuir a los logros de este estudio. En primer lugar, ya que no fue posible encontrar alguna de la información necesaria para extender la muestra para estimaciones econométricas más allá de 1993, y dado que un corto análisis gráfico (ver la gráfica II) sugiere que los cambios en la regulación de 1992 tuvieron un efecto en la evolución de nuestra variable dependiente, sería importante hacer un esfuerzo adicional para ampliar la muestra para analizar la significancia estadística de dichos cambios.

En segundo lugar, sería interesante tener una idea de qué tan cara ha resultado la política que fomenta el uso de los colones manteniendo los tipos de interés en ciertos niveles y qué tanto estos costos han influido en los índices domésticos de inflación.

Finalmente, y más en el campo financiero de la investigación, resultaría de utilidad contar con medidas respecto a qué tan vulnerable se ha vuelto el sistema bancario ante modificaciones al tipo de cambio, dada la creciente participación de créditos denominados en dólares.



*Apéndice* (resultados actualizados)

Los cuadros A.1 y A.2 son actualizaciones de los cuadros 4 y 5. Básicamente ampliamos todas las series hasta julio de 2007 y

**CUADRO A. 1.** SALIDAS DE LA REGRESIÓN USANDO RD COMO PROXY DE Td

<i>Regresiones</i>	<i>Variable dependiente: Dol</i>				
	<i>1.1</i>	<i>1.2</i>	<i>1.3</i>	<i>1.4</i>	<i>1.5</i>
Constante	-28.461 (-5.580)	-28.381 (-5.802)	-11.217 (-2.090)	-37.377 (-9.386)	-20.869 (-4.865)
Rdom	-1.003 (-4.964)	-1.233 (-6.709)	-1.485 (-8.222)	-0.913 (-6.129)	-1.152 (-7.990)
ExD	0.011 (0.070)	-0.351 (-2.218)	-0.722 (-4.305)	-0.400 (-3.197)	-0.723 (-5.581)
SpdC	3.087 (10.015)	3.323 (12.296)	3.396 (13.455)	3.075 (14.237)	3.142 (15.868)
RD (-1)	0.603 (23.269)	0.556 (21.327)	0.518 (20.205)		
RchtRD				0.264 (28.570)	0.532 (27.884)
Dinf	-0.013 <sup>a</sup> (-0.082)				
RchtDinf		0.655 (3.8812)		0.658 (4.878)	
RchtVolinf			23.028 (6.219)		21.669 (7.582)
Dum96	4.210 <sup>a</sup> (2.619)	0.141 (0.078)	-2.936 (-1.657)	1.302 <sup>a</sup> (0.915)	-1.446 <sup>a</sup> (-1.044)
Dum06	6.608 (3.459)	4.054 (2.095)	0.591 (0.301)	-5.314 (-3.502)	-7.877 (-5.343)
R <sup>2</sup> ajustado	0.930	0.936	0.943	0.959	0.965
Estadístico F	1,370,1	1,842.22	2,089.95	2,924.92	3,445.07
Probabilidad (estadístico F)	0.000	0.000	0.000	0.000	0.000
Error estándar de la regresión	4.548	4.362	4.098	3.469	3.198
Criterio de informa- ción Akaike	5.912	5.828	5.704	5.370	5.208
Criterio Schwarz	6.057	5.973	5.848	5.515	5.352
Estadístico Durbin- Watson	0.299	0.303	0.322	0.232	0.303

NOTA: Estadístico t en paréntesis.

<sup>a</sup> No hay significancia estadística al 5%.

**CUADRO A. 2.** SALIDA DE LA REGRESIÓN USANDO RDM2 COMO PROXY DE Td

Regresiones	Variable dependiente: Dol				
	2.1	2.2	2.3	2.4	2.5
Constante	-51.168 (-10.211)	-50.025 (-10.108)	-33.807 (-6.092)	-59.415 (-13.394)	-43.390 (-8.838)
Rdom	-0.910 (-4.957)	0.982 (-5.652)	-1.252 (-7.360)	-0.690 (-4.480)	-0.953 (-6.368)
ExD	-0.301 (-2.163)	-0.448 (-3.071)	-0.818 (-5.354)	-0.531 (-4.224)	-0.862 (-6.638)
SpdC	3.052 (10.936)	12.171 (1.857)	13.612 (1.723)	2.793 (12.660)	2.908 (14.359)
RDM2 (-1)	1.936 (26.405)	1.857 (23.561)	1.723 (22.534)		
RchtRDM2				1.913 (28.138)	1.793 (27.404)
Dinf	0.164 <sup>a</sup> (1.124)				
RchtDinf		0.435 (2.667)		0.466 (3.333)	
RchtVolinf			19.715 (5.681)		10.097 (6.471)
Dum96	0.532 <sup>a</sup> (0.349)	-2.544 (-1.522)	-5.691 (-3.481)	-1.432 <sup>a</sup> (-0.991)	-4.346 <sup>a</sup> (-3.102)
Dum06	3.661 (2.121)	2.323 (1.302)	-0.942 (-0.526)	-3.945 (-2.570)	-6.567 (-4.381)
R <sup>2</sup> ajustado	0.943	0.945	0.952	0.958	0.964
Estadístico F	2,069.90	2,142.57	2,455.25	2,851.47	3,347.59
Probabilidad (estadístico F)	0.00	0.00	0.00	0.00	0.00
Error estándar de la regresión	4.118	4.048	3.784	3.513	3.244
Criterio de informa- ción Akaike	5.713	5.679	5.544	5.396	5.236
Criterio Schwarz	5.858	5.824	5.689	5.540	5.381
Estadístico Durbin- Watson	0.351	0.349	0.385	0.210	0.280

NOTA: Estadístico t en paréntesis.

<sup>a</sup> No hay significancia estadística al 5%.

seguimos el mismo procedimiento econométrico. Además, para tomar en cuenta la modificación del régimen del tipo de cambio que tuvo lugar en octubre de 2006 se incluyó también una variable *dummy*. A pesar de que se capturaron solamente 10 meses

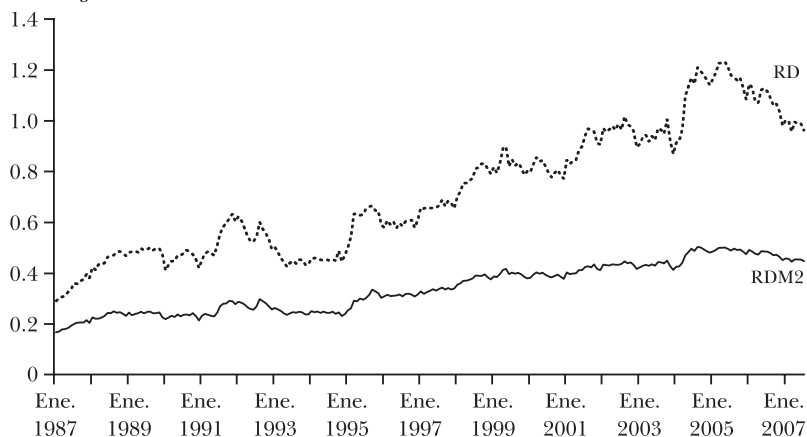
después de la modificación del régimen del tipo de cambio, que es ciertamente un período corto para evaluar sus implicaciones en el proceso de dolarización, el ejercicio de actualización resulta útil para fortalecer nuestras conclusiones originales.

Lo que es más importante extraer de la comparación entre los resultados econométricos originales con los actualizados es que las conclusiones principales no resultan afectadas. Aún tenemos evidencia fuerte de histéresis en el proceso de dolarización y todavía es posible identificar dos efectos que pueden aislarse: Falta de credibilidad y factores de red.

Como regla general, los resultados actualizados parecen superar en rendimiento a los originales respecto a su significancia, ahora obtenemos mayores  $R$  cuadradas y estadísticas  $F$ . Además se encontró que variables menos independientes no eran significativas a un nivel del 5%.

La variable *dummy* incluida para tomar en cuenta la modificación del régimen de tipo de cambio se encontró significativa de manera consistente al explicar nuestra *proxy* de dolarización. Después de sólo 10 meses de que comenzara a operar el sistema de bandas cambiarias controladas, la evidencia sugiere que ha habido un impacto sobre la variable dependiente. No podemos decir que se está induciendo un proceso de desdolarización pero al menos el cambio ha tenido un efecto importante a nivel estadístico. Al observar la gráfica A.I es posible ver que las *proxies* de dolarización al menos han dejado de crecer. Dado que encontramos agentes que tienen buena memoria a largo plazo y que hay costos por el cambio de monedas, se requiere de un

**GRÁFICA A.I.** COSTA RICA: COEFICIENTES DE DOLARIZACIÓN, ENERO DE 1987 A JULIO DE 2007



período de tiempo más largo para saber si es posible inducir a los agentes a la desdolarización.

## REFERENCIAS

- Baliño, T., A. Bennette y E. Borensztein (1999), *Monetary Policy in Dollarized Economies*, FMI (IMF Occasional Paper, nº 171).
- Calvo, A. G., y C. A. Végh (1992), "Currency Substitution in Developing Countries: An Introduction", *Revista de Análisis Económico*, vol. 7, nº 1, pp. 3-27.
- Dornbusch, R., F. Sturzenegger y H. Wolf (1990), "Extreme Inflation: Dynamics y Stabilization", *Brooking Papers on Economic Activity*, nº 2, pp. 2-84
- Fenske, K., y M. Melvin (1992), "Dollarization y Monetary Reform: Evidence from the Cochabamba Region of Bolivia", *Revista de Análisis Económico*, vol. 7, nº 1, pp. 125-38.
- Gomis-Poqueras, P., C. Serrano y A. Somuano (2000), *Currency Substitution in Latin America: Lessons from the 1990's*; véase en: <http://econ.worldbank.org/docs/1094.pdf>, agosto de 2005.
- Guidotti, P. E., y C. A. Rodríguez (1992), "Dollarization in Latin America: Gresham's Law in Reverse?", *IMF Staff Papers*, vol. 39, nº 3, pp. 518-44.
- Johansen, S., y K. Juselius (1990), "Maximum Likelihood Estimation y Inferences on Cointegration with Applications to the Demand for Money", *Oxford Bulletin of Economics y Statistics*, vol. 52, pp. 169-210.
- Kikut, A. C., y E. Méndez (2003), *La Dolarización en Costa Rica*, Banco Central de Costa Rica (DIE/03-2003-DI).
- Lebre de Freitas, M. (2004), "The Dynamics of Inflation y Currency Substitution in a Small Open Economy", *Journal of International Money y Finance*, vol. 23, pp. 133-42.
- Osterwald-Lenum, M. (1992), "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Prueba Statistics", *Oxford Bulletin of Economics y Statistics*, vol. 54, pp. 461-72.
- Peiers, B., y J. M. Wrase (1997), *Dollarization Hysteresis y Network Externalities: Theory y Evidence from an Informal Bolivian Credit Market*, Federal Reserve Bank of Philadelphia (Working Paper, nº 97-21).
- Rostowsky, J. (1992), "The Benefits of Currency Substitution During High Inflation y Stabilization", *Revista de Análisis Económico*, vol. 7, nº 1, pp. 91-108.

- Tandon, A., e Y. Wang (2003), "Confidence in Domestic Money y Currency Substitution", *Economic Inquiry*, vol. 41, n<sup>o</sup> 3, pp. 407-19.
- Uribe, M. (1997), "Hysteresis in a Simple Model of Currency Substitution", *Journal of Monetary Economics*, vol. 40, pp. 185-202.



*Alberto Humala*

# South American disinflation and regime switches: unobserved volatility components?

## I. INTRODUCTION

Inflation in South America has decreased substantially in recent times, both in levels and in volatility. Castillo, Humala and Tuesta (2007) found for Peru that permanent and transitory shocks to inflation were subject to switches in regime and that inflation dynamics had shifted towards long-term and short-term price stability. Remarkably, they also found that under recent inflation-intolerant policies, both permanent and transitory shocks were less volatile. From these results, an open question stands out. Do the decrease in short-term and long-term uncertainty result from domestic monetary policy being reoriented to a unique objective of price stability (in many cases in the form of inflation targeting)? Or is it the outcome of some other factors

*Paper prepared by A. Humala of Central Reserve Bank of Peru. The author thanks participants in the Economic Research Workshop at the Central Reserve Bank of Peru, for useful comments and suggestions. I also thank Camila Cisneros and Rosario Sánchez for invaluable research assistance. Any remaining errors are the authors' own responsibility. The views expressed herein are those of the author and do not necessarily reflect those of the Central Reserve Bank of Peru. Author's mail and address: (alberto.humala@bcrp.gob.pe), Jr. Miró Quesada 441, Lima-Perú. (511) - 6132785.*

reinforcing monetary policy, such as globalization or more-anticyclical fiscal policies? This paper's objective is to evaluate regime switches in inflation uncertainty in South America to answer a related question: does the whole region benefit from reduced volatility in inflation at short- and long-term alike?, that is, irrespective of the monetary policies in place? Complementarily, the analysis focuses on detecting feasible common patterns associated to any group effect in that volatility reduction.

Methodologically, this paper follows closely the work of Castillo, Humala, and Tuesta (2007). It uses an unobserved component model of inflation subject to regimen switching, borrowed from Kim and Nelson (1999). It decomposes inflation into its trend part and its transitory, stationary, path and considers shocks to both components being subject to independent regime shifts. An expectations-maximization algorithm is used to estimate the model subject to Markov switching. Preliminary, unrestricted autoregressive representation signals the number of regimes for each inflationary process. After testing appropriately for unit root in the inflation time series, a random walk representation is imposed in the trend component of inflation and the Markov switching heteroskedasticity model is estimated.

The rest of the paper is organized as following. The second section depicts the model and its rationale, with a brief reference to alternative representations. The third section describes data and unit root testing. In section four, the empirical results are presented and organized into sub groups of hyperinflationary experiences and non-hyperinflationary cases. Next section discusses and interprets results and, finally, the last section concludes and sets up research agenda.

## II. UNOBSERVED COMPONENT MODEL OF INFLATION

Inflation dynamics has shifted towards more stable scenarios in most South American countries during the last decade or so. Not only the level of inflation has declined but also, and importantly, the uncertainty (volatility) about inflation. Ball and Cecchetti (1990) put forward an interesting discussion about the relationship between level and volatility of inflation and reported positive parameters connecting the two for a sample of industrialized countries. Various South American economies suffered hyperinflation processes during the 1980s and most others experienced highly uncertain and inflationary scenarios. The recent move to price stability seems to reflect changes in the



conduct of monetary policy throughout the region. This sections borrowed from Kim and Nelson (1999) to present a Markov switching heteroskedasticity model that decomposes inflation into a trend part and a transitory component; both parts being subject to shocks that shift between regimes and that affect average level of inflation. The model allows for conditional and unconditional heteroskedasticity. Castillo, Humala, and Tuesta (2007) have recently applied empirically this model to inflation dynamics in Peru. This work extends the approach to all other South American economies.

The rationale behind decomposing the inflation time series is that deviations from the long-term path of inflation are due to changes in monetary policy, while that short-term variations might respond to temporary supply or demand shocks that will not impinge upon long term inflation patterns. Volatility in the first type of inflation (what is usually called core inflation) would reflect shocks to inflation trend and it is here conceived as coming from either a low-variance scenario or from a high-variance regime. Likewise, short-term shocks to temporary inflation might respond also to two alternative low-variance or high-variance regimes. Both types of shocks are indeed independent. Thus, the following unobserved component model of inflation dynamics is postulated:

$$(1) \quad \pi_t = \pi_t^T + \mu_2 S_{1,t} + \mu_3 S_{2,t} + \mu_4 S_{1,t} S_{2,t} + (h_0 + h_1 S_{2,t}) \eta_t$$

$$(2) \quad \pi_t^T = \pi_{t-1}^T + (Q_0 + Q_1 S_{1,t}) \varepsilon_t$$

Where  $\varepsilon_t \sim N(0,1)$  and  $\eta_t \sim N(0,1)$  are shocks to the stochastic trend component and to the transitory autoregressive component, respectively, of the inflation series.  $S_{1,t}$  and  $S_{2,t}$  are two independent first-order two-state Markov chains which represent the regime shifting in volatility of the stochastic component and the transitory component, respectively. Both of them are unobserved state variables. Notice out that those shifts in regime would also affect the average level of inflation through the  $\mu_i$  parameters,  $i=2,3,4$ . For each shock, the state variable defines a low-variance state by taking on the value 0 and a high-variance regime by taking on the value 1. These discrete Markov processes are represented by the transition probabilities:

$$(3) \quad \Pr[S_{1,t} = 0 / S_{1,t-1} = 0] = p_{00}, \quad \Pr[S_{1,t} = 1 / S_{1,t-1} = 1] = p_{11}$$

$$(4) \quad \Pr[S_{2,t} = 0 / S_{2,t-1} = 0] = q_{00}, \quad \Pr[S_{2,t} = 1 / S_{2,t-1} = 1] = q_{11}$$

Volatility of shocks to the permanent (transitory) component is associated to the value  $Q_0 (h_0)$  if those shocks come from a low-volatility regime and  $Q_0 (h_0) + Q_1 (h_1)$  otherwise. There are up to four different economic states resembling possible combinations of regime occurrence at time  $t$ . The first regime corresponds to a low-variance state for both Markov chains ( $S_{1,t} = 0$  and  $S_{2,t} = 0$ ); regime 2 stands for a low-variance regime in inflation trend and a high-variance state in the transitory part ( $S_{1,t} = 0$  and  $S_{2,t} = 1$ ); regime 3 is for a high-variance and low-variance states, respectively ( $S_{1,t} = 1$  and  $S_{2,t} = 0$ ); and, finally, the fourth regime corresponds to high-variance states in both Markov chains ( $S_{1,t} = 1$  and  $S_{2,t} = 1$ ). If permanent shocks come from high-variance states the inflation mean rises through the parameter  $\mu_2$ . The parameter  $\mu_3$  signals the effect of high volatility in transitory shocks and  $\mu_4$  stands for any feasible combined effect of high-variance states in both shocks.

The paper's objective is to detect whether inflation uncertainty, both at short- and long-term, is mostly at the low-variance states in the recent years for all South American economies. That is, whether  $S_{1,t}$  or  $S_{2,t}$  is close to the lower bound of the state probabilities, irrespective of the monetary policy features of each economy under analysis. There is actually no adjustment for any money or interest rate variable that could signal stance of monetary policy. This approach is univariate in nature and contributes to the analysis of the level-volatility relationship in inflation dynamics by introducing regime shifting into an unobserved component model as in Kim (1993). Econometric significance of estimated parameters (the four transition probabilities, the three  $\mu$  parameters, and the state-dependent volatility factors,  $Q$ 's and  $h$ 's, for each regime) are assessed against economic interpretation for each country. An expectation-maximization-type algorithm, as in Kim and Nelson (1999), is used to estimate all eleven parameters for each South American country independently.<sup>1</sup>

### III. THE DATA

Times series for inflation are taken as Consumer Price Index (CPI)'s quarterly percentage changes (after seasonally adjusting

<sup>1</sup> All ten major South American economies are considered (excluding French Guiana, Guyana, and Suriname). Empirical results for Peru, are directly taken from Castillo, Humala and Tuesta (2007), but outlier adjustment and unit root testing are further conducted for this case.

the index). Data samples for all countries but Brazil and Peru span over the period 1957:1 to 2006:2; that is, a total of 198 observations for each country. For Brazil, data is only available for the period 1980:1 to 2006:2 (106 observations) while that for Peru, the sample spreads out to starting in 1949:1 (229 observations). All data is taken from the International Financial Statistics Database build by the International Monetary Fund.

Visual inspection and basic statistics show a heterogeneous pattern for inflation dynamics among all countries, but a major stylized factor is that in most of them, their highest inflation levels took place during mid- or end-1980s (the so-called “lost decade” in South America).<sup>2</sup> Likewise, for most of them, inflation has recently reached their in-sample lowest. Two broad distinct groups could be identified, those economies with hyperinflation experiences (Argentina, Bolivia, Brazil, Peru, and, with a nearby hyperinflation episode, Chile) and those economies without that extreme experience, although with very unstable inflation records (Colombia, Ecuador, Paraguay, Uruguay, and Venezuela).

Preliminary, in order to identify the number of feasible regime switching in inflation dynamics for each country, unrestricted autoregressive representations of the corresponding inflation time series are estimated using the MSVAR software for Ox. Regime switching is allowed for all parameters (intercept, autoregressive, and variance). The number of lags is automatically selected, in their linear representation, using PcGets for Ox.<sup>3</sup> For all hyperinflationary countries, there seem to be evidence of three regimes corresponding to a price-stability state; a high-level, high-uncertainty inflation scenario; and a hyperinflation state. The other countries switch between a low and stable inflation regime and high and uncertain inflation state.

The Markov switching heteroskedastic model of inflation assumes a random walk representation for the inflation trend component. Therefore, unit root testing is in order to validate this assumption. In most cases, standard unit roots, Augmented

<sup>2</sup> With the exception of Chile, which experienced its largest inflation levels in the mid-1970. Inflation graphs and mean and standard deviation table statistics for all countries in the sample are available from the author upon request.

<sup>3</sup> This Ox package has also been used to detect automatically the presence of additive outliers in the linear representation of the series. In all cases, outliers are suggested for those periods that correspond to the peaks of the hyperinflation regimes in the nonlinear estimations or to the highest inflation levels in the other countries.

Dickey Fuller (ADF) and Phillips-Perron tests, reject the null hypothesis of a unit root in the inflation time series.<sup>4</sup> Given the Kim-Nelson's assumption of both inflation components being subject to regime switching, there is a point in assessing the validity of these tests for detecting unit root under the presence of Markov chains in the time series. Nelson, Piger, and Zivot (2000) have actually described a number of situations in which those tests should be adjusted when regime switching is present in the data generating process. Fortunately, in the Markov switching heteroskedastic model, the nature of the nonlinearity seems to validate the use of the standard tests. However, Perron and Rodríguez (2003) and Rodríguez (2004) have already shown that the presence of additive outliers in a time series tends to over reject a true unit root process. This seems to be the case in hand. Therefore, in order to account for the feasible presence of outliers, we borrow from Rodríguez (2004) a couple of outlier-adjusted ADF tests to detect unit root in our series.<sup>5</sup> These tests first detect the presence of additive outliers, then correct the standard testing equation by using dummy variables, and then implement these so modified ADF tests. In none of the cases the null of unit root is rejected.<sup>6</sup> Therefore, it is argued here that the assumption of a random walk representation for inflation trend is safe. A quite interesting issue, brought up by this testing, is the actual detection of the additive outliers in the series. Two alternative methods have been used for the linear representation of the inflation series. The automatic selection done by the PcGets software from Ox and the selection procedure in the outlier-modified ADF tests from Rodríguez (2004). Furthermore, the regime classification done by the MSVAR software, for a nonlinear representation of the series (especially for the hyperinflation cases) signal also a selection of feasible outliers.<sup>7</sup>

For the hyperinflationary cases, estimation is conducted by sub-samples, as to include in each of them only up to two regimes.

<sup>4</sup> Tests results have been obviated to save space. They are available from the author upon request.

<sup>5</sup> Results from these tests are also available from the author upon request.

<sup>6</sup> For Brazil, one of the tests rejects the null, but the other does not. Results for Bolivia show a somewhat explosive root, as previously detected by Rodríguez (2004).

<sup>7</sup> A table comparing dates in which outliers are considered for hyperinflation cases, with all these procedures, is available from the author upon request. This issue of efficiently detecting additive outliers should merit further empirical research. However, it goes beyond the scope of this paper.

With the expectations-maximization algorithm, considering up to three regimes would complicate the likelihood function as the number of total possible combinations of states rises considerably (with two regimes and two Markov chains, we have up to 16 different combinations). For all the other cases, the entire sample is considered at once in estimation of the model parameters.

#### IV. EMPIRICAL RESULTS

Though inflation dynamics show particular features in all countries under study, there are a number of common patterns that suggest some external factors might be as well influencing inflation trends or the roots of them. Sargent, Williams and Zha (2006), for example, argue that fiscal deficits are the bottom causes of accelerating inflation throughout hyperinflation countries in South America. Interestingly, an important impulse for those deficits was the oil international price crisis of the early 1970s. Differences among the countries about how they dealt with fiscal policy imbalances surely determine the heterogeneity in inflation dynamics, but most certainly, inflation-intolerant monetary policies in place in most countries during the 1990s and 2000s have curve down inflation levels and uncertainty. Notably, exceptions are associated to countries with loose monetary policies.

A significant difference among the countries is the presence of up to three states or regimes in inflation dynamics in those countries with experiences of hyperinflation and only up to two in those that did not reach those levels of inflation. Nonetheless, the similarities between the first two regimes among all countries are significant. In general, for most countries, the volatility of shocks to trend inflation has switched to the stable regime first, followed by the decrease in volatility of transitory shocks. In a few cases, reduction in volatility to transitory shocks has preceded that of permanent shocks.

In the figures that follow, each graph depicts the inflation series with the probability that shocks come from the highly uncertain regimes. The upper graph is for the permanent shocks and the lower graph for the temporary shocks. For the hyperinflationary cases, three sub samples (Argentina, Bolivia, and Peru) or two sub samples (Brazil and Chile) are presented to cover up all three regimes in the entire sample. Notice out that some overlapping in the sub samples is in place. Most of them

**TABLE 1.** REGIME SWITCHING HETEROSKEDASTICITY MODEL OF INFLATION, 1957-2006

<i>Parameters</i>	<i>Argentina</i>					
	<i>1957:2-1987:1</i>		<i>1990:3-2006</i>		<i>1972:1-1993:1</i>	
	<i>Estimates</i>	<i>St. Dev.</i>	<i>Estimates</i>	<i>St. Dev.</i>	<i>Estimates</i>	<i>St. Dev.</i>
$p_{11}$	0.906	0.056	0.834	0.108	0.940	0.062
$p_{00}$	0.969	0.018	0.965	0.024	0.961	0.030
$q_{11}$	0.887	0.068	0.926	0.064	0.895	0.227
$q_{00}$	0.901	0.046	0.967	0.036	0.960	0.045
$Q_{\phi}$	1.156	0.382	0.152	0.136	6.240	0.824
$h_0$	1.722	0.364	0.573	0.091	0.000	0.012
$Q_J$	24.407	3.652	46.112	9.752	47.937	7.221
$h_1$	1.771	0.858	2.275	0.687	8.541	3.148
$\mu_2$	8.732	1.226	0.037	0.621	-2.037	11.548
$\mu_3$	5.150	4.938	6.647	3.141	-11.187	6.089
$\mu_4$	3.395	9.360	4.059	4.425	30.942	27.130
$Q_J/Q_0$	21.113		302.565		7.683	
$h_J/h_0$	1.029		3.973		37,297.428	
Log likelihood	408.917		156.532		380.121	

**TABLE 2.** REGIME SWITCHING HETEROSKEDASTICITY MODEL OF INFLATION, 1958-2006

<i>Parameters</i>	<i>Bolivia</i>					
	<i>1958:2-1982:1</i>		<i>1978:1-1990:4</i>		<i>1985:4-2006</i>	
	<i>Estimates</i>	<i>St. Dev.</i>	<i>Estimates</i>	<i>St. Dev.</i>	<i>Estimates</i>	<i>St. Dev.</i>
$p_{11}$	0.525	0.148	0.930	0.059	0.986	0.019
$p_{00}$	0.776	0.135	0.976	0.025	0.992	0.009
$q_{11}$	0.824	0.086	0.953	0.071	0.289	0.149
$q_{00}$	0.940	0.030	0.870	0.183	0.855	0.046
$Q_{\phi}$	0.022	0.134	0.629	0.828	0.161	0.055
$h_0$	1.096	0.193	2.614	1.096	0.566	0.064
$Q_J$	0.000	0.026	53.406	9.567	17.442	3.755
$h_1$	6.049	1.093	3.247	1.507	0.907	0.264
$\mu_2$	8.911	2.748	2.557	2.818	3.417	0.572
$\mu_3$	2.699	0.443	2.099	8.704	2.767	1.178
$\mu_4$	-4.244	6.337	7.360	12.126	4.299	2.629
$Q_J/Q_0$	0.022		84.930		108.208	
$h_J/h_0$	5.521		1.242		1.602	
Log likelihood	240.529		211.512		167.341	

**TABLE 3.** REGIME SWITCHING HETEROSKEDASTICITY MODEL OF INFLATION, 1980-2006

<i>Parameters</i>	<i>Brazil</i>			
	<i>1980:2-1996:3</i>		<i>1995:3-2006:2</i>	
	<i>Estimates</i>	<i>St. Dev.</i>	<i>Estimates</i>	<i>St. Dev.</i>
$p_{11}$	0.955	0.037	0.362	0.181
$p_{00}$	0.978	0.024	0.704	0.109
$q_{11}$	0.975	0.027	0.678	0.405
$q_{00}$	0.712	0.314	0.930	0.058
$Q_0$	2.766	1.463	0.000	0.219
$h_0$	1.250	2.088	0.308	0.047
$Q_1$	29.663	4.448	0.238	0.242
$h_1$	1.285	1.959	2.462	0.912
$\mu_2$	19.192	3.198	3.979	2.997
$\mu_3$	14.125	10.673	1.263	0.144
$\mu_4$	-2.128	12.585	-0.588	5.194
$Q_1/Q_0$	10.725		-238,071.000	
$h_1/h_0$	1.028		7.987	
Log likelihood	267.528		67.568	

**TABLE 4.** REGIME SWITCHING HETEROSKEDASTICITY MODEL OF INFLATION, 1957-2006

<i>Parameters</i>	<i>Chile</i>			
	<i>1957:2-1991:2</i>		<i>1975:3-2006:2</i>	
	<i>Estimates</i>	<i>St. Dev.</i>	<i>Estimates</i>	<i>St. Dev.</i>
$p_{11}$	0.936	0.063	0.986	0.022
$p_{00}$	0.992	0.009	0.994	0.006
$q_{11}$	0.861	0.111	0.929	0.065
$q_{00}$	0.916	0.046	0.947	0.040
$Q_0$	1.026	0.196	0.432	0.156
$h_0$	0.892	0.190	0.295	0.080
$Q_1$	24.121	4.160	9.710	3.509
$h_1$	2.165	0.458	2.261	0.371
$\mu_2$	4.444	1.041	-0.581	0.883
$\mu_3$	8.211	2.025	4.937	0.812
$\mu_4$	-0.239	0.000	6.582	4.439
$Q_1/Q_0$	23.506		22.460	
$h_1/h_0$	2.427		7.676	
Log likelihood	386.974		218.991	

TABLE 5. REGIME SWITCHING-HETEROSKEDASTICITY OF INFLATION IN COUNTRIES WITHOUT HIPERINFLATION, 1957-2006

Parameter	Colombia		Ecuador		Paraguay		Uruguay		Venezuela	
	Estimates	St. Dev.	Estimates	St. Dev.	Estimates	St. Dev.	Estimates	St. Dev.	Estimates	St. Dev.
$\hat{p}_{11}$	0.259	0.414	0.785	0.083	0.523	0.163	0.988	0.026	0.935	0.040
$\hat{p}_{00}$	0.955	0.026	0.970	0.013	0.935	0.032	0.984	0.045	0.902	0.048
$\hat{q}_{11}$	0.995	0.006	0.717	0.104	0.972	0.017	0.738	0.116	0.566	0.277
$\hat{q}_{00}$	0.994	0.007	0.650	0.112	0.917	0.042	0.934	0.026	0.878	0.084
$\hat{Q}_0$	0.270	0.096	0.222	0.064	-0.265	0.074	0.862	0.105	0.000	0.030
$\hat{h}_0$	0.632	0.061	0.373	0.086	0.466	0.069	0.000	0.007	0.247	0.094
$\hat{Q}_t$	0.795	0.670	3.382	1.092	0.265	0.282	1.539	0.532	1.422	0.199
$\hat{h}_1$	2.422	0.356	1.013	0.159	1.398	0.175	5.859	0.940	5.674	0.790
$\hat{\mu}_2$	-0.916	1.309	1.074	0.267	0.539	0.494	2.432	3.035	2.527	2.643
$\hat{\mu}_3$	2.641	0.530	2.328	0.686	-0.142	0.766	2.487	0.765	0.841	0.600
$\hat{\mu}_4$	5.737	2.289	5.125	1.281	5.020	1.136	3.960	3.422	2.129	3.429
$Q_0/Q_0$	2.946		15.226		-1.000		1.786		474,048.0	
$h_t/h_0$	3.831		2.715		3.000		12,156.4		22.939	
Log likelihood	394.505		383.865		400.285		483.081		404.264	



are significant at usual levels. Notably, for Brazil, we found the less number of significant parameters, probably due to its shorter sample.

### 1. Hyperinflation countries

Inflation dynamics in all countries that experienced hyperinflation processes seem to shift among three regimes, with moves from the lowest inflation state to the high inflation, high-variance state before sliding down into the hyperinflation episode. As a result, in all the hyperinflation cases data had been split into sub samples containing up to two regimes for estimation.<sup>8</sup> A common feature is that these three regimes alternate in the order just described. There are no cases in which inflation dynamics has shifted directly from the low-inflation case to the hyperinflation scenario (or vice versa) without entering first the other intermediate regime. It should be mentioned out here that an important feature of the data in these cases is the presence of (additive) outliers that complicates convergence of the estimation algorithm. Frequently, those outliers correspond to economic policy identifiable attempts to curve down inflation in the form of the start of stabilization programs that unleashed price liberalization at once.<sup>9</sup> See Castillo, Humala, and Tuesta (2007) and Rodríguez (2004) for accounts of the presence of those outliers. The hyperinflationary countries are Argentina, Bolivia, Brazil, Chile, and Peru.

Argentina had a period of relative price stability up to mid-1970s (Figure 1). High volatility in shocks to trend inflation started soon afterwards and continued for most of that decade and the first half of the 1980s. The second upper panel of Figure 1 show high inflation and hyperinflation alternating. The latter one started to mount by the second part of the 1980s in the form of higher relative volatility. At the beginning of the

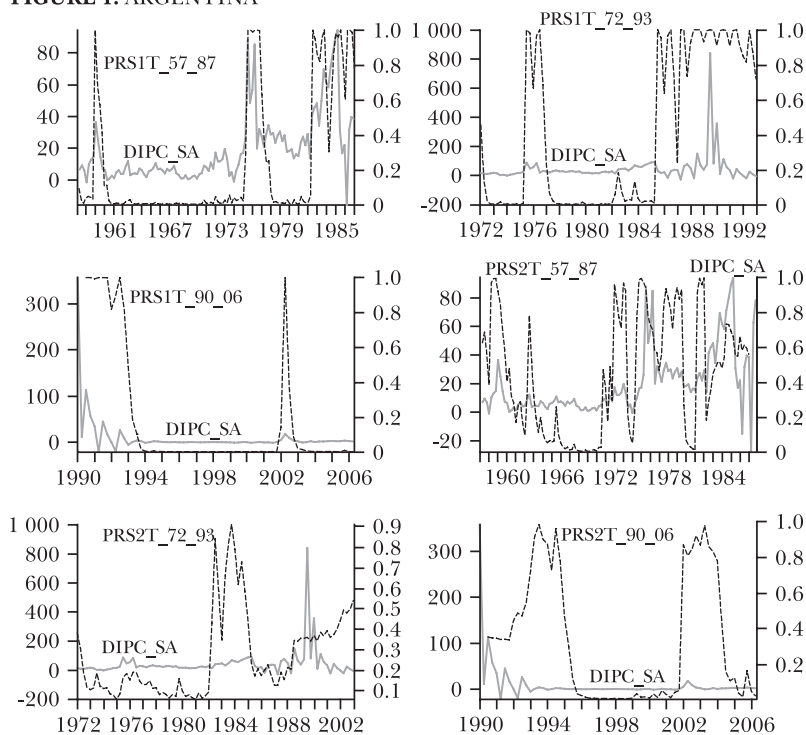
<sup>8</sup> The reference for sample definition is the regime classification from the unrestricted autoregressive representation.

<sup>9</sup> As with the case of detecting the outliers, dealing with them is a crucial issue. For instance, in those cases in which the outliers respond to policy decisions (the start of a stabilization program), are those observations to be averaged over a close-by sub sample?, should they be dropped from estimation? Dummy variables should be used instead? This paper favors the first approach. However, it could also be considered that type of outliers capture constrained forces (price controls) accumulated from the past and, therefore, should be taken as they are. There are no easy answers to this issue without further empirical research.

1990s, stabilization programs and the adoption of a fixed exchange rate system contributed to bring down inflation to more stable levels. Large volatility of transitory shocks accompanied for most parts of the unstable regimes in trend inflation (actually, even more frequently than volatility in trend inflation). The last upper graph in Figure 1 shows the significant decrease in volatility of shocks to trend inflation after 1993. It remained so up to current times only to be interrupted by the exchange rate regime collapse of 2001. A major empirical finding is that volatility in transitory shocks to inflation (the last lower graph) only ceases a while after trend inflation curves down permanently. It took a couple of years after 1993 and almost three years after 2001 to shift to the calm regime in transitory shocks. Once the 2001 crisis was overcome entirely, price stability and low-volatility of shocks to both inflation components are in place.

Two empirical facts are worth mention about the Argentine inflation. First, although they were independent processes, its hyperinflation dynamics and timing is quite similar to that of

FIGURE 1. ARGENTINA



Peru. Second, as it would be seen below, its inflation dynamics seem to influence importantly inflation dynamics in Paraguay and Uruguay during the 2001 currency board crisis.

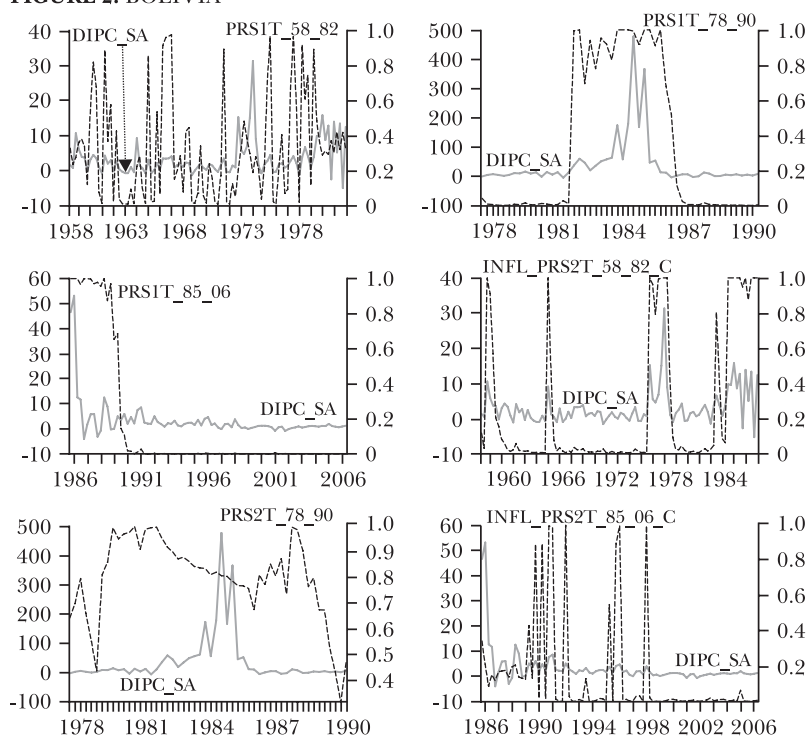
For Bolivia although there were not extended periods of price volatility at short- or long-term up to mid-1970s, there were numerous occasions in which shocks were highly volatile. By 1973 and 1974, it seems that pressures on permanent and temporary inflation started to accumulate (first panels in Figure 2). It was by the end of the 1970s and the first years of the 1980s that the inflation process reached highly uncertain levels and, subsequently, ended up in a hyperinflation process (middle panel in Figure 2). Most of the 1980s were marked by shocks of high volatility to temporary inflation. As in the Argentine case, volatility in temporary shocks decreases only after a while than volatility in permanent inflation diminishes. Successful stabilization programs (the second after Chile and earlier than in the rest of South America) curved down hyperinflation by 1985, but inflation remained unstable and high for most of the rest of that decade. The 1990s started with trend inflation stable. It took a while longer for the transitory shocks to diminish in uncertainty, till finally gets them into a stable regime after 1998 (around eight years later than for trend inflation shocks). Bolivia was the first proper hyperinflation case in South America and lasted for around 3 to 4 years.

Peru suffered its hyperinflation at the end of the 1980s (in similar years as in Argentina).<sup>10</sup> For most of the sample up till mid-1970s it benefited from price stability at the long-term, although there were sporadic volatile shocks to temporary inflation. It was after 1975 that pressures on inflation started to mount up, both at its trend and stationary components. As in the previous cases, the hyperinflation episode is clearly spotted by the large probabilities of high volatility regimes for both types of shocks. Hyperinflation lasted for around three years, 1988 to 1990. Once more, after a successful stabilization program at the beginning of the 1990s, shocks to trend inflation decreased in volatility by 1994. Up to four to five years later, volatility in temporary shocks to inflation also decreased. Thereafter (around 1999), both core and non-core inflation seems to be in a regime of price stability.

Brazil experienced a somewhat different type of hyperinflation

<sup>10</sup> Results for Peru are directly taken from Castillo, Humala, and Tuesta (2007). For graphs of inferred probabilities and parameter estimates, readers are kindly referred to that paper.

FIGURE 2. BOLIVIA

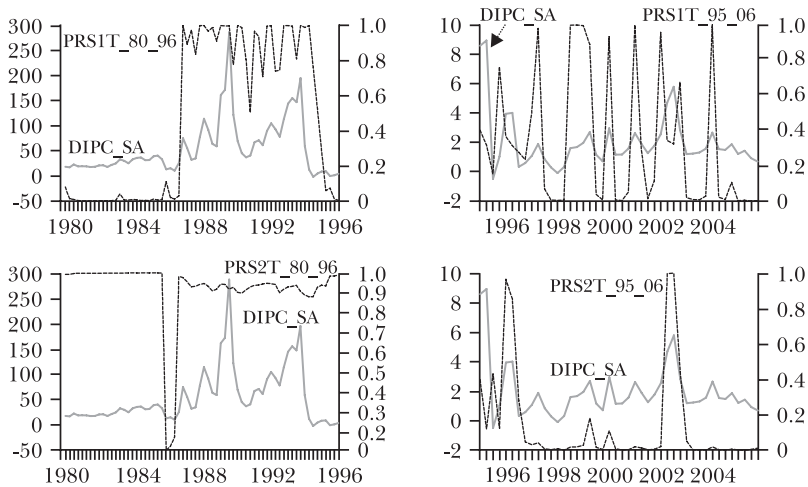


than the other cases in South American. First panel in Figure 3 shows volatility of shocks to trend inflation increased substantially by mid-1980s and ended up in hyperinflation levels by the end of the decade. Temporary shocks were volatile throughout that first sub sample. Stabilization programs succeeded only partially to restore stability, since by mid-1990s there was a resurgence of inflation to almost-hyperinflation levels (second panel). Differently to the other cases, volatility in trend inflation remained high for most of the 1990s and the beginning of the 2000s, while that for temporary shocks, there were only few cases of high volatility during this sub sample. It seems that as recently as 2005 onwards that price stability at both components have been reached in Brazil.

Although Chile did not reach the peaks in inflation the other South American hyperinflation cases did, it was the first economy to reach well above average level of inflation in the region by the mid-1970s (Figure 4). And it was the base for the first successful stabilization program in the region. In the first sub sample, up to 1990, probabilities of the volatile regimes again

clearly spot the near hyperinflation of the mid-1970s. Shifts towards the hyperinflation regime were around 1972-1973 and it lasted for 3 to 4 years.<sup>11</sup> Temporary shocks were highly volatile for most of this sample. Once more, although trend inflation lowers in volatility by the beginning of the 1980s (second panel of graphs), it took most of the decade for the transitory shocks to collapse to the stable regime. As early as the beginning of the 1990s, both inflation components reached their stable regimes and have remained in it ever after.

**FIGURE 3. BRAZIL**



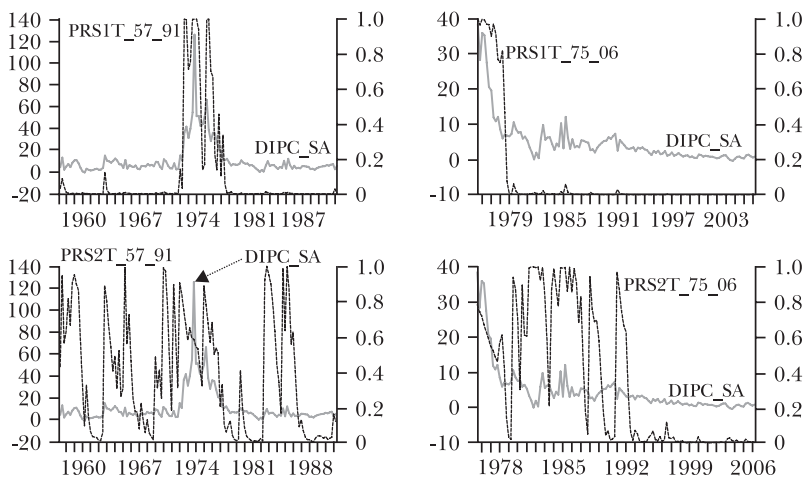
**2. Non-hyperinflation countries**

The pattern of inflation dynamics for those countries without the extreme experience of hyperinflation is also heterogeneous and different in the timing of shifts. Nonetheless, a number of regularities could also be single out in reference with the volatility of inflation. Recall, parameter estimation for these cases involves entire samples and includes up to two regimes for shifting volatility.

Colombia is one of the cases in South America that during the so-called lost decade of the 1980s had a moderate but stable inflation level. Actually, volatility in trend inflation is shown as higher during the 1960s and 1970s (Figure 5) and only again by the mid-1980s. Thereafter, it remains in the low-variance

<sup>11</sup> Oil price crisis as well as political turbulence in Chile might have influenced inflation dynamics.

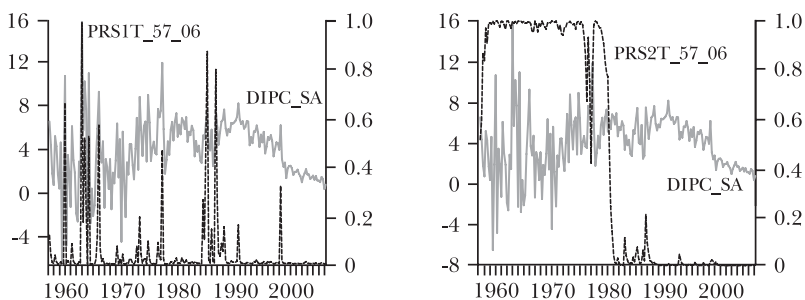
FIGURE 4. CHILE



regime. Transitory shocks are classified into the high-variance regime for most of the sample up to 1980, but remained stable thereafter. This might be the first empirical case, that the model represents the transitory shocks collapsing to the stable regime even before shocks to trend inflation. Notice out, that in both cases, stability was reached well before the beginning of the 1990s. There might be a case in arguing that even though the 1980s brought up moderate but stable inflation in Colombia, the country has benefited from very low and stable inflation levels in the 1990s, especially thereafter the second half of the decade. And the processes that generate those types of inflation levels are quite different. Econometrically, unrestricted non-linear representation of the inflation time series suggests the possibility of a third regime differentiating these two decades, but testing for them showed less clear results.

The experience of Ecuador is depicted in Figure 6. Its in sample

FIGURE 5. COLOMBIA



highest inflation observed is actually at the beginning of the 2000s, when they faced some political turbulence and adopted full dollarization. Nevertheless, once more, high-variance regimes seem to start around mid-1970s and flourish during most of the 1980s, especially mid- and end-1980s. Surprisingly, estimation results show that temporary shocks to inflation were in the high-variance regime for most of the entire sample, even well in the 2000s. It seems that trend inflation indeed has entered a somewhat stable regime after the adoption of full-dollarization, but it remains to be seen if short-term shocks to inflation also diminishes in volatility thereafter.

FIGURE 6. ECUADOR

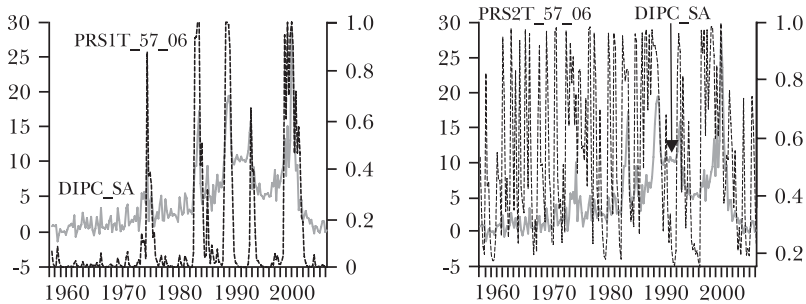
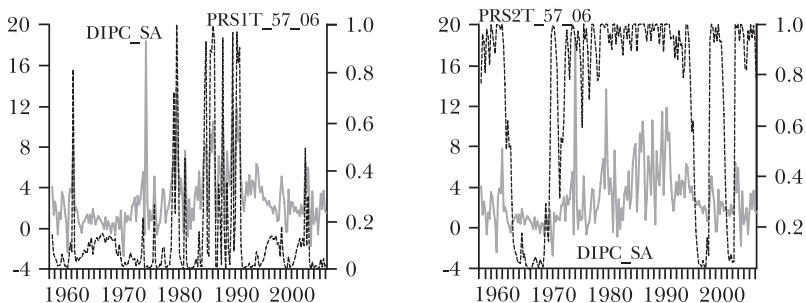


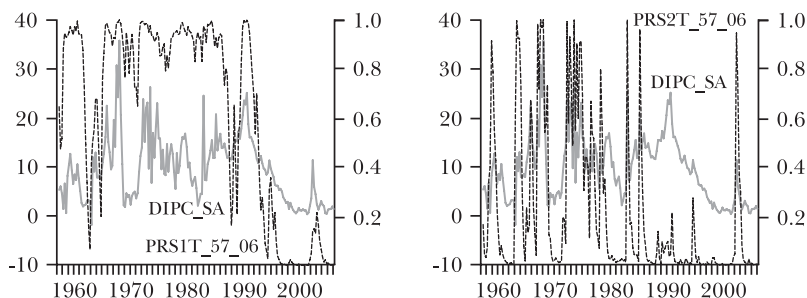
Figure 7 shows shifts in volatility of shocks to inflation in Paraguay. High-variance regime for trend inflation concentrates from mid-1970s to most of the 1980s and shortly during the 2001 currency board crisis in Argentina. High-variance stance for the transitory shocks is spread out throughout the sample. Although, not as clearer as in the previous cases mentioned, after the decrease in volatility of trend inflation by the start of the 1990s, volatility in transitory shocks seemed to decrease partially by mid-1990s but reassumed the high-variance stance during the 2001 crisis and remained in it thereafter.

FIGURE 7. PARAGUAY



Inflation dynamics in Uruguay has been quite volatile during most of the sample up until mid-1990s (Figure 8). Well before the oil price crisis of the mid-1970s, trend inflation was subject to highly volatile shocks. At the start of the 1990s, economic policies brought down inflation trend into the low-variance regime and remained there afterwards but for a short while that coincide with the 2001 Argentine crisis. The transitory component of inflation is classified into the high-variance stance for most of the sample, as well, up till mid-1980s, although seems to be a bit more sporadic. Again, short-term volatility seems to return briefly to the volatile regime during the 2001 crisis in Argentina.

**FIGURE 8. URUGUAY**



Volatility of both inflation components in Venezuela has been classified into the high-variance regime for most of the sample, including the 2000s (Figure 9). Although, Venezuela had relative low inflation records until mid-1980s, by the end of that decade, trend inflation started to rise considerably for its standard, but still low for South American standards. By the last part of the 1990s and most of the 2000s, although inflation decreased it remained at high level by the then standards in the region. Temporary shocks, although somewhat less frequently, were also at the volatile regime for most of the entire sample, with a tendency to shift to the low-variance regime in recent years (despite the fact that trend inflation remains at the high-variance state).

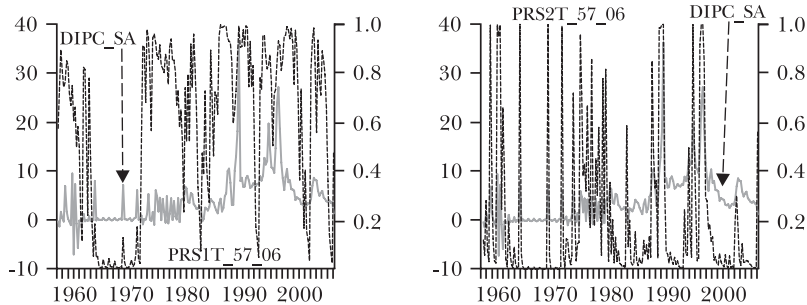
## V. SOUTH AMERICAN INFLATION TRENDS

A number of important trends and stylized facts stand out from the previous analysis. First of all, there is indeed a major shift towards price stability throughout the region, but there are also



a couple of noticeable exceptions (Ecuador and Venezuela). A third country, Paraguay could also be considered an exception in the case of transitory inflation. It is worth noting out that all these three cases correspond to economies that had moderate and relative stable inflation records during the 1980s. All of the hyperinflation economies, but Brazil, have benefited from price stability for quite a while. Their stabilization programs seem to have worked efficiently in curving down trend inflation and inflation expectations. Less volatile short-term shocks to inflation follow up price stability, although not immediately. Further analysis, though, is required to advance more precise conclusions on each country case.

**FIGURE 9. VENEZUELA**



## VI. CONCLUSIONS

The recent slowdown of inflation levels and volatility in South American economies is studied here to verify whether it corresponds to a downward trend in permanent inflation, in its short-term component or in both. A regime switching model of inflation dynamics is estimated for each economy, allowing for conditional and unconditional heteroskedasticity. Two regimes, of low- and high- variance, are considered for shocks to trend inflation and two for shocks to short-term inflation. Long-term shocks volatility decreases due to more inflation-intolerant monetary policy in place throughout the region and short-term shocks volatility decreases probably because of lower inflation expectations.

All time series in South American inflation show indeed a regime shifting behaviour. Countries with hyperinflation records move among three regimes, of price stability, high and unstable inflation, and hyperinflation. These countries seem to have reached earlier and for the longest period, price stability in the

recent years. Countries with moderate inflation during the 1980s seem to be the more reluctant to bring inflation down permanently.

A common pattern is that short-term inflation slows down in volatility consistently only after trend inflation has shifted convincingly towards a low-variance regime. The lag by which this occurs varies among countries (from two to three years to around a decade).

The timing of the downward trend in permanent inflation is not the same among all countries. It remains to be studied whether the shift to price stability in trend inflation is due to the adoption of an inflation-intolerant monetary policy or other factors, such as globalization or anticyclical fiscal policies, have also contributed in defining such a pattern. Furthermore, the effects are permanent so that volatility of shocks should remain low as long monetary policy anchors inflation expectations? A more structural approach to a multivariate model, as in Sims and Zha (2005) is in need to disentangle the roots to this downward trend in inflation. From an econometric perspective, the presence of outliers in the time series under analysis, in particular in the hyperinflation periods, poses a major concern in the efficiency of detection and in how to deal with them. Further empirical research is required.

## REFERENCES

- Ball, Laurence, and Stephen Cecchetti (1990), Inflation uncertainty at short and long horizons, *Brooking Papers on Economic Activity*, n° 1, pp. 215-54.
- Castillo, Paul, Alberto Humala, and Vicente Tuesta (2007), *Monetary policy, regime shifts, and inflation uncertainty in Peru (1949-2006)*, Central Reserve Bank of Peru, March (Working Paper, n° 2007-005).
- Kim, Chang-Jin (1993), "Unobserved-component time series models with Markov-switching heteroskedasticity: changes in regime and the link between inflation rates and inflation uncertainty", *Journal of Business and Economic Statistics*, vol. 11, n° 3, July, pp. 341-49.
- Kim, Chang-Jin, and Charles R. Nelson (1999), *State-space models with regime switching. Classical and Gibbs-sampling approaches with applications*, The MIT Press, Cambridge Massachusetts, London.

- Krolzig, Hans-Martin (1997), *Markov-switching vector autoregressions. Modelling, statistical inference, and application to business cycle analysis*, Springer (Lecture notes in economics and mathematical systems).
- Nelson, Charles R., Jeremy Piger and Eric Zivot (2000), *Markov regime-switching and unit root tests*, Board of Governors of The Federal Reserve System, September (International Finance Discussion Papers, n° 683).
- Perron, Pierre, and Gabriel Rodríguez (2003), "Searching for additive outliers in nonstationary time series", *Journal of Time Series Analysis*, vol. 24, n° 2, pp. 193-220.
- Rodríguez, Gabriel (2004), "An empirical note about additive outliers and nonstationarity in Latin-American inflation series", *Empirical Economics*, vol. 29, pp. 361-72.
- Sargent, Thomas, Noah Williams and Tao Zha (2006), *The conquest of South American inflation*, Mimeo ([http://homepages.nyu.edu/~ ts43/](http://homepages.nyu.edu/~ts43/)).
- Sims, Christopher A., and Tao Zha (2005), "Were there regime switches in US monetary policy?", *American Economic Review*, vol. 96, n° 1, March, pp. 54-81.

Diego Bastourre  
Jorge Carrera  
Javier Ibarlucia

*Commodity prices in Argentina.  
What does move the wind?*

Juan Carlos Castañeda Fuentes  
Juan Carlos Catalán Herrera

*Emigrant remittances and the real exchange  
rate in Guatemala: an adjustment-costs story*

Xiomara Archibald  
Denny Lewis-Bynoe  
Winston Moore

*Labour market flexibility in small island  
developing states*

Carmen Broto  
Javier Díaz-Cassou  
Aitor Erce-Domínguez

*The sources of capital flows volatility: empirical  
evidence for emerging countries*

María J. Nieto  
Garry J. Schinasi

*El marco para salvaguardar la estabilidad  
financiera en la Unión Europea: hacia  
una referencia analítica  
para evaluar su eficacia*

Kenneth Sullivan

*Presentación proactiva del desempeño*

Banco Central Europeo

*La titulización en la zona del euro*

PUBLICACIONES DEL CEMLA

CEMLA (ed.)

*Estimación y uso de variables no observables  
en la región*

Serie PROGRAMA DE INVESTIGACIÓN CONJUNTA

William A. Allen

*Metas de inflación: la experiencia británica*

Serie ENSAYOS

Rodolfo de Jesús Haro García

*Metodologías para la estimación matemática de  
la matriz de insumo-producto simétrica: a partir  
de las matrices de oferta y utilización  
asimétricas en una economía abierta*

Serie ESTUDIOS



El *Centro de Estudios Monetarios Latinoamericanos* es una asociación regional de bancos centrales de América Latina y el Caribe. Además participan, como miembros colaboradores, bancos centrales extrarregionales, organismos supervisores y entidades financieras regionales. El propósito principal de la Institución es, desde 1952, la cooperación entre sus miembros para promover un mejor conocimiento de temas monetarios y financieros en la región. Entre sus modalidades de acción el Centro realiza actividades de capacitación, divulgación y estudios, así como programas plurianuales de asistencia técnica en áreas de infraestructura del sector financiero. Asimismo, el CEMLA actúa como secretaria técnica en las reuniones de gobernadores y técnicos de banca central de nuestra región.

El CEMLA ofrece regularmente desde 1955 publicaciones periódicas que, al cabo del año, pasan revista y analizan los principales temas financieros y monetarios que ocupan la atención de los bancos centrales de América Latina y el Caribe. En el *Boletín* se presentan temas actuales de interés general que reúnen a una amplia audiencia para mantenerla a la vanguardia en los principales debates en materia de teoría y política económica, monetaria y financiera. *Monetaria* y *Money Affairs* (publicaciones bilingüe y en inglés, respectivamente) difunden estudios elaborados por investigadores de bancos centrales, instituciones financieras internacionales, universidades y otras entidades académicas. En los artículos publicados predomina un enfoque teórico o cuantitativo sobre temas especializados en los campos financiero y monetario.

## monetaria

*Suscripción anual: 80.00 dólares (América Latina y el Caribe: 50.00 dólares; estudiantes y maestros: 40.00 dólares). Ejemplar suelto: 22.00 dólares (América Latina y el Caribe: 15.00 dólares; estudiantes y maestros: 11.00 dólares).*

*Suscripciones y pedidos:*

*Aída Salinas Domínguez*

*CEMLA*

*Durango nº 54, México, D. F., 06700, México*

*Tel.: (5255) 5061-6651*

*Telefax: (5255) 5061-6659*

*E-mail: publicaciones@cemla.org*

## MIEMBROS DEL CEMLA

### ASOCIADOS

Banco Central de la República Argentina	Banco de Guatemala
Centrale Bank van Aruba	Bank of Guyana
Central Bank of the Bahamas	Banque de la République d'Haïti
Central Bank of Barbados	Banco Central de Honduras
Central Bank of Belize	Bank of Jamaica
Banco Central de Bolivia	Banco de México
Banco Central do Brasil	Bank van de Nederlandse Antillen
Eastern Caribbean Central Bank	Banco Central de Nicaragua
Cayman Islands Monetary Authority	Banco Central del Paraguay
Banco Central de Chile	Banco Central de Reserva del Perú
Banco de la República (Colombia)	Banco Central de la República Dominicana
Banco Central de Costa Rica	Centrale Bank van Suriname
Banco Central de Cuba	Central Bank of Trinidad and Tobago
Banco Central del Ecuador	Banco Central del Uruguay
Banco Central de Reserva de El Salvador	Banco Central de Venezuela

### COLABORADORES

#### *Bancos centrales*

Deutsche Bundesbank (Alemania)	Banca d'Italia
Bank of Canada	De Nederlandsche Bank (Países Bajos)
Banco de España	Bangko Sentral ng Pilipinas
Federal Reserve System (Estados Unidos)	Banco de Portugal
Banque de France	European Central Bank

#### *Otras instituciones*

Superintendencia de Bancos y Seguros (Ecuador)	Superintendencia de Bancos (República Dominicana)
Superintendencia del Sistema Financiero (El Salvador)	Banco Centroamericano de Integración Económica
Superintendencia de Bancos (Guatemala)	Banco Latinoamericano de Exportaciones, S. A.
Comisión Nacional de Bancos y Seguros (Honduras)	Deutscher Genossenschafts- und Raiffeisenverband e. V. (Confederación Alemana de Cooperativas)
Superintendencia de Bancos (Panamá)	Fondo Latinoamericano de Reservas