



## Cuadernos de economía

ISSN 0717-6821 *versión on-line*

Tabla de contenido

Cuad. econ. v.44 n.129 Santiago mayo 2007

### **SUBASTANDO LA ENERGÍA ELÉCTRICA PARA CLIENTES REGULADOS: EQUILIBRIO CON INFORMACIÓN COMPLETA Y AVERSIÓN AL RIESGO**

CARAVIA, FRANCISCO; SAAVEDRA, EDUARDO

- [resumen en español](#) | [inglés](#)
- [texto en español](#)
- [pdf en español](#)

### **EFFECTOS DISTRIBUTIVOS DE LA REFORMA DE LA SEGURIDAD SOCIAL: EL CASO URUGUAYO**

FORTEZA, ALVARO

- [resumen en español](#) | [inglés](#)
- [texto en español](#)
- [pdf en español](#)

### **¿PUEDE EL DISEÑO DE UN TORNEO DEPORTIVO AFECTAR SU ASISTENCIA?**

SERTSIOS, GIORGO

- [resumen en español](#) | [inglés](#)
- [texto en español](#)
- [pdf en español](#)

### **CONVERGENCIA Y ESTABILIDAD DE LOS TIPOS DE CAMBIO EUROPEOS: UNA APLICACIÓN DE EXPONENTES DE LYAPUNOV**

OLMEDO, ELENA; GIMENO, RICARDO; ESCOT, LORENZO; MATEOS, RUTH

- [resumen en español](#) | [inglés](#)
- [texto en español](#)
- [pdf en español](#)

### **EVIDENCE OF A BANK LENDING CHANNEL FOR ARGENTINA AND COLOMBIA**

GÓMEZ-GONZÁLEZ, JOSÉ; GROSZ, FERNANDO

- [resumen en español](#) | [inglés](#)
- [texto en inglés](#)
- [pdf en inglés](#)

© 2008 *Instituto de Economía, Pontificia Universidad Católica de Chile*

**Casilla 76, Correo 17  
Santiago - Chile  
Teléfonos 354 4314 - 354 4312**



[cuadecon@faceapuc.cl](mailto:cuadecon@faceapuc.cl)

## SUBASTANDO LA ENERGÍA ELÉCTRICA PARA CLIENTES REGULADOS: EQUILIBRIO CON INFORMACIÓN COMPLETA Y AVERSIÓN AL RIESGO\*

FRANCISCO CARAVIA  
*Fiscalía Nacional Económica de Chile*  
EDUARDO SAAVEDRA  
*ILADES-Universidad Alberto Hurtado*

*This paper studies the non-cooperative equilibrium of an electricity auction in which bidding firms are risk averse and have complete information. It assumes a centralized dispatch and stochastic hydraulic generation. We find that the auction allocates the energy contract to the generating firm that sets the spot market price: the thermal firm under moderate hydrological conditions and the hydraulic firm under a severe drought (rationing). In both cases the equilibrium price is greater than the regulated nodal price. Results are robust to introducing compensations for supply shortages. Finally, we find that if firms were risk-neutral they would bid the same prices and equal to the regulated nodal price, regardless of supply-shortage risk.*

*JEL:* D44, D81, L43, L94

*Keywords:* Subastas, Energía Eléctrica, Aversión al Riesgo, Oferta Incierta, Racionamiento Eléctrico.

### 1. INTRODUCCIÓN

Este artículo presenta un ejercicio teórico motivado en la modificación a la Ley Eléctrica en Chile<sup>1</sup>, la que liberaliza el precio de la energía que pagan los clientes regulados. La idea central de este cambio legal es la determinación del precio a través de subastas de bloques de energía con que se abastecen las em-

\* Agradecemos los comentarios de Soledad Arellano, Juan Pablo Montero y de los participantes en el Encuentro de la SECHI 2005 a una versión previa de este trabajo. Los posibles errores que persistan y las opiniones expresadas en este artículo son de nuestra exclusiva responsabilidad.

E-mail: saavedra@uahurtado.cl

<sup>1</sup> Ley 20.018, promulgada en mayo de 2005.

presas distribuidoras, con lo cual el precio que pagan los consumidores regulados ya no está directamente fijado por el regulador, sino que por el mercado. Motivado en esto, nos parece pertinente analizar las propiedades de los precios de equilibrio resultantes en este tipo de subastas, en particular cuando no han sido analizadas por la teoría en el contexto en que opera un sistema eléctrico como el de Chile central: con parque generador hídrico y térmico, sujeto a incertidumbre de oferta<sup>2</sup>.

Tres preguntas que nos interesa responder son: ¿serán estos precios mayores, iguales o menores a aquellos que fijaría un regulador bien informado? ¿de qué variables económicas depende una eventual diferencia entre el precio de la subasta competitiva y el idealmente regulado? ¿cuáles son las consecuencias en bienestar de introducir más competencia en el sector eléctrico?

Conforme a las preguntas previas, este artículo analiza el equilibrio que resulta de una subasta hipotética en que una empresa distribuidora contrataría su energía. Para ello, se modela un mercado con dos empresas generadoras, una hidráulica y otra térmica, ambas aversas al riesgo, que compiten por contratar el abastecimiento para clientes regulados de un único distribuidor independiente, vendiendo o comprando en el mercado *spot* sus diferencias de producción. Se supone que la generación hidráulica es estocástica con dos estados posibles: normal y seco; mientras que por simpleza se supone que la generación térmica es segura. Asimismo, como una primera aproximación al tema, se supone que en todo momento las empresas generadoras compiten entre sí y no siguen estrategias colusivas en las subastas de energía.

En términos de la secuencia del juego, se modela en primer lugar la decisión de qué precio cobrar por la energía a clientes regulados, competencia que se realiza a través de una subasta con sobre cerrado primer precio e incertidumbre de oferta. En la segunda etapa las empresas venden sus excedentes o compran energía en el mercado *spot*, lo que ocurre una vez despejada la incertidumbre hidrológica que caracteriza a este mercado.

Por último, y no menos importante, este artículo supone que la información de costos de ambas tecnologías es conocida por las empresas participantes en la subasta. Si bien este supuesto permite encontrar formas cerradas a las soluciones de la subasta, sus resultados no tienen por qué replicarse en caso de que hubiese información asimétrica entre los participantes<sup>3</sup>.

El principal resultado de este trabajo es que si no hay colusión en la subasta, ésta es adjudicada a la generadora que fija el precio del mercado *spot* del sistema en el estado de la naturaleza seco. En otras palabras, cuando la sequía es moderada es la empresa térmica la que margina y, por lo tanto, la que está dispuesta a ofrecer un precio más bajo en la subasta. Por otro lado, si la sequía es

<sup>2</sup> Una adecuada descripción de cómo operaba el sistema eléctrico en Chile previo a la modificación legal comentada se encuentra en Díaz, Galetovic y Soto (2001). Véase Comisión Nacional de Energía de Chile (2005) para una descripción del funcionamiento más reciente del sector.

<sup>3</sup> Este supuesto no significa que el regulador conozca los costos de las empresas que participan de la subasta. Es más, podemos suponer que el regulador no los conoce, lo que es consistente con la liberalización del precio de la energía para clientes regulados. Sólo suponemos un regulador bien informado cuando comparamos los resultados de la subasta con el precio nudo regulado.

extrema, al punto de producir desabastecimiento, entonces será la firma hidráulica la que ofrecerá el menor precio en la subasta. En cualquier caso, los precios de equilibrio son mayores al precio nudo que fijaría un regulador bien informado en cuanto a los costos de las empresas, generando rentas esperadas positivas para las generadoras. Es más, los precios de equilibrio son superiores al hipotético precio que maximiza el bienestar de los consumidores y deja a ambas empresas con utilidades esperadas no negativas. Sin embargo, tal situación que definimos como de máximo bienestar neto de los consumidores no es implementable, pues en general el regulador carece de información del grado de aversión al riesgo de los inversionistas en el sector.

Este resultado muestra el *trade off* en la decisión de política regulatoria, pues el esquema de precio regulado originado en una subasta libre de colusión no maximiza el bienestar neto de los consumidores, debido a las diferencias de costos entre ambas tecnologías usadas; mientras que por otro lado, un esquema de precio regulado basado sólo en información completa de costos de las empresas desincentiva al menos a una generadora para abastecer a clientes regulados, debido a que en este caso la firma con costos mayores, siendo aversa al riesgo, no obtiene beneficios esperados de participar.

Otras dos conclusiones de este artículo son las siguientes. Se encuentra que los resultados principales son robustos a un diseño diferente de las responsabilidades contractuales frente a una sequía extrema. En otras palabras, deba una empresa hidroeléctrica compensar o no a los usuarios por energía no abastecida, no cambia los resultados de la subasta. Asimismo, se encuentra que si ambas generadoras fueran neutrales al riesgo ofrecerían el mismo precio en la subasta e igual al precio nudo regulado, haya o no riesgo de desabastecimiento. Sólo en este caso es irrelevante para el bienestar social si utilizar el sistema regulado, con un regulador bien informado respecto de los costos de las empresas, o cambiarse a uno de subasta no colusiva.

La decisión de subastar el precio que pagarán los clientes regulados por su energía consumida, reemplazando el precio nudo o regulado, conlleva ventajas y desventajas para la sociedad que van más allá del *trade off* regulatorio ya mencionado. La principal ventaja es que el gobierno deja su rol de regulador en un mercado que es potencialmente competitivo, condición que en este artículo nosotros sólo lo suponemos y cuya veracidad dependerá crucialmente del diseño de la subasta que finalmente se implemente. De ser así, esta liberalización en la industria eléctrica debiera entregar mayor transparencia y certeza jurídica a potenciales inversionistas en el mercado de la generación eléctrica. Ello por cuanto el proceso fijación de precios para los clientes regulados es muy complejo debido a la gran cantidad de variables que considera el regulador en dicho proceso, existiendo espacio para un eventual manejo discrecional por parte de la autoridad al utilizar dichos parámetros<sup>4</sup>.

<sup>4</sup> Las empresas eléctricas han cuestionado la objetividad y la rigidez del proceso de fijación tarifaria, lo que ha producido diversas disputas entre generadoras, distribuidoras y regulador en Chile. Basañes, *et al.* (2001) reportan disputas en este tenor ya en las fijaciones tarifarias de los años 90s. Este cuestionamiento es avalado estadísticamente, pues las series de precio nudo y precios promedios de contratos con clientes libres no cointegran, como se esperaría cuando ni empresas ni regulador tienen mejor información en el largo plazo. Este análisis estadístico puede solicitarse directamente a los autores.

Por otro lado, el mayor problema de usar las subastas para tarifificar es la posibilidad de colusión entre las generadoras, en particular por el alto grado de concentración del segmento generación en países como Chile. Esta posibilidad se ve incrementada por la falta de incentivos de las empresas distribuidoras para monitorear las condiciones competitivas de estas subastas, ya que al fin y al cabo un distribuidor es un mero intermediario entre el consumidor y el generador eléctrico.

La literatura de subastas en la industria eléctrica ha surgido principalmente por la reestructuración del mercado *spot* en aquellos países en donde el despacho se descentralizó. Por ejemplo, el trabajo de Klemperer y Meyer (1989), que considera un modelo oligopólico con función de oferta bajo incertidumbre, es tomado y adaptado principalmente para los mercados eléctricos de USA y Reino Unido por Green y Newbery (1992), Von der Fehr y Harbord (1993), Newbery (1998); y en menor medida por los trabajos de Stacchetti (1999) y Fabra *et al.* (2004). En esta literatura el despacho eléctrico es realizado de acuerdo a los precios subastados independientemente de su costo marginal de producción. Sin embargo, en Chile como en otros países se mantiene el esquema de despacho centralizado (por orden de mérito o a costo marginal), con lo que la literatura sobre subastas en el mercado eléctrico sólo tangencialmente podría ayudar a entender las consecuencias económicas de la liberalización del precio a clientes regulados como el introducido en Chile.

Como consecuencia de lo anterior, el principal aporte de nuestro trabajo radica en el estudio de los precios de la energía de equilibrio resultantes de una subasta no colusiva cuando el despacho continúa siendo de acuerdo a condiciones de costo de las empresas generadoras. Una omisión consciente de nuestro trabajo es que no se estudia el diseño mismo de la subasta; sólo se supone que el regulador hace un diseño óptimo de la subasta de forma tal que las empresas generadoras ofrecen precios que reflejan sus verdaderos costos de oportunidad.

Nos parece que a pesar de ciertos supuestos simplificadores en la modelación, como el suponer información completa entre los oferentes, nuestros resultados son también importantes para la política energética de Chile. El trabajo muestra la racionalidad de poner techos razonablemente elevados a los precios determinados por la subasta, por encima del precio que resulta de la información de costos de las empresas. Si la subasta es capaz de imitar las condiciones de mercado competitivo, ello tendría consecuencias positivas sobre la asignación eficiente de recursos en la industria<sup>5</sup>, además de entregar las señales correctas a los inversionistas en generación, situación que ha sido deficitaria y es considerada como un potencial cuello de botella para el desarrollo del sector y del propio país<sup>6</sup>.

<sup>5</sup> Esto es de importancia si la demanda por energía es elástica. Al respecto, Benavente *et al.* muestran una no despreciable elasticidad de demanda residencial por energía eléctrica, estimándose en 0,048 en el corto plazo (un mes) y 0,39 en el largo plazo.

<sup>6</sup> La creencia de que la falta de inversiones pueda llevar a desabastecimiento ha sido analizada por Galetovic *et al.* (2002 y 2002a), quienes encuentran que la posibilidad de racionamiento es baja. No obstante, esos estudios son pertinentes a un contexto de bajo crecimiento económico, lo cual no es necesariamente cierto para Chile a partir del año 2004 en adelante.

El trabajo se estructura de la siguiente forma. En la sección dos se plantea el modelo teórico. Su resolución es entregada en la sección tres. La sección cuatro analiza los resultados encontrados desde el punto de vista del bienestar social. Enseguida, la sección cinco estudia la robustez de los resultados al considerar compromisos contractuales alternativos para las generadoras contratadas. Por último, la sección seis concluye.

## 2. EL MODELO

Se supone un modelo abstracto de un sistema eléctrico abastecido por dos firmas, ambas aversas al riesgo, una de generación hidráulica y otra térmica<sup>7</sup>. La capacidad de generación de la firma hidráulica depende de la hidrología del año, donde la capacidad máxima de generación se produce en un año lluvioso. Para la firma térmica su capacidad productiva es constante. Se supone que las capacidades de producción de ambas firmas están dadas y no pueden ser modificadas, al menos en el corto plazo.

Los costos marginales de generación para un año con una hidrología promedio (indistintamente normal o lluvioso) son  $c_H$  para la firma hidráulica (H) y  $c_T$  para la firma térmica (T), con  $c_H < c_T$ , lo que implica que la firma hidráulica es llamada primero a producir en el mercado *spot*. En el caso de sequía extrema, el costo marginal de generación hidráulica es igual al costo de falla del sistema y, por lo tanto, la relación de costos de producción entre las firmas queda  $c_H^F > c_T$ , lo que implicaría que sólo en caso de desabastecimiento es la planta térmica la llamada a producir en primer lugar. Este caso de sequía extrema se reserva para analizar la robustez de nuestros resultados.

El estado de la naturaleza lluvioso tiene una probabilidad  $\alpha$ , en el cual sólo la firma hidroeléctrica es llamada a producir; mientras que el estado de la naturaleza seco tiene una probabilidad  $(1-\alpha)$  y ambas firmas producen. Como se mencionó, la capacidad de generación de la firma hidráulica depende del estado de la naturaleza que se observa, siendo ésta  $k_\alpha$  en el estado lluvioso y  $k_{1-\alpha}$  en el estado seco, con  $k_\alpha > k_{1-\alpha}$ . La firma térmica tiene una capacidad de generación,  $k_T$ , independiente del estado de la naturaleza que se observa.

Dada una demanda contratada  $\bar{f}$  de la distribuidora, las firmas generadoras pueden comprar (vender) el faltante (sobrante) de la energía en el mercado *spot* de acuerdo al estado de la naturaleza que se presenta. Dado el supuesto que sólo la firma hidráulica es llamada a producir en un año lluvioso, esto implica que su capacidad en ese estado es suficiente para cubrir la demanda de los clientes regulados de la distribuidora y vender en el mercado *spot* el sobrante, esto es  $k_\alpha > \bar{f}$ . Para el caso de un año seco no extremo, la capacidad de producción de ambas firmas logra cubrir la demanda de la distribuidora. Consistente con esto,

<sup>7</sup> Esta modelación de un mercado de generación eléctrica con mezcla de tecnologías hidráulica y térmica ya ha sido utilizada en trabajos previos, como los de Arrellano (2004) y Serra (1997).

se supone sin pérdida de generalidad que no existe un mercado de clientes libres, ya que su incorporación sería redundante con el diseño de la subasta que hace la distribuidora.

Finalmente, sea  $u(x)$  la función de utilidad del beneficio de los dueños de estas empresas. Se asume que  $u(x)$  es continua, creciente y dos veces diferenciable, con  $u'(x) > 0$  y  $u''(x) < 0$  para todo  $x > 0$  y  $u(0) = 0$ , en donde  $x$  representa los pagos o beneficios efectivos de cada empresa en un estado de la naturaleza determinado.

Previo al análisis de las características de equilibrio de la subasta, veamos qué predice el modelo respecto de la situación anterior a la entrada en vigencia de la modificación legal, en donde las distribuidoras pagan a las generadoras el precio nudo regulado. Se trata por tanto de una subasta por un bloque de energía con el precio determinado exógenamente por el regulador. El siguiente resultado muestra qué debiera teóricamente suceder con los contratos de abastecimiento de las distribuidoras en este caso (ésta y todas las demás proposiciones, lema y corolarios se demuestran en el Anexo).

**Proposición 1:** Sean  $f^H$  y  $f^T$  las ofertas de las firmas hidráulica y térmica, respectivamente. Si el precio nudo regulado refleja el costo esperado del sistema  $P_N = \alpha c_H + (1-\alpha)c_T$  y  $u(x)$  es creciente, estrictamente cóncava y satisface las condiciones de Inada, entonces la firma hidráulica preferirá contratar en el mercado regulado sólo hasta su capacidad de generación en el estado de sequía,  $f^H = k_{1-\alpha}$ ; mientras que la firma térmica estará dispuesta a abastecer toda la demanda de la distribuidora, aunque ésta sea mayor a su propia capacidad,  $f^T = \bar{f}$ .

Nótese en primer lugar que no se determina qué firma finalmente abastecerá la demanda del sector regulado, o en qué proporciones esto se hará. Simplemente, este resultado nos dice que entre ambas empresas deberían contratar toda la demanda de la distribuidora y, sin importar los costos marginales en uno u otro estado de la naturaleza, la firma hidráulica no corre el riesgo de llegar a ser deficitaria; sin embargo, el generador térmico será deficitario en el estado normal-lluvioso. Lo segundo, y más importante en términos de política pública, es que este resultado nos dice que si ello ha sido diferente en la práctica, la razón debe encontrarse en la diferencia entre el precio nudo y los costos efectivamente esperados por los empresarios del sector generación. Tal diferencia se debe a los ajustes que se le hace al precio nudo efectivo por los contratos con clientes libres, o a diferencias entre el regulador y las empresas generadoras en cuanto a la estimación de los costos de operación de las plantas o a las probabilidades de ocurrencia de los eventos.

De aquí en adelante supondremos que el precio nudo o regulado no es restrictivo y, por lo tanto y de acuerdo a la ley, la subasta cambia su estrategia desde una que decía “oferto tal cantidad al precio máximo fijado por el regula-

dor” a una que dice “oferto a tal precio la cantidad requerida”. En consecuencia, los resultados del resto del trabajo son comparables con el de la Proposición 1 teniendo en cuenta esta diferencia sustantiva en el modo y las consecuencias en precios de la subasta por abastecer a los clientes regulados.

### 3. EQUILIBRIOS EN LA SUBASTA NO COOPERATIVA

Suponemos que las firmas compiten por precio en una subasta única y competitiva por abastecer el mercado *forward* o de clientes regulados de una única empresa distribuidora, cuya demanda,  $\bar{f}$ , es conocida por ambos generadores e inelástica<sup>8</sup>. El tipo de subasta es de sobre cerrado primer precio. La secuencia del juego es la siguiente: observada la demanda, las empresas generadoras eligen simultánea e independientemente sus ofertas  $b_i \in [0, P]$ ,  $i = H, T$ , donde  $P$  es el precio de reserva de mercado determinado por el regulador a partir de la disposición legal (precio máximo). Si las ofertas difieren, la firma con el menor precio debe abastecer el total de la demanda, independiente del estado de la naturaleza que ocurra. Suponemos si ambas ofertas son iguales, la empresa que margina en el estado de la naturaleza seco obtiene el contrato con la distribuidora<sup>9</sup>. Una vez asignado el ganador, la naturaleza se revela y se producen los pagos correspondientes. En suma, las empresas generadoras deben elegir el precio a abastecer el bloque de energía que licita la empresa distribuidora para los clientes regulados, tal que maximice el pago (utilidad) esperado para sus dueños.

De acuerdo a los supuestos previos, las utilidades esperadas de los dueños de la firma hidroeléctrica son:

$$(1) \quad \Pi_H^E(b_H) = \begin{cases} \alpha \cdot u(b_H \bar{f} + c_H(k_\alpha - \bar{f}) - c_H k_\alpha) + \\ (1 - \alpha) \cdot u(b_H \bar{f} + c_T(k_{1-\alpha} - \bar{f}) - c_H k_{1-\alpha}) & \text{si } b_H < b_T \\ \\ \alpha \cdot u(c_H k_\alpha - c_H k_\alpha) + \\ (1 - \alpha) \cdot u(c_T k_{1-\alpha} - c_H k_{1-\alpha}) & \text{si } b_H \geq b_T \end{cases}$$

<sup>8</sup> Este supuesto es equivalente a que la distribuidora licite un único bloque de energía. Ciertamente, si licitara varios bloques de energía los resultados podrían diferir, lo que no es analizado en este trabajo.

<sup>9</sup> Este supuesto es más conveniente que el alternativo de suponer que a ofertas iguales ambas generadoras contratan en partes iguales. Nuestro supuesto nos permite encontrar un equilibrio de Nash refinado para este juego, equilibrio que es bastante intuitivo según vemos en la Proposición 3, más adelante.



Mientras que los pagos de la firma térmica son:

$$(2) \quad \Pi_T^E(b_T) = \begin{cases} \alpha \cdot u(b_T \bar{f} - c_H \bar{f}) + \\ (1 - \alpha) \cdot u(b_T \bar{f} + c_T(k_T - \bar{f}) - c_T k_T) & \text{si } b_T \leq b_H \\ \\ \alpha \cdot u(c_T 0 - c_T 0) + \\ (1 - \alpha) \cdot u(c_T k_T - c_T k_T) & \text{si } b_T > b_H \end{cases}$$

Nótese que las utilidades esperadas de las firmas que no tienen contratos se simplifican bastante, siendo igual a  $(1 - \alpha) \cdot u(\Delta c \cdot k_T)$  para la empresa hidráulica y de 0 para la empresa térmica, donde  $\Delta c \equiv c_T - c_H$ .

Sea  $P_i^*$ , para  $i = \{H, T\}$ , el precio tal que la utilidad esperada para la firma  $i$  de ganar la subasta es igual a la utilidad esperada de perderla. Es decir,  $P_H^*$  y  $P_T^*$ , respectivamente resuelven:

$$(3) \quad \begin{aligned} \alpha \cdot u((P_H^* - c_H) \bar{f}) + (1 - \alpha) \cdot u((P_H^* - c_T) \bar{f} + \Delta c k_{1-\alpha}) &= (1 - \alpha) \cdot u(\Delta c k_{1-\alpha}) \\ \alpha \cdot u((P_T^* - c_H) \bar{f}) + (1 - \alpha) \cdot u((P_T^* - c_T) \bar{f} + \Delta c k_T) &= 0 \end{aligned}$$

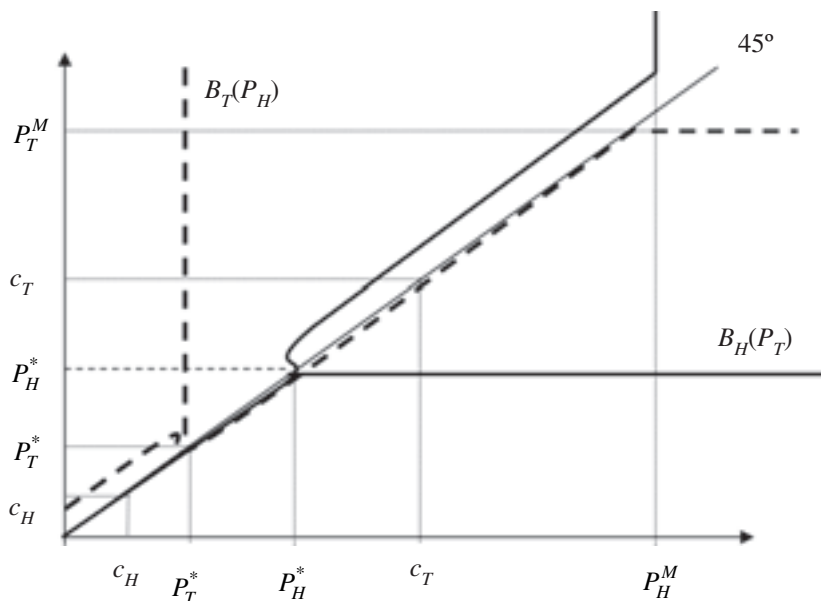
Hay que destacar que dada la concavidad en  $u(x)$ , ambos precios son únicos.

**Lema 1:** Si los dueños de las empresas generadoras son aversos al riesgo, el precio al cual los pagos de ganar la subasta se iguala al de perderla es mayor para la firma hidráulica que para la térmica, es decir,  $P_H^* > P_T^*$ .

El Lema 1 no parece muy intuitivo a primera vista, pues la firma con menores costos (generador hidráulico) tiene un precio mayor que la deja indiferente entre ganar o perder la subasta. La explicación es que considerando que la utilidad esperada de la firma hidráulica por perder la subasta es positiva y por lo tanto mayor al de la firma térmica, la firma hidráulica enfrenta menores incentivos a disminuir su oferta comparados a los de la firma térmica. Gracias a este Lema es posible encontrar las correspondencias de reacción de cada una de estas empresas,  $B_H(P_T)$  y  $B_T(P_H)$ , que se muestran en la Figura 1.

En línea gruesa se muestra la correspondencia de reacción para la firma hidráulica,  $B_H(P_T)$ , la que bajo  $P_H^*$  en rigor es cualquier precio mayor que  $P_T$ . La explicación es simple. Cuando la firma térmica oferta un precio mayor o igual al precio que pondría la firma hidráulica cuando actúa como monopolista, la firma hidráulica como mejor respuesta debe ofrecer el precio máximo que le permite la ley, el que llamamos  $P_H^M$ , debido a que con esto obtiene el máximo de beneficios posibles. Si el precio de la firma térmica es menor que  $P_H^M$  y mayor que  $P_H^*$ ,

FIGURA 1  
CORRESPONDENCIAS DE REACCION DE LAS FIRMAS GENERADORAS



la firma hidráulica ofrece como mejor respuesta un centavo menos que su rival, se adjudica la subasta y obtiene beneficios y pagos positivos. Si la firma térmica ofrece un precio menor o igual a  $P_H^*$ , la firma hidráulica, como mejor respuesta, ofrece cualquier precio mayor con el fin de no adjudicarse la subasta pues obtendría beneficios negativos.

De manera similar se encuentra la correspondencia de reacción de la firma térmica frente a los distintos precios que puede ofrecer la firma hidráulica,  $B_T(P_H)$ , graficada con línea punteada en la Figura 1. La única diferencia está dada porque la firma térmica sabe que ofreciendo lo mismo que la hidráulica se quedará con el contrato. Esto lleva a que la mejor respuesta de la firma térmica cuando la firma hidráulica ofrece un precio menor que  $P_T^M$  y mayor que  $P_T^*$  es ofrecer lo mismo que su rival; mientras que si tal precio es inferior a  $P_T^*$  entonces su mejor respuesta es ofrecer cualquier precio superior al ofertado por la firma hidráulica.

De las correspondencias de reacción se desprende directamente el siguiente resultado:

**Proposición 2:** Si los dueños de las empresas generadoras son aversos al riesgo y ante una misma oferta la subasta se le adjudica a la empresa que margina en el estado seco de la naturaleza, entonces existen múltiples equilibrios de Nash en estrategias puras. Todos ellos cumplen con que  $P_T = P_H$  y  $(P_T, P_H) \in [P_T^*, P_H^*]$ , siendo la empresa térmica la que obtiene el contrato.

El resultado de la subasta es que la firma térmica se adjudica el contrato de la distribuidora ofreciendo un precio menor que la firma hidráulica, aun teniendo un costo marginal de operación mayor a su rival en ambos estados de la naturaleza. Si bien este resultado no es intuitivo debido a la diferencia de costos a favor de la empresa hidráulica, es precisamente esta diferencia de costos la que lleva a la firma hidráulica a no ganar la subasta en precios, pues al ser llamada primero a producir en el estado seco obtiene beneficios positivos por vender su capacidad de producción en el mercado *spot* a un precio  $c_T$  mayor que su costo marginal  $c_H$ , lo que a su vez le impone un costo de oportunidad al precio mínimo a ofrecer en la subasta competitiva. Por otra parte, debido a que la firma térmica no obtiene beneficios al vender su capacidad en el mercado *spot*, porque su costo marginal fija el precio de dicho mercado, el precio mínimo que ofrece en la subasta es menor que su rival adjudicándose el bloque de energía subastado por la distribuidora.

Es importante señalar que la multiplicidad de equilibrios es propia de todo modelo tipo Bertrand con costos marginales constantes, pero diferentes entre las firmas, cualquiera sea el criterio de asignación de la subasta en caso de igualdad en las ofertas. En este caso se supuso que ante una misma oferta la subasta es asignada a la empresa térmica, lo cual llevó a una multiplicidad de equilibrios de Nash caracterizados por  $P_T^* \leq P_H \leq P_H^*$  y  $P_T = P_H$ . Este resultado no difiere mayormente si la regla de desempate fuera una en que se asigna el 50% del abastecimiento a cada una de las generadoras cuando las ofertas son iguales. En tal caso, la multiplicidad de equilibrios de Nash son caracterizados por  $P_T^* \leq P_H \leq P_H^*$  y  $P_T = P_H - \varepsilon$  ( $\varepsilon \approx 0$ ) cuando  $P_H > P_T^*$  y  $P_T = P_H$  cuando  $P_H = P_T^*$ .

Un corolario de esta proposición nos dice que si las generadoras fueran neutrales al riesgo, el resultado que se obtiene es que las dos firmas ofrecen el mismo precio y dicha oferta es igual al precio nudo fijado por el regulador, cuando éste conoce los costos de cada tecnología. Así, sólo en este caso de neutralidad al riesgo y simetría de información entre firmas y regulador respecto de costos, es irrelevante si las distribuidoras subastan sus necesidades de energía o si el regulador fija el precio de transferencia de energía entre generadores y distribuidores. En consecuencia, sólo en este caso particular la situación previa y la posterior a la modificación legal entregarían el mismo resultado para los consumidores. Como se dijo, este corolario aplica en tanto el precio nudo corresponda a su valor teórico y refleje fielmente los costos esperados del sistema, lo que requiere suponer que el regulador tiene información completa respecto de los costos de las firmas.

**Corolario 1:** *Si las firmas fueran neutrales al riesgo y  $P_N$  es el precio regulado que representa los costos esperados del sistema, entonces sólo en este caso existe un único equilibrio de Nash, precisamente uno en que los precios subastados por ambas firmas son aquellos que las dejan indiferentes entre ganarse o no la subasta, e iguales al precio nudo; es decir,  $P_H^* = P_T^* = P_N$ .*

Debido a la existencia de múltiples equilibrios de Nash en estrategias puras que encontramos cuando los agentes son aversos al riesgo, determinamos a continuación un equilibrio refinado que elimina aquellos equilibrios de Nash que contienen estrategias débilmente dominadas. Se utiliza el refinamiento de la “mano temblorosa” desarrollado por Selten (1975), de acuerdo a la adaptación para juegos estáticos presentada en Mas-Colell, *et al.* (1995). Para aplicar este refinamiento se supone que el espacio de estrategias de los jugadores es discreto.

**Proposición 3:** *Dados los supuestos acerca de la aversión al riesgo, la única estrategia que es Equilibrio de Nash de la “mano temblorosa” es  $(P_H, P_T) = (P_H^*, P_H^*)$ .*

Este resultado nos dice que la firma térmica se adjudica la subasta con el precio que iguala la utilidad esperada de ganar la subasta con la de perderla para la firma hidráulica. Esto implica que ambas empresas generadoras tienen una estrategia que débilmente domina al resto de los posibles precios de elección. Como este resultado tiene consecuencias económicas importantes, éstas son analizadas en la siguiente sección.

Es importante destacar que la unicidad del equilibrio refinado no depende tampoco del supuesto de asignación de la subasta cuando los precios son iguales; es más, tampoco depende de este supuesto el resultado de poder de mercado encontrado que se desprende de este único equilibrio. La razón es muy simple, si se modifica la regla de asignación a una en que a iguales precios el contrato con la distribuidora se divida en partes iguales, el conjunto de equilibrios de Nash no cambia; sin embargo, en tal caso la firma térmica tendría incentivos a reducir su oferta al precio inmediatamente inferior (recuérdese que se ha supuesto que los precios ofertados son discretos). Es fácil ver que en tal caso el único equilibrio de Nash refinado sería  $(P_H, P_T) = (P_H^*, P_{n-1}^*)$ , donde  $P_{n-1}^*$  es el precio inmediatamente inferior a  $P_H^*$ , resultado que es similar al nuestro en tanto  $P_{n-1}^*$  sea lo suficientemente cercano de  $P_H^*$ <sup>10</sup>.

#### 4. PODER DE MERCADO Y BIENESTAR

El equilibrio refinado que caracteriza el resultado de la competencia en la subasta por energía es único y, además, permite que la firma térmica obtenga beneficios esperados positivos. Esto por cuanto el precio adjudicado es mayor que el precio que iguala sus beneficios de ganar la subasta con los de perderla.

<sup>10</sup> Véase la demostración de la Proposición 3 para detalles de notación así como de los supuestos necesarios para poder utilizar el refinamiento de la “mano temblorosa” a este ejemplo. Es importante notar que si los saltos discretos en precios fuesen muy grandes, entonces podría ser rentable para la firma térmica compartir el contrato en lugar de quedárselo para ella sola pero a un precio menor.

Este resultado es cierto aún cuando se haya supuesto que las firmas no se coluden para ofertar precios en la subasta. Asimismo, también la firma hidráulica obtiene beneficios esperados positivos debido a que vende su capacidad disponible en el mercado *spot* en el estado de naturaleza seco, recibiendo un precio igual a  $c_T$  mayor que su costo marginal  $c_H$ . Ambos beneficios esperados positivos producen a su vez utilidades esperadas positivas tanto para el generador hidráulico como para el térmico, gracias a que el precio subastado por ambas firmas es el mismo e igual a  $P_H^*$ <sup>11</sup>.

El sobreprecio que se encuentra en equilibrio tiene relación con el poder de mercado inframarginal que posee la firma hidráulica en el estado de naturaleza seco. Renta que también capta la empresa térmica en la interacción estratégica durante la subasta, debido a que el generador hidráulico no está dispuesto a correr riesgos mayores bajando su puja por debajo de  $P_H^*$ .

Sabiendo que a igual precio el generador hidráulico obtiene rentas mayores que el generador térmico, definimos  $P^S$  como el mínimo precio que deja al generador térmico con una utilidad esperada igual a cero. Es decir,  $P^S$  satisface:

$$(4) \quad \alpha \cdot u\left(\left(P^S - c_H\right)\bar{f}\right) + (1 - \alpha) \cdot u\left(\left(P^S - c_T\right)\bar{f}\right) = 0$$

Es fácil ver que  $P^S$  es precisamente igual a  $P_T^*$ . Luego, la interacción estratégica no cooperativa en la subasta genera un sobreprecio o poder de mercado, por cuanto en equilibrio  $P_H^* > P^S = P_T^*$ . En consecuencia, ambas empresas generadoras obtienen utilidades esperadas positivas, siendo mayor la utilidad esperada del generador hidráulico según lo vemos de inmediato:

$$(5) \quad \begin{aligned} 0 < \Pi_T^E\left(P_H^*\right) &= \alpha \cdot u\left(\left(P_H^* - c_H\right)\bar{f}\right) + (1 - \alpha) \cdot u\left(\left(P_H^* - c_T\right)\bar{f}\right) \\ &< \alpha \cdot u\left(\left(P_H^* - c_H\right)\bar{f}\right) + (1 - \alpha) \cdot u\left(\left(P_H^* - c_T\right)\bar{f}\right) + \Delta c \cdot k_{1-\alpha} \\ &= (1 - \alpha) \cdot u\left(\Delta c \cdot k_{1-\alpha}\right) = \Pi_H^E\left(P_H^*\right) \end{aligned}$$

Lamentablemente, como discutimos luego de la siguiente proposición, no es posible regular  $P^S$  con la sola información de costos de las empresas. Ello por cuanto la empresa térmica obtendría pérdidas si en lugar de subastar el precio de la energía a clientes regulados se tarificara éste a precio nudo,  $P_N$ . La Proposición 4 resume los resultados hasta acá encontrados.

<sup>11</sup> Hay que destacar que en este modelo no hay inversiones que deban pagarse con los ingresos por venta en el mercado *spot* o los contratos. Tal supuesto es sin pérdida de generalidad, pues en la práctica los costos hundidos de las empresas generadoras son pagados por el pago a la potencia.

**Proposición 4:** Sean ambas empresas generadoras igualmente aversas al riesgo; sea  $P_N$  el precio nudo regulado por un regulador benevolente con información simétrica respecto de costos; sea  $P^S$  el precio que maximiza el bienestar neto de los consumidores y que permite a ambas generadoras obtener utilidades esperadas no negativas. Entonces,  $P_T^* > P_N$  y  $P_H^* > P^S = P_T^*$ .

El primer resultado de esta proposición nos dice que el generador térmico, por tener costos mayores, cubrirá exactamente sus costos esperados de abastecer a los clientes regulados en caso que el regulador tarificara estos precios a costo marginal esperado o precio nudo. Sin embargo, tal tarifa conllevaría una desutilidad esperada para este generador averso al riesgo, siéndole más conveniente no participar en el abastecimiento de los clientes regulados. La causa de fondo es que al precio nudo, esta empresa obtiene rentas esperadas iguales a cero de participar en la subasta, pero tal decisión le lleva a tomar riesgos que no enfrenta de no contratarse para abastecer a la empresa distribuidora.

Este resultado nos indica que la modificación legal en Chile, que liberaliza el precio que pagan los clientes regulados por su energía consumida, avanza en la dirección correcta al permitir que el precio finalmente pagado esté por encima del precio nudo determinado por el regulador bien informado.

El segundo resultado de la Proposición 4 nos plantea un dilema de política pública. Existe un precio óptimo que permite maximizar el bienestar neto de los consumidores, definido como el menor precio que deja a ambas empresas al menos tan bien como si participaran en la subasta. A dicho precio, la firma hidráulica obtiene una utilidad esperada positiva, renta que se deriva de su ventaja absoluta de costos. Sin embargo, como este precio depende del grado de aversión al riesgo de los dueños de las empresas generadoras, parámetro que suponemos desconocido por el regulador, entonces no es posible implementar esta solución ni a través de la tarificación directa ni tampoco a través de regular el monto máximo de la subasta. En otras palabras, si bien es teóricamente consistente imponer un techo por encima del precio nudo determinado por el regulador, tal techo puede ser tan alto como para ser simplemente un punto focal de posible colusión; o, puede ser tan bajo como para inhibir la participación de uno o ambos generadores en la subasta por abastecer de energía a los clientes regulados.

Si suponemos que el techo máximo es definitivamente no restrictivo (suficientemente alto), el gobierno enfrenta un *trade off* de política pública. Lo bueno es que comparado al mecanismo de fijación tarifaria por precio nudo, la subasta genera incentivos a invertir en el sector al permitírsele a las empresas generadoras vender a clientes regulados por sobre el precio nudo regulado que resulta de costos. Lo malo es que el regulador queda amarrado de manos y no puede evitar que las empresas extraigan rentas a los consumidores por encima de lo socialmente deseable. Más aún, esta liberalización en la industria eléctrica genera un enorme desafío en términos de política de competencia, por cuanto no es posible detectar en la práctica si el exceso de rentas es causado por la aversión al riesgo de los empresarios o por colusión en la subasta.

Finalmente, el Corolario 2 nos indica que si las empresas generadoras fuesen neutrales al riesgo, el único equilibrio de Nash que resulta de la subasta no colusiva es igual a este precio que maximiza el bienestar neto de los consumidores. Sólo en este caso particular, que además supone un regulador con información simétrica de costos, el precio que maximiza el bienestar neto de los consumidores es implementable, ya sea a través del precio nudo regulado o a través de la subasta competitiva. Los cuidados deben ser en el primer caso evitar la manipulación de dicho precio por parte del regulador y en el segundo caso fiscalizar adecuadamente que las empresas no se coludan en la subasta.

**Corolario 2:** *Si las firmas fueran neutrales al riesgo, entonces el único equilibrio de Nash no cooperativo de la subasta es igual al precio que maximiza el bienestar neto de los consumidores e igual al precio nudo regulado,  $P_H^* = P_T^* = P^S = P_N$ .*

Un último comentario respecto de estos resultados. En este modelo la aversión al riesgo pasa a ser un mecanismo de compromiso que les permite a las empresas generadoras a no pujar a la baja en la subasta. Cuando estas empresas son neutrales al riesgo, y eso es de conocimiento común, desaparece este mecanismo de compromiso y las empresas son llevadas por las fuerzas de la competencia hacia una situación en que la más ineficiente no obtiene ni beneficios ni utilidades esperadas de participar en la subasta.

Es golpeante e inquietante que no sea óptimo para estas empresas diversificar su riesgo, aun cuando pertenecieran a grupos económicos diversificados. Es un resultado golpeante en tanto mayor riesgo es preferido a menor riesgo (en un rango relevante), y la razón es que mayor riesgo genera un mecanismo de compromiso que sostiene un equilibrio de Nash que deja mayores beneficios. Es un resultado inquietante en tanto la subasta competitiva incentiva a no diversificar riesgos, lo cual supone una pérdida de eficiencia para la sociedad<sup>12</sup>.

## 5. RACIONAMIENTO ELÉCTRICO Y EQUILIBRIO

Una variante interesante de considerar es qué ocurre con los resultados de la subasta si se supone que la sequía es extrema al punto de producir desabastecimiento o racionamiento eléctrico. Dado esto, es pertinente considerar el alcan-

<sup>12</sup> En un contexto dinámico, el hecho de que la aversión al riesgo lleve a ablandar la competencia en la subasta incentiva al propietario precisamente a diversificar menos el riesgo de sus negocios. En otras palabras, este efecto podría llevar a un portafolio de inversiones menos diversificado como un mecanismo de compromiso para una etapa posterior en la subasta. Antecedentes de este tipo de juegos pueden encontrarse en Fudenberg y Tirole (1984). Ciertamente un análisis más profundo de este hecho requeriría endogeneizar la decisión de cuán averso al riesgo se quiere ser (diversificación del portafolio de inversiones del grupo controlador de cada generadora). Agradecemos a Juan Pablo Montero este comentario.

ce de la responsabilidad contractual de las generadoras deficitarias para compensar a los usuarios por la energía no servida. Supondremos dos escenarios polares. En el primero, las generadoras deficitarias, sean térmicas o hidráulicas, quedan exentas de pagar compensaciones a los usuarios regulados en el caso de racionamiento eléctrico producto de una sequía extrema. En el segundo caso, las empresas generadoras deben compensar a los usuarios desconectados por cada unidad de energía no entregada al precio de costo de falla del sistema, sin importar la causa del racionamiento.

Supongamos que la capacidad total de generación de ambas firmas en el estado de sequía es menor que la demanda licitada por la distribuidora,  $k^T + k_{1-\alpha} < f$ . Recuérdese que el desabastecimiento de energía eleva el costo de oportunidad del agua en ese estado de la naturaleza, por encima del costo de operación de la generadora térmica, cumpliéndose que  $c_H^F > c_T > c_H$ .

*Caso i) Generadoras deficitarias no compensan a usuarios racionados.*

Las funciones de pagos o utilidad esperada de las firmas hidráulica y térmica, respectivamente son:

$$(6) \quad \Pi_H^E(b_H) = \begin{cases} \alpha \cdot u(b_H \bar{f} + c_H(k_\alpha - \bar{f}) - c_H k_\alpha) + \\ (1 - \alpha) \cdot u(b_H(k_{1-\alpha} + k_T) - c_H^F(k_{1-\alpha} + k_T)) & \text{si } b_H \leq b_T \\ \alpha \cdot u(c_H k_\alpha - c_H k_\alpha) + \\ (1 - \alpha) \cdot u(c_H^F k_{1-\alpha} - c_H^F k_{1-\alpha}) & \text{si } b_H > b_T \end{cases}$$

$$(7) \quad \Pi_T^E(b_T) = \begin{cases} \alpha \cdot u(b_T \bar{f} - c_H \bar{f}) + \\ (1 - \alpha) \cdot u(b_T(k_{1-\alpha} + k_T) - c_H^F k_{1-\alpha} - c_T k_T) & \text{si } b_T < b_H \\ \alpha \cdot u(0) + (1 - \alpha) \cdot u(c_H^F k_T - c_T k_T) & \text{si } b_T \geq b_H \end{cases}$$

Comparando (6) con (1) se observa un cambio importante en la utilidad esperada del generador hidráulico, en particular en caso de perder la subasta<sup>13</sup>. En este caso, la firma hidráulica no obtiene beneficios, pues marca el costo mar-

<sup>13</sup> Es necesario hacer notar que se ha modificado también la regla de adjudicación de la subasta en caso de que ambas firmas ofrezcan el mismo precio, siendo esta vez la generadora hidráulica y no la térmica quien se la adjudicaría. Este supuesto se requiere para encontrar un único equilibrio no cooperativo que elimine equilibrios con estrategias débilmente dominadas.



ginal del sistema o precio *spot* en ambos estados de la naturaleza por cuanto su costo marginal en un estado de sequía extrema es el costo de falla del sistema. En cuanto al generador térmico, se observa al comparar (7) con (2) que esta firma obtiene ahora pagos esperados positivos si no se adjudica la subasta. El generador térmico es el primero llamado a producir en el caso de darse el estado de sequía extrema. Definiendo  $P'_i$  con  $i = \{H, T\}$ , como el precio que iguala las utilidades esperadas para la firma  $i$  de ganar la subasta con las utilidades esperadas de perderla, se obtiene el siguiente resultado:

**Proposición 5:** Si ambas empresas son igualmente aversas al riesgo, el costo de falla del sistema es tal que  $c_H < c_T < c_H^F$  y no hay pago de compensaciones a usuarios no abastecidos, entonces i) los precios que dejan indiferentes a ambas firmas de participar o no en la subasta cumplen con  $P'_H < P'_T$ ; ii) hay un conjunto de precios que son equilibrios de Nash y todos ellos cumplen con  $P_H = P_T$  y  $(P_H, P_T) \in [P'_H, P'_T]$ ; iii) el único par de estrategias que débilmente domina a los otros equilibrios de Nash y, por lo tanto, resiste el refinamiento que le impone el criterio de la “mano temblorosa” es  $(P_H, P_T) = (P'_T, P'_T)$ , de forma que es la empresa hidráulica el generador que se adjudica la subasta.

La Proposición 5 confirma los resultados anteriores. Esto es, la firma que obtiene beneficios positivos al no adjudicarse la subasta tiene menores incentivos a ofrecer un precio menor que la otra firma; así, es la empresa que margina en el estado seco de la naturaleza la que se adjudica la subasta. Todos los resultados ya analizados se mantienen, en particular esos relativos al poder de mercado que estas firmas obtienen como producto de su interacción estratégica cuando son aversos al riesgo y al hecho de que no es posible implementar el óptimo para los consumidores con la sola información de costos de las empresas.

*Caso ii) Generadoras deficitarias compensan a usuarios no servidos, a todo evento*

En este caso las generadoras deben, frente a cualquier estado e hidrología, compensar a los usuarios por cada unidad de energía no entregada al precio del costo de falla. Con esto, las funciones de pagos de las firmas generadoras son ahora:

$$(8) \quad \Pi_H^E(b_H) = \begin{cases} \alpha \cdot u(b_H \bar{f} + c_H(k_\alpha - \bar{f}) - c_H k_\alpha) + \\ (1 - \alpha) \cdot u(b_H(k_{1-\alpha} + k_T) - c_H^F(\bar{f} - (k_{1-\alpha} + k_T)) - c_H^F(k_{1-\alpha} + k_T)) & \text{si } b_H \leq b_T \\ \alpha \cdot u(c_H k_\alpha - c_H k_\alpha) + \\ (1 - \alpha) \cdot u(c_H^F k_{1-\alpha} - c_H^F k_{1-\alpha}) & \text{si } b_H > b_T \end{cases}$$

$$(9) \quad \Pi_T^E(b_T) = \begin{cases} \alpha \cdot u(b_T \bar{f} - c_H \bar{f}) + \\ + (1 - \alpha) \cdot u(b_T (k_{1-\alpha} + k_T) - c_H^F (\bar{f} - (k_{1-\alpha} + k_T)) - c_H^F k_{1-\alpha} - c_T k_T) & \text{si } b_T < b_H \\ \alpha \cdot u(0) + (1 - \alpha) \cdot u(c_H^F k_T - c_T k_T) & \text{si } b_T \geq b_H \end{cases}$$

Comparando (8) con (1) y (6) se observa una situación muy parecida al caso de racionamiento previo para el operador hidráulico, y muy diferente al caso cuando no hay falla de abastecimiento en el sistema. Lo mismo ocurre para el generador térmico si se comparan (9) con (2) y (7). Luego, no es la obligación de pagar compensaciones, lo que marca la diferencia en cuanto a qué tipo de tecnología abastecerá a clientes regulados (qué firma ganará la subasta), sino que lo es simplemente el qué firma es la que margina en el estado seco de la naturaleza. Una diferencia adicional, cuantitativa solamente, es que el precio que deja indiferente a cada generador entre ganar o perder la subasta es mayor ahora que en el caso en que no hay obligaciones contractuales en caso de racionamiento para el generador contratado, pues hay una pérdida esperada extra a causa de las compensaciones que se deben pagar a los usuarios en este caso:  $c_H^F \cdot (\bar{f} - (k_{1-\alpha} + k_T))$ . Si definimos como  $P_i^n$  con  $i = \{H, T\}$  a ese precio, entonces  $P_i^n > P_i^i$ .

Por la similitud al caso previo no se analiza esta situación en mayor detalle, quedando su equilibrio completamente caracterizado en la siguiente proposición:

**Proposición 6:** *Si ambas empresas son igualmente aversas al riesgo, el costo de falla del sistema es tal que  $c_H < c_T < c_H^F$  y el generador contratado compensa a los usuarios no abastecidos, entonces i) los precios que dejan indiferentes a ambas firmas de participar o no en la subasta cumplen con  $P_H^n < P_T^n$ ; ii) hay un conjunto de precios que son equilibrios de Nash que cumplen con  $P_H = P_T$  y  $(P_H, P_T) \in [P_H^n, P_T^n]$ ; iii) el único par de estrategias que débilmente domina a los otros equilibrios de Nash y, por lo tanto, resiste el refinamiento que le impone el criterio de la “mano temblorosa” es  $(P_H, P_T) = (P_T^n, P_T^n)$ , de forma que es la empresa hidráulica el generador que se adjudica la subasta.*

Como se aprecia de las Proposiciones 5 y 6, cualquiera sea el tratamiento a las compensaciones a usuarios por energía no abastecida, cuando la sequía lleva a racionamiento eléctrico será la firma hidráulica la que se adjudique la subasta por abastecer a clientes regulados. En consecuencia, el factor que finalmente determina qué empresa generadora abastecerá a los clientes regulados es el grado de profundidad que se espere de la sequía en este estado de la naturaleza, cuando el grado de aversión al riesgo es el mismo para ambos tipos de generadores.

## 6. CONCLUSIONES

Este artículo realiza un análisis teórico del precio de equilibrio que resultaría en una subasta por energía cuando la información entre oferentes es completa, la oferta con una de las tecnologías usadas es estocástica y los inversionistas son aversos al riesgo. La oferta incierta está dada por dos estados posibles de hidrología, uno en que sólo la firma hidráulica produce (normal o lluvioso) y otro en que ambas firmas son llamadas a generar y abastecen la demanda licitada (sequía). Este trabajo está motivado en la reciente modificación a la ley eléctrica en Chile que liberaliza el precio nudo que pagan los clientes regulados, precisamente por la vía de subastar este precio entre empresas generadoras. Se supone que una única distribuidora licita por una única vez su demanda por energía en una subasta cuya variable de decisión es el precio que pagarán los clientes regulados, en donde los oferentes son sólo una empresa térmica y otra hidroeléctrica.

El resultado más importante cuando la sequía no genera desabastecimiento es que la firma térmica siempre se llevará los contratos a un precio inferior a su costo operacional, pero mayor al precio nudo calculado por un regulador bien informado. Se encuentra que hay un único equilibrio de Nash que no contiene estrategias débilmente dominadas, precisamente aquel que deja indiferente al generador hidráulico entre contratar con la distribuidora o vender toda su producción en el mercado *spot*. En este caso de sequía no extrema, ambas firmas obtienen beneficios esperados positivos que a su vez conllevan además utilidades esperadas positivas, las que no se disipan como producto de la subasta no colusiva. Así, la subasta genera precios que son mayores al que maximiza el bienestar neto de los consumidores, aunque el artículo discute acerca de la dificultad práctica de implementar este precio teórico.

La explicación a este resultado de “competencia imperfecta” se explica porque la aversión al riesgo suaviza la competencia, permitiendo a las empresas obtener beneficios esperados positivos. Si bien en este trabajo la aversión al riesgo es exógena, se discute en el texto la posibilidad de que el grupo controlador de cada empresa generadora podría ver desincentivado su interés por diversificar el riesgo de sus inversiones como un mecanismo de compromiso para sostener estas rentas extranormales en la etapa de la subasta. Asimismo, a mayor aversión al riesgo más suave es la competencia en la subasta. Este tipo de comportamiento es equivalente al conocido efecto de diferenciación de productos en mercados oligopólicos.

Cuando se supone que las empresas generadoras son neutrales al riesgo, los resultados cambian muy ligeramente en cuanto a la unicidad del equilibrio, pero cambian fuertemente en cuanto a que se eliminan las ganancias extranormales que obtienen las empresas generadoras como producto de la liberalización del mercado. Así, bajo neutralidad al riesgo se encuentra que estas firmas ofrecen el mismo precio y dicha oferta es igual al precio nudo fijado por el regulador, siempre que el regulador conozca los costos de las empresas, por lo que es indiferente que las distribuidoras subasten sus necesidades de energía o que el re-

gulador fije el precio de transferencia de energía entre generadores y distribuidores a dicho precio. Este nuevo precio de equilibrio disipa las rentas y maximiza el excedente neto de los consumidores. Hay que notar, no obstante, que esta situación se obtiene como producto de que se ha modificado el supuesto de aversión al riesgo. En consecuencia, el supuesto de aversión al riesgo es central para determinar hasta dónde las empresas son capaces de extraer rentas de los consumidores sin entrar a suponer colusión en la subasta.

Cuando se considera que el estado de sequía lleva a desabastecimiento, con lo cual el precio *spot* alcanza el costo de falla del sistema, la subasta se la adjudica la firma hidráulica. Este resultado no representa un cambio de fondo ya que nuevamente se adjudica la subasta aquella firma que marca el costo marginal del sistema en el estado seco de la naturaleza; es decir, aquella que tiene los costos operacionales mayores y tiene, por lo tanto, incentivos a estar contratada cuando se produce el racionamiento eléctrico. Se mantienen los resultados en términos de poder de mercado y bienestar de los consumidores. Estos resultados además no cambian cualquiera sea la responsabilidad contractual de la generadora contratada con los usuarios no abastecidos.

Finalmente, hay una serie de posibles extensiones de este trabajo. Por ejemplo, dentro del mismo análisis de competencia en la subasta e información simétrica entre firmas, se puede suponer que las empresas generadoras ofrecen cantidad (bloques de energía), reflejado por el porcentaje de la demanda que desean cubrir, además del precio a pagar por esas unidades; o, elevar a tres estados de la naturaleza para fundir el análisis bajo sequía no extrema con ese bajo desabastecimiento; o, analizar las consecuencias de permitir la entrada en la industria, ya sea de nuevos actores, tecnologías o de las mismas empresas ya establecidas. Se puede además estudiar el caso de colusión en la subasta, para lo cual debiera suponerse que se subastan secuencialmente bloques de energía por tiempo indefinido. Como recomendación de política pública, las tres extensiones anteriores son necesarias de estudiar para darle plena validez práctica a los resultados teóricos encontrados en este trabajo. Por último, abandonar el supuesto de información completa entre las firmas es ciertamente más interesante y realista, de forma de incorporar las creencias que tienen las empresas en la subasta respecto de parámetros de decisión de su rival; así como analizar los juegos de señales que aparecerían bajo esta modelación más compleja, en caso que se tratara de subastas secuenciales de bloques de energía.

## REFERENCIAS

- Arellano, S. (2004), "Market Power in Mixed Hydro-Thermal Electric", *Documento de Trabajo* N° 187, Centro de Economía Aplicada, Universidad de Chile.
- Basañes, F., E. Saavedra y R. Soto (2001), "Post-privatization, Renegotiations and Disputes in Chile", en *Second Generation Reforms in Infrastructure Services*, F. Basañes y R. Willig (Ed.), Banco Inter-Americano de Desarrollo.

- Benavente, J., A. Galetovic, R. Sanhueza y P. Serra (2005), "Estimando la Demanda Residencial por Electricidad en Chile: El Consumidor es Sensible al Precio", *Cuadernos de Economía* 42 (mayo): 31-61.
- Comisión Nacional de Energía de Chile (2005), "Estadísticas Sector Energía Chile: 1990-2004", en <http://www.cne.cl/estadisticas/anuario/index.html> (15/12/2006).
- Díaz, C., A. Galetovic y R. Soto (2001), "Anatomía de una Crisis Eléctrica", *Revista de Análisis Económico* 16(1): 3-57.
- Fabra, N., N. Von der Fehr y D. Harbord (2004), "Designing Electricity Auctions", *Working Paper Center for Study of Energy Markets (CSEM)*. University of California Energy Institute.
- Fudenberg, D. y J. Tirole (1984), "The Fat Cat Effect, The Puppy Dog Ploy, and the Lean and Hungry Look", *American Economic Review, Papers and Proceedings* 74: 361-368.
- Galetovic, A., J.C. Olmedo y H. Soto (2002), "Una Estimación del Costo Social de Eliminar los Déficit de Abastecimiento Eléctrico en el SIC", *Revista de Análisis Económico* 17(2): 3-30.
- Galetovic, A., J.C. Olmedo y H. Soto (2002a), "¿Qué Tan Probable es una Crisis Eléctrica?", *Estudios Públicos* 87: 175-212.
- Green, R. y D. Newbery (1992), "Competition in the British Electricity Spot Market", *Journal of Political Economy* 100: 929-953.
- Klemperer, P. y M. Meyer (1989), "Supply Function Equilibria in Oligopoly under Uncertainty", *Econometrica* 57: 1243-1277.
- Mas-Colell, A., M. Whinston y J. Green (1995), *Microeconomic Theory*, Oxford University Press. Nueva York NY.
- Newbery, D. (1998), "Competition, Contracts, and Entry in the Electricity Spot Market", *The RAND Journal of Economics* 29: 726-749.
- Selten, R. (1975), "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games", *International Journal of Game Theory* 4: 25-55.
- Serra, P. (1997), "Energy Pricing Under Uncertain Supply", *Energy Economics* 19: 417-434.
- Stacchetti, E. (1999), "Auction Design for the Colombian Electricity Market", *Technical Report 72*, Economics Series Working Paper, Centro de Economía Aplicada, Universidad de Chile.
- Von der Fehr, N. y D. Harbord (1993), "Spot Market Competition in the UK Electricity Industry", *The Economic Journal* 103: 531-546.

## ANEXO

**Demostración Proposición 1:** Tomemos en principio el supuesto de que la distribuidora hace una demanda ilimitada de energía,  $\bar{f} \rightarrow \infty$ . Así, teniendo en cuenta que el precio nudo corresponde al costo esperado del sistema,  $P_N = \alpha c_H + (1 - \alpha) c_T$ , entonces la firma hidráulica resuelve:

$$(A1) \quad \begin{aligned} \prod_H^E(f^H) &= \alpha \cdot u((P_N - c_H) f^H) + (1 - \alpha) \cdot u((P_N - c_T) f^H + (c_T - c_H) \cdot k_{1-\alpha}) \\ &= \alpha \cdot u((1 - \alpha) \cdot \Delta c \cdot f^H) + (1 - \alpha) \cdot u(-\alpha \cdot \Delta c \cdot f^H + \Delta c \cdot k_{1-\alpha}) \end{aligned}$$

Como la función de utilidad esperada es estrictamente cóncava, las condiciones de primer orden del problema son suficientes; éstas son:

$$(A2) \quad \alpha \cdot (1 - \alpha) \cdot \Delta c \cdot u'((1 - \alpha) \cdot \Delta c \cdot f^H) - \alpha \cdot \Delta c \cdot (1 - \alpha) \cdot u'(-\alpha \cdot \Delta c \cdot f^H + \Delta c \cdot k_{1-\alpha}) = 0$$

o, simplificando:

$$(A3) \quad u'((1 - \alpha) \cdot \Delta c \cdot f^H) = u'(-\alpha \cdot \Delta c \cdot f^H + \Delta c \cdot k_{1-\alpha})$$

Como la función de utilidad es estrictamente cóncava,  $u'(x)$  es creciente para todo  $x > 0$ , entonces sólo un único argumento de esta función puede cumplir con (A3). Así,

$$(A4) \quad (1 - \alpha) \cdot \Delta c \cdot f^H = -\alpha \cdot \Delta c \cdot f^H + \Delta c \cdot k_{1-\alpha}$$

Reordenando, simplificando y resolviendo para  $f^H$  se determina la elección óptima de la firma hidráulica.

$$(A5) \quad f^H = k_{1-\alpha}$$

Por último, el supuesto de demanda ilimitada de la distribuidora no es restrictivo en tanto se ha supuesto que la generación hidráulica no abastece a todos los clientes regulados cuando se está en el estado seco de la naturaleza, esto es  $k_{1-\alpha} < \bar{f}$ . Esto completa la primera parte de la demostración.

Por otro lado, la firma térmica resuelve:

$$(A6) \quad \begin{aligned} \prod_T^E(f^T) &= \alpha \cdot u((P_N - c_H) f^T) + (1 - \alpha) \cdot u((P_N - c_T) f^T) \\ &= \alpha \cdot u((1 - \alpha) \cdot \Delta c \cdot f^T) + (1 - \alpha) \cdot u(-\alpha \cdot \Delta c \cdot f^T) \end{aligned}$$

La condición de primer orden de este problema es suficiente e igual a:

$$(A7) \quad \alpha \cdot (1-\alpha) \cdot \Delta c \cdot u'((1-\alpha) \cdot \Delta c \cdot f^T) - \alpha \cdot \Delta c \cdot (1-\alpha) \cdot u'(-\alpha \cdot \Delta c \cdot f^T) = 0$$

con lo que simplificando:

$$(A8) \quad u'((1-\alpha) \cdot \Delta c \cdot f^T) = u'(-\alpha \cdot \Delta c \cdot f^T)$$

La única posibilidad de que se cumpla (A8), cuando  $u''(x)$  es negativa, es que  $u'(x) = 0$ . Como hemos supuesto que esta función cumple con las condiciones de Inada, entonces es cierto que ello sólo ocurre si  $f^T$  tiende a infinito, pues:

$$(A9) \quad \lim_{f^T \rightarrow \infty} u'((1-\alpha) \cdot \Delta c \cdot f^T) = 0 \quad \text{y} \quad \lim_{f^T \rightarrow \infty} u'(-\alpha \cdot \Delta c \cdot f^T) = 0$$

Como se sabe, la demanda de la empresa distribuidora está acotada a  $\bar{f}$ . Luego, como la función objetivo de la firma térmica es quasicóncava, mientras más cerca de satisfacer la condición de primer orden más pagos esperados, entonces la firma térmica deseará tomar toda la demanda de los clientes regulados, esto es  $f^T = \bar{f}$ . ■

**Demostración Lema 1:** Por definición de  $P_H^*$  y  $P_T^*$  se debe cumplir para la firma hidráulica y térmica, respectivamente, que:

$$(A10) \quad \alpha u((P_H^* - c_H) \bar{f}) + (1-\alpha) u((P_H^* - c_T) \bar{f} + k_{1-\alpha} \Delta c) = (1-\alpha) u(k_{1-\alpha} \Delta c)$$

donde  $\Delta c = c_T - c_H$

$$(A11) \quad \alpha u((P_T^* - c_H) \bar{f}) + (1-\alpha) u((P_T^* - c_T) \bar{f}) = 0$$

Contrario a lo que se busca demostrar, supóngase que  $P_T^* > P_H^*$ . Evaluando  $P_T^*$  en (A10) se tiene:

$$(A12) \quad \alpha u((P_T^* - c_H) \bar{f}) + (1-\alpha) u((P_T^* - c_T) \bar{f} + k_{1-\alpha} \Delta c) > (1-\alpha) u(k_{1-\alpha} \Delta c)$$

Usando (A11) en (A12):

$$(A13) \quad -(1-\alpha) u((P_T^* - c_T) \bar{f}) + (1-\alpha) u((P_T^* - c_T) \bar{f} + k_{1-\alpha} \Delta c) > (1-\alpha) u(k_{1-\alpha} \Delta c)$$

Reordenando y simplificando:

$$(A14) \quad u((P_T^* - c_T) \bar{f} + k_{1-\alpha} \Delta c) > u(k_{1-\alpha} \Delta c) + u((P_T^* - c_T) \bar{f})$$

resultado que es contrario al supuesto de concavidad estricta de  $u(x)$ , pues en ese caso se debe cumplir que  $u\left(\left(P_T^* - c_T\right)\bar{f} + k_{1-\alpha}\Delta c\right) < u\left(k_{1-\alpha}\Delta c\right) + u\left(\left(P_T^* - c_T\right)\bar{f}\right)$ . Luego,  $P_H^* > P_T^*$ . ■

**Demostración Proposición 2:** De las correspondencias de reacción de las firmas generadoras, se observa que el único intervalo de estrategias en donde ninguna de las firmas quiere desviarse unilateralmente es entre  $[P_H^*, P_T^*]$ , que es precisamente donde las correspondencias de reacción se superponen. Este intervalo para cada jugador determina el conjunto de equilibrios de Nash para las firmas generadoras, correspondiendo al par  $(P_H, P_T) \in [P_H^*, P_T^*]$  para la firma hidráulica y térmica, respectivamente. ■

**Demostración Corolario 1:** Dado el supuesto que las firmas son neutrales al riesgo, entonces de la definición de  $P_H^*$  y  $P_T^*$  se debe cumplir para la firma hidráulica que:

$$(A15) \quad \alpha\left(P_H^* - c_H\right)f + (1-\alpha)\left(\left(P_H^* - c_T\right)f + \Delta c \cdot k_{1-\alpha}\right) = (1-\alpha)\Delta c \cdot k_{1-\alpha}$$

y para la firma térmica:

$$(A16) \quad \alpha\left(P_T^* - c_H\right)f + (1-\alpha)\left(P_T^* - c_T\right)f = 0$$

Considerando la definición del precio nudo como  $P_N = \alpha c_H + (1-\alpha)c_T$ , con un poco de álgebra en cada una de las ecuaciones anteriores se encuentra que  $P_H^* = P_T^* = P_N$ , demostrando el corolario. ■

**Demostración Proposición 3:** Por notación y sin pérdida de generalidad supondremos que existen  $n+1$  precios en el set o conjunto de equilibrios de Nash encontrado en la Proposición 1, donde los límites son  $P_T^* = P_0$  y  $P_H^* = P_n$  para ambas firmas por igual. Sean además los vectores  $\sigma_T^\varepsilon = (\beta_0, \beta_1, \dots, \beta_n)$  y  $\sigma_H^\varepsilon = (\gamma_0, \gamma_1, \dots, \gamma_n)$  una secuencia completa de estrategias mixtas perturbadas para ambas firmas, en donde a cada creencia se le asocian los precios ofertados  $(P_0^{*T}, \dots, P_n^{*T})$  y  $(P_0^{*H}, \dots, P_n^{*H})$ , respectivamente. Luego,  $\sigma^*$  es un equilibrio de Nash que resiste el refinamiento de la “mano temblorosa” si existe una secuencia completa de estrategias mixtas  $(\sigma^\varepsilon)$  tal que se cumplan las siguientes condiciones:

$$\begin{aligned} & -(\sigma^\varepsilon) \text{ converge a } \sigma^* \\ & -\sigma_i^* \in BR_i(\sigma_{-i}^\varepsilon) \quad \forall i \end{aligned}$$



con  $BR_i$  siendo la correspondencia de mejor respuesta de la firma  $i$  a una acción de la otra firma. Como se ha supuesto que en caso de que las ofertas sean iguales la subasta se le adjudica a la empresa que margina en el período de sequía, que en este caso es la firma térmica, entonces los pagos esperados para esta empresa dependerán de si su precio ofrecido es mayor, igual o menor al de su rival:

$$(A17) \quad \Pi_T^E(P_j^{*T}) = \begin{cases} U^T(P_j^{*T}) \equiv \alpha \cdot u([P_j^{*T} - c_H] \cdot \bar{f}) + (1-\alpha) \cdot u([P_j^{*T} - c_T] \cdot \bar{f}) & \text{si } P_j^{*T} \leq P_j^{*H} \\ 0 & \text{si } P_j^{*T} > P_j^{*H} \end{cases}$$

para  $j, l = 0, 1, \dots, n$ . Es posible observar que  $U^T(\cdot)$  es creciente en  $P_j^{*T}$ , en tanto esta oferta sea menor o igual que la oferta de la firma hidráulica. Asimismo, los pagos esperados para la firma hidráulica también dependerán de cómo es su oferta relativa a la oferta de la firma térmica:

$$(A18) \quad \Pi_H^E(P_j^{*H}) = \begin{cases} A \equiv (1-\alpha) \cdot u(\Delta c \cdot k_{1-\alpha}) & \text{si } P_j^{*H} \geq P_j^{*T} \\ U^H(P_j^{*H}) \equiv \alpha \cdot u([P_j^{*H} - c_H] \cdot \bar{f}) + (1-\alpha) \cdot u([P_j^{*H} - c_T] \cdot \bar{f} + \Delta c \cdot k_{1-\alpha}) & \text{si } P_j^{*H} < P_j^{*T} \end{cases}$$

Luego, la matriz de pagos del set de equilibrios de Nash, para cada posible par de estrategias o precios ofertados elegidos por estas empresas, es la siguiente<sup>14</sup>:

CUADRO A.1.  
MATRIZ DE PAGOS

	$\gamma_0$	$\gamma_1$	$\dots$	$\gamma_{v-1}$	$1-\gamma_0-\dots-\gamma_{v-1}$
$\beta_0$	$[U^T(P_0^{*T}), A]$	$[U^T(P_0^{*T}), A]$		$\dots$	$[U^T(P_0^{*T}), A]$
$\beta_1$	$[0, U^H(P_0^{*H})]$	$[U^T(P_1^{*T}), A]$		$\dots$	$[U^T(P_1^{*T}), A]$
$\cdot$	$\cdot$	$\cdot$	$\cdot$		$\cdot$
$\cdot$	$\cdot$	$\cdot$	$\cdot$		$\cdot$
$\cdot$	$\cdot$	$\cdot$	$\cdot$		$\cdot$
$\beta_{v-1}$	$\cdot$	$\cdot$		$[U^T(P_{n-1}^{*T}), A]$	$[U^T(P_{n-1}^{*T}), A]$
	$\cdot$	$\cdot$			$\cdot$
$1-\beta_0-\dots-\beta_{v-1}$	$[0, U^H(P_0^{*H})]$	$[0, U^H(P_1^{*H})]$	$\dots$	$[0, U^H(P_{n-1}^{*H})]$	$[U^T(P_n^{*T}), A]$

<sup>14</sup> Cada par de estrategias mixtas  $(\beta_j, \gamma_l)$  corresponde al par de precios ofertados  $(P_j^{*T}, P_l^{*H})$ , para  $(j, l) = (0, 1, \dots, n-1)$ .

Dadas las conjeturas que cada firma hace de qué ofertará su rival, para la firma térmica jugar cada posible estrategia de precios le significa recibir una utilidad esperada diferente:

$$(A19) \quad \begin{aligned} \Pi_T^E(P_0^{*T}) &= \gamma_0 \cdot U^T(P_0^{*T}) + \gamma_1 \cdot U^T(P_0^{*T}) + \dots + \gamma_{n-1} \cdot U^T(P_0^{*T}) + (1 - \gamma_0 - \dots - \gamma_{n-1}) \cdot U^T(P_0^{*T}) \\ &= U^T(P_0^{*T}) \end{aligned}$$

$$(A20) \quad \begin{aligned} \Pi_T^E(P_1^{*T}) &= \gamma_0 \cdot 0 + \gamma_1 \cdot U^T(P_1^{*T}) + \dots + \gamma_{n-1} \cdot U^T(P_1^{*T}) + (1 - \gamma_0 - \dots - \gamma_{n-1}) \cdot U^T(P_1^{*T}) \\ &= (1 - \gamma_0) U^T(P_1^{*T}) \\ &\vdots \end{aligned}$$

$$(A21) \quad \begin{aligned} \Pi_T^E(P_j^{*T}) &= \gamma_0 \cdot 0 + \gamma_1 \cdot 0 + \dots + \gamma_j \cdot U^T(P_j^{*T}) + \dots + (1 - \gamma_0 - \dots - \gamma_{n-1}) \cdot U^T(P_j^{*T}) \\ &= (1 - \gamma_0 - \dots - \gamma_{j-1}) U^T(P_j^{*T}) \\ &\vdots \end{aligned}$$

$$(A22) \quad \begin{aligned} \Pi_T^E(P_n^{*T}) &= \gamma_0 \cdot 0 + \gamma_1 \cdot 0 + \dots + \gamma_{n-1} \cdot 0 + (1 - \gamma_0 - \dots - \gamma_{n-1}) \cdot U^T(P_n^{*T}) \\ &= (1 - \gamma_0 - \dots - \gamma_{n-1}) U^T(P_n^{*T}) \end{aligned}$$

lo que implica que si para cada posible  $j$ ,  $\gamma_j \rightarrow 0$  entonces  $\Pi_T^E(P_j^{*T}) = U^T(P_j^{*T})$ .

A su vez, como esta función es creciente en el precio ofertado, entonces  $P_{n+1}^T = P_H^*$  es la mejor respuesta de la firma térmica a cualquier estrategia posible de la firma hidráulica.

Siguiendo el mismo análisis, la firma hidráulica obtiene un pago esperado por jugar cada una de la estrategias iguales a:

$$(A23) \quad \begin{aligned} \Pi_H^E(P_0^{*H}) &= \beta_0 \cdot A + \beta_1 \cdot U^H(P_0^{*H}) + \dots + \beta_{n-1} \cdot U^H(P_0^{*H}) + (1 - \beta_0 - \dots - \beta_{n-1}) \cdot U^H(P_0^{*H}) \\ &= \beta_0 \cdot A + (1 - \beta_0) \cdot U^H(P_0^{*H}) \end{aligned}$$

$$(A24) \quad \begin{aligned} \Pi_H^E(P_1^{*H}) &= \beta_0 \cdot A + \beta_1 \cdot A + \beta_1 \cdot U^H(P_1^{*H}) + \dots + (1 - \beta_0 - \dots - \beta_{n-1}) \cdot U^H(P_1^{*H}) \\ &= (\beta_0 + \beta_1) \cdot A + (1 - \beta_0 - \beta_1) U^H(P_1^{*H}) \\ &\vdots \end{aligned}$$

$$(A25) \quad \begin{aligned} \Pi_H^E(P_j^{*H}) &= \beta_0 \cdot A + \dots + \beta_j \cdot A + \beta_{j+1} \cdot U^H(P_j^{*H}) + \dots + (1 - \beta_0 - \dots - \beta_{n-1}) \cdot U^H(P_j^{*H}) \\ &= \left( \sum_{m=1}^j \beta_j \right) \cdot A + \left( 1 - \sum_{m=1}^j \beta_j \right) U^H(P_j^{*H}) \\ &\vdots \end{aligned}$$

$$(A26) \quad \begin{aligned} \Pi_H^E(P_n^{*H}) &= \beta_0 \cdot A + \beta_1 \cdot A + \dots + \beta_{n-1} \cdot A + (1 - \beta_0 - \dots - \beta_{n-1}) \cdot A \\ &= A \end{aligned}$$

con lo que si para cualquier  $j$  se cumple que  $\beta_j \rightarrow 0$ , entonces se debe cumplir que  $\Pi_H^E(P_j^{*H}) = U^H(P_j^{*H})$  para todo  $j \neq n$  y  $\Pi_H^E(P_n^{*H}) = A$ . Adicionalmente, por definición de  $P_H^*$  –que es la oferta que deja indiferente a la empresa hidráulica entre ganarse o perder la subasta– entonces este mismo precio es también la mejor respuesta de generador.

Queda así demostrado que el par  $(P_H^*, P_H^*)$  es el único equilibrio de Nash que sobrevive el refinamiento de la “mano temblorosa” y, por lo tanto, las estrategias  $(P_H^*, P_H^*)$  débilmente dominan al resto de los equilibrios de Nash en estrategias puras. Finalmente, como a igual precio ofrecido la subasta es adjudicada a la empresa que margina en el estado de naturaleza seco, entonces la empresa térmica obtiene este contrato de abastecimiento. ■

**Demostración Proposición 4:** Recordando que el precio nudo regulado se define como  $P_N = \alpha \cdot c_H + (1 - \alpha) \cdot c_T$ . Luego, la utilidad esperada obtenida por la firma térmica con este precio es igual a:

$$\begin{aligned} \Pi_T^E(P_N) &= \alpha \cdot u\left(\left[\alpha \cdot c_H + (1 - \alpha) \cdot c_T\right] - c_H\right) \cdot \bar{f} + (1 - \alpha) \cdot u\left(\left[\alpha \cdot c_H + (1 - \alpha) \cdot c_T\right] - c_T\right) \cdot \bar{f} \\ (A27) \quad &= \alpha \cdot u\left((1 - \alpha) \cdot \Delta c \cdot \bar{f}\right) + (1 - \alpha) \cdot u\left(-\alpha \cdot \Delta c \cdot \bar{f}\right) \end{aligned}$$

Por la desigualdad de Jensen:

$$\begin{aligned} &< u\left(\alpha(1 - \alpha) \cdot \Delta c \cdot \bar{f} - \alpha(1 - \alpha) \cdot \Delta c \cdot \bar{f}\right) \\ (A28) \quad &= u(0) \\ &= \Pi_T^E(P_T^*) \end{aligned}$$

Como la función  $\Pi_T^E(P)$  es estrictamente creciente en  $P$ , luego si  $\Pi_T^E(P_T^*) > \Pi_T^E(P_N)$  entonces  $P_T^* > P_N$ . Esto completa la demostración de i).

Para demostrar ii) basta con observar que  $P_T^*$  y  $P^S$  resuelven exactamente la misma ecuación:

$$(A29) \quad \alpha \cdot u\left(\left(P_T^* - c_H\right) \cdot \bar{f}\right) + (1 - \alpha) \cdot u\left(\left(P_T^* - c_T\right) \cdot \bar{f}\right) = 0$$

y

$$(A30) \quad \alpha \cdot u\left(\left(P^S - c_H\right) \cdot \bar{f}\right) + (1 - \alpha) \cdot u\left(\left(P^S - c_T\right) \cdot \bar{f}\right) = 0$$

Como el lado izquierdo de esta ecuación es una función estrictamente cóncava, en tanto es una combinación lineal de  $u(\cdot)$  que es estrictamente cóncava también, entonces esa función tiene una única solución cuando igualada a cero,  $P_T^* = P^S$ . ■

**Demostración Corolario 2:** Ya se vio en la demostración previa que  $P^S = P_T^*$ , lo que unido al resultado del Corolario 1,  $P_H^* = P_T^* = P_N$  completa esta prueba. ■

**Demostración Proposición 5:** Dada la definición de  $P_i'$ , con  $i = \{H, T\}$ , para la firma hidráulica se cumple que:

$$(A31) \quad \alpha u\left((P_H' - c_H)\bar{f}\right) + (1 - \alpha)u\left((P_H' - c_H^F)(k_{1-\alpha} + k_T)\right) = 0$$

y para la firma térmica se cumple que:

$$(A32) \quad \alpha u\left((P_T' - c_H)\bar{f}\right) + (1 - \alpha)u\left(P_T'(k_{1-\alpha} + k_T) - c_H^F k_{1-\alpha} - c_T k_T\right) = (1 - \alpha)u\left((c_H^F - c_T)k_T\right)$$

Contrario a lo que se quiere demostrar, supóngase que  $P_H' > P_T'$ . Reemplazando  $P_H'$  en (A32):

$$(A33) \quad \alpha u\left((P_H' - c_H)\bar{f}\right) + (1 - \alpha)u\left(P_H'(k_{1-\alpha} + k_T) - c_H^F k_{1-\alpha} - c_T k_T\right) > (1 - \alpha)u\left((c_H^F - c_T)k_T\right)$$

Reemplazando (A31) en (A33) y simplificando se obtiene:

$$(A34) \quad -u\left((P_H' - c_H^F)(k_{1-\alpha} + k_T)\right) + u\left(P_H'(k_{1-\alpha} + k_T) - c_H^F k_{1-\alpha} - c_T k_T\right) > u\left((c_H^F - c_T)k_T\right)$$

Reordenando convenientemente (A35) encontramos que:

$$(A35) \quad \begin{aligned} & u\left(P_H'(k_{1-\alpha} + k_T) - c_H^F k_{1-\alpha} - c_T k_T\right) - u\left((P_H' - c_H^F)(k_{1-\alpha} + k_T)\right) > \\ & u\left(P_H'(k_{1-\alpha} + k_T) - c_H^F k_{1-\alpha} - c_T k_T - \left(P_H'(k_{1-\alpha} + k_T) - c_H^F k_{1-\alpha} - c_H^F k_T\right)\right) \end{aligned}$$

Lo cual contradice una consecuencia inmediata del supuesto de concavidad estricta de la función de utilidad de los dueños de las empresas generadoras. Por lo tanto,  $P_T' > P_H'$ .

Lo que sigue de la demostración es trivial una vez comprobadas las Proposiciones 3 y 4. La segunda parte de la Proposición 5 se demuestra intercambiando las correspondencias de reacción construidas para la Proposición 2 (la correspondencia de la firma hidráulica corresponde a la de la térmica y vicever-

sa). Es fácil ver que la firma hidráulica se adjudica la subasta ofertando lo mismo que la firma térmica, pues en este caso a igual oferta es la firma hidráulica la que obtiene el contrato con la distribuidora. Las ofertas de cada firma que son equilibrio de Nash están caracterizadas por  $P_H = P_T$  y  $(P_H, P_T) \in [P_H', P_T']$ .

La tercera parte de la demostración es similar. Basta con seguir el refinamiento de la “mano temblorosa” resuelto en la demostración de la Proposición 3, reemplazando las ecuaciones y los pagos de la firma hidráulica por los de la térmica y viceversa. Con todo, es la firma hidráulica la que obtiene el contrato de abastecimiento y a un precio igual a  $P_T'$ . ■

**Demostración Proposición 6:** De acuerdo a la definición de  $P_i''$ , se cumple que:

$$(A36) \quad \alpha \cdot u\left(\left(P_H'' - c_H\right)\bar{f}\right) + (1 - \alpha) \cdot u\left(P_H''(k_{1-\alpha} + k_T) - c_H^F \bar{f}\right) = 0$$

$$(A37) \quad \begin{aligned} & \alpha \cdot u\left(\left(P_T'' - c_H\right)\bar{f}\right) + (1 - \alpha) \cdot u\left(P_T''(k_{1-\alpha} + k_T) - c_H^F \bar{f} + c_H^F k_T - c_T k_T\right) \\ & = (1 - \alpha) \cdot u\left(\left(c_H^F - c_T\right)k_T\right) \end{aligned}$$

Supóngase lo contrario a lo que se desea mostrar, es decir  $P_H'' > P_T''$ . Reemplazando  $P_H''$  en (A37) y siguiendo el mismo procedimiento de la demostración anterior, se llega a:

$$(A38) \quad -u\left(P_H''(k_{1-\alpha} + T) - c_H^F \bar{f}\right) + u\left(P_H''(k_{1-\alpha} + k_T) - c_H^F \bar{f} + c_H^F k_T - c_T k_T\right) > u\left(\left(c_H^F - c_T\right)k_T\right)$$

Lo que contradice el supuesto de función de utilidad estrictamente cóncava. Por lo tanto, se demuestra que  $P_T'' > P_H''$ , lo que significa que la firma hidráulica se adjudica la subasta ofreciendo  $P_H = P_T$  y  $(P_H, P_T) \in [P_H'', P_T'']$ .

Los equilibrios de Nash y el equilibrio de Nash refinado o de la “mano temblorosa”, correspondientes a las partes ii) y iii) de esta proposición, se resuelven siguiendo el mismo procedimiento mencionado en la demostración de la Proposición 5. Esto es, intercambie las ecuaciones y los pagos de la firma hidráulica por los de la térmica y viceversa. Esto cumple con la demostración. ■

## EFFECTOS DISTRIBUTIVOS DE LA REFORMA DE LA SEGURIDAD SOCIAL. EL CASO URUGUAYO\*

ALVARO FORTEZA

*Universidad de la República, Uruguay*

*We present in this paper an estimation of the distributive effects of the reform of the Uruguayan pensions system initiated in 1995. The estimation is based on simulations done with an overlapping generations model adapted and calibrated to the Uruguayan reality. We compute the expected changes in the generational accounts of several groups of workers, considering different generations, gender and income level. The simulations are done in the framework of a small open economy macroeconomic model, and hence we can simultaneously and consistently assess the micro and macro impact of the reform.*

*JEL:* H55, J14

*Keywords:* Pensions, Generational Accounting.

### 1. INTRODUCCIÓN

Las reformas de pensiones que tuvieron lugar en América Latina en las últimas décadas modificaron aspectos fundamentales del funcionamiento de los sistemas de seguridad social, por lo cual cabe esperar que los beneficios y las cargas fiscales que recaen en distintos colectivos de trabajadores hayan cambiado sensiblemente. Pero estos cambios distributivos no se produjeron en forma instantánea, sino que se harán notar a lo largo de todo el ciclo de vida de los trabajadores actuales y futuros. Por lo tanto, el análisis de los efectos distributi-

\* El presente trabajo fue realizado a solicitud de la Conferencia Interamericana de la Seguridad Social (CISS) y con una ayuda financiera aportada por esta institución. Las opiniones vertidas en este documento son de entera responsabilidad de su autor y no involucran a la CISS. Se agradece la eficiente colaboración de Anna Caristo. Comentarios a versiones anteriores de este documento realizados por participantes en un taller organizado por el CISS y por árbitros de esta revista fueron muy útiles para mejorar el artículo. Los errores que puedan permanecer son responsabilidad exclusiva del autor.

Email: alvarof@decon.edu.uy

vos de la reforma es un tema esencialmente dinámico y requiere formular hipótesis sobre la evolución futura, con y sin reforma, de las transferencias netas que realizarán los distintos colectivos de trabajadores analizados. En este documento se presentan los resultados de un análisis de este tipo basado en un modelo de simulación de generaciones solapadas.

Uno de los objetivos fundamentales de las reformas de pensiones que tuvieron lugar en la región en estos años fue contribuir a restablecer la sostenibilidad fiscal que se encontraba comprometida. La insostenibilidad del sendero fiscal impone un desafío al análisis de la política fiscal, porque implica que en algún momento deberá realizarse un ajuste que devuelva la sostenibilidad y ese ajuste no es conocido al momento de realizarse el análisis. Si el inevitable ajuste fiscal futuro no se tiene en cuenta, se obtendrá una visión distorsionada de la situación inicial porque no se estarán contabilizando todas las cargas que se asocian a la trayectoria fiscal que se analiza. En el presente estudio se propone resolver esta dificultad analizando las trayectorias fiscales con los parámetros disponibles (con y sin reforma de la seguridad social), pero incorporando ajustes fiscales futuros que devuelvan la sostenibilidad a todas las trayectorias fiscales bajo análisis. Este supuesto no hace más que reconocer el hecho de que una trayectoria fiscal insostenible es en realidad una trayectoria incompleta en la que no se ha especificado todavía quién pagará las cuentas. El documento desarrolla esta propuesta metodológica para analizar el impacto distributivo de las reformas de pensiones en presencia de senderos fiscales insostenibles y presenta una aplicación en un caso particular, el de Uruguay, que resulta especialmente ilustrativo dado el importante déficit que mostraban las cuentas de la seguridad social al iniciarse la reforma de la seguridad social en este caso.

En la próxima sección se exponen brevemente las principales características del sistema de pensiones de Uruguay y de la reforma que se inició a mediados de los años noventa. La metodología del estudio se presenta en la sección 3. En la sección 4 se exponen los principales resultados. La sección 5 concluye el documento con algunos comentarios finales.

## 2. EL SISTEMA URUGUAYO DE PENSIONES Y LA REFORMA

Hacia mediados de los años noventa, el principal sistema de jubilaciones y pensiones de Uruguay era servido en forma exclusiva por una institución pública, el Banco de Previsión Social (BPS). Se trataba de un sistema de prestaciones definidas, de reparto puro, maduro y que había alcanzado una amplia cobertura<sup>1</sup>. Enfrentaba serias dificultades financieras, asociadas a una población envejecida y niveles de evasión elevados. El gobierno entendía que, debido a su inequidad actuarial, el sistema provocaba importantes distorsiones en los incen-

<sup>1</sup> Al iniciarse la reforma, aproximadamente un 95 por ciento de la población de 60 años y más de Uruguay recibía una pasividad del BPS.

tivos microeconómicos de los trabajadores y empresarios, induciendo la evasión. También se cuestionaba la escasa transparencia y falta de focalización de la redistribución de ingresos realizada a través del sistema de pensiones.

En la segunda mitad de 1995, el Parlamento aprobó una ley de reforma del sistema de pasividades amparado por el BPS que modificó aspectos sustanciales del régimen de jubilaciones y pensiones. Por esta ley se introdujo un sistema mixto con dos pilares, uno de reparto y otro de ahorro individual<sup>2</sup>. El pilar de reparto sigue siendo cuantitativamente fundamental en el sistema uruguayo reformado, ya que todos los trabajadores recibirán al menos parte de su prestación de este pilar y se ha estimado que aproximadamente tres cuartas partes del gasto total en pasividades del sistema seguirá siendo servido por el pilar de reparto en el nuevo régimen (Camacho, 1997). Sin perjuicio de ello, la introducción del pilar de ahorro individual significó un cambio fundamental en el funcionamiento del sistema. Por un lado, la mayor parte de los trabajadores tiene ahora una cuenta de ahorro para su pasividad. El funcionamiento de este pilar requiere de la existencia de un registro individualizado de los aportes realizados, es decir, de una historia laboral, cosa que no se había instrumentado bajo el viejo régimen. Como ocurrió en otros países de la región, la historia laboral se empezó a registrar efectivamente a partir de la aprobación de una reforma que, al introducir un pilar de ahorro individual, no podía operar sin historia laboral<sup>3</sup>. Por otro lado, los fondos de pensiones cuentan ya hoy con fondos acumulados que los convierten en los principales inversores institucionales del país.

La reforma modificó también parámetros fundamentales del pilar de reparto. Se aumentó la edad mínima de jubilación de las mujeres que se igualó a la de los hombres en 60 años (antes estaba en 55). Se modificó la forma de cálculo del salario básico jubilatorio, extendiéndose el período base sobre el cual se realiza el cálculo. Se modificaron las tasas de reemplazo, reduciendo el valor inicial (es decir, la tasa de reemplazo correspondiente al retiro con la edad y años de servicio mínimos) y aumentando su “pendiente” (haciendo más pronunciado el aumento de la tasa a medida que el trabajador posterga el retiro). Se modificaron los mínimos y máximos de prestaciones.

En el presente trabajo se presenta una estimación de los efectos distributivos de la reforma, realizada a partir de simulaciones de un modelo de generaciones solapadas adaptado y calibrado para la realidad uruguayana<sup>4</sup>. Se proyectan los cambios esperados en las cuentas generacionales de distintos colectivos de trabajadores, diferenciados por generación, género y nivel de ingreso. La reali-

<sup>2</sup> Si bien el sistema uruguayo no tiene estrictamente dos pilares independientes o paralelos, es frecuente referirse a los subsistemas de reparto y de ahorro individual como “pilares”.

<sup>3</sup> Forteza (2003) analiza la economía política de la reforma, argumentando que la introducción de un nuevo jugador, las administradoras de fondos de pensiones, desplazó el equilibrio político del sistema, limitando los márgenes de discrecionalidad de los administradores de las instituciones públicas para otorgar pasividades.

<sup>4</sup> El “Modelo Informático de la Seguridad Social” (MISS) es un software programado en GAMS, de libre acceso disponible a solicitud del lector.



zación de las simulaciones en el marco de un modelo macroeconómico de una pequeña economía abierta permite evaluar conjuntamente las proyecciones micro y macroeconómicas, garantizando su coherencia interna. El modelo garantiza que se verifican las restricciones presupuestales y de recursos de todos y cada uno de los agentes de la economía. Si se proyecta una transferencia a favor de un agente, el modelo obliga a establecer claramente de dónde surgen los recursos. Esta es posiblemente la principal fortaleza del instrumento analítico elegido, particularmente cuando se trata de analizar efectos distributivos.

Existen algunos estudios que han analizado los efectos de la reforma de la seguridad social uruguaya en la distribución del ingreso. Noya *et al.* (1999) realizan ejercicios de simulación microeconómica, distinguiendo un significativo número de categorías de trabajadores, lo cual les permite modelizar en forma detallada la diversidad de regímenes de pensiones cubiertos por el BPS. Las simulaciones realizadas con el modelo MISS son algo más estilizadas, ya que no incorporan una riqueza de detalles tan grande. Como contraparte de esta menor riqueza de detalles, el modelo MISS ofrece mayores garantías de consistencia global de los resultados microeconómicos presentados.

Bucheli (1999) presenta resultados de simulación de los efectos distributivos de la reforma basados en un antecedente inmediato del modelo MISS (por una presentación del modelo usado por Bucheli, ver Forteza, 1999b). Como es usual en la literatura, Bucheli (1999) analiza los efectos de la reforma en los ingresos de los trabajadores de varias generaciones sin evaluar si el sendero fiscal es sostenible o no. Los resultados son válidos para los trabajadores analizados, siempre que el gobierno logre mantener el sendero por un tiempo suficientemente largo como para completar el ciclo vital de los trabajadores analizados. Pero este tipo de análisis no brinda un panorama completo de los efectos distributivos de la reforma, porque no establece quién carga con el peso de financiar a la seguridad social cuando el sendero fiscal que se analiza es insostenible. En este sentido, brinda una visión demasiado optimista de los senderos insostenibles. Las estimaciones que se presentan en este documento están libres de ese problema. Por otra parte, las simulaciones que se presentan aquí están basadas en información de los registros administrativos de historia laboral que no estaba disponible cuando Bucheli realizó sus estimaciones.

### 3. METODOLOGÍA

Como se indicó en la introducción, los resultados de simulación que se presentan en este trabajo fueron obtenidos con un modelo de generaciones solapadas. Se trata de una adaptación para una economía pequeña y abierta del modelo de Auerbach y Kotlikoff (1987). El modelo fue calibrado para reproducir los resultados agregados del BPS al iniciarse la reforma. Los detalles del modelo se exponen en Forteza (2002).

En el presente documento se utilizan las cuentas generacionales para analizar los efectos de la reforma sobre distintos colectivos de trabajadores. Se en-

tiende por cuenta generacional a la suma descontada de los aportes realizados y transferencias recibidas por el trabajador a lo largo de toda su vida (ver, entre otros, Auerbach, Gokhale and Kotlikoff, 1994; Auerbach, Kotlikoff and Leibfritz, 1999; Kotlikoff and Raffelhuschen, 1999). Este indicador parece apropiado para evaluar cómo se distribuyen los costos de la reforma entre los distintos grupos de trabajadores.

Un aspecto que resulta particularmente delicado en la evaluación de los efectos distributivos de los programas públicos de pensiones radica en que usualmente una parte significativa de sus ingresos proviene de impuestos generales y no siempre contemporáneos (dado que es posible trasladar parte del peso del financiamiento hacia el futuro a través de la deuda pública). Cuando esto ocurre, se puede obtener una estimación sesgada excesivamente optimista de los beneficios netos que recibe el trabajador del sistema de pensiones si sólo se contabilizan los aportes específicos de seguridad social. Una forma de tratar de escapar de ese problema es estimar la contribución indirecta que realiza el trabajador al sistema de pensiones a través de los impuestos generales. Esta estimación es engorrosa, ya que requiere realizar supuestos sobre los impuestos pagados por cada grupo de trabajadores y, lo que es más difícil, sobre los impuestos que pagará en el futuro.

En el presente trabajo se adopta una estrategia diferente. Se simulan las trayectorias de la economía bajo las viejas y nuevas reglas del sistema de pensiones. En el marco de esas simulaciones, se calculan las transferencias totales netas (contribuciones menos prestaciones) que realizan los trabajadores al gobierno y al sistema de pensiones. Por último, se calculan las variaciones en las transferencias netas que realizan los trabajadores como consecuencia de la reforma. La suma descontada de los cambios en las transferencias es igual al cambio de la cuenta generacional provocado por la reforma. La variación de las cuentas generacionales se usa entonces para evaluar la distribución de los costos de la reforma entre distintos colectivos de trabajadores.

Al adoptar esta estrategia de evaluación de los efectos de la reforma, se toma como punto de referencia la trayectoria simulada de la economía bajo las viejas reglas suponiendo que esas reglas son sostenibles en el período de la simulación. Más específicamente, en el escenario de referencia se supuso que todas las generaciones anteriores a la que nace en 2033 están regidas durante toda su vida por las normas que estaban vigentes antes de la reforma<sup>5</sup>. Si sus cuentas generacionales no están balanceadas y, en consecuencia, la política fiscal simulada no garantiza la solvencia intertemporal del gobierno, se supone que el peso del ajuste fiscal recae en las generaciones siguientes. En la medida en que la trayectoria de la deuda pública se vuelve explosiva en este escenario, el statu quo no puede mantenerse indefinidamente y debe suponerse que en algún momento el gobierno encara una reforma fiscal. En las simulaciones realizadas para este

<sup>5</sup> Este supuesto permite evaluar los efectos puros de la reforma hasta el 2050, cuando el nuevo sistema de seguridad social está maduro. En ese año, la generación 2032 cumple 18 años.

trabajo, se supuso que esa reforma futura inevitable recae sobre las generaciones 2033 y siguientes de manera uniforme<sup>6</sup>. En este sentido, el statu quo es un escenario en que las actuales generaciones le transfieren una carga impositiva creciente a las generaciones futuras. En la medida en que la reforma revierte esta situación, le impone un costo a las generaciones presentes, reduciendo la carga de las generaciones futuras. Al estimar el cambio en las cuentas generacionales se evalúa la distribución de esos costos entre distintos colectivos de trabajadores.

Una ventaja adicional de evaluar los efectos distributivos de la reforma a través del cambio en las cuentas generacionales que deriva de la reforma es que la medida resulta más fácil de obtener y está menos sujeta a errores y a decisiones contables arbitrarias que las usuales mediciones de cuentas generacionales en niveles. El cálculo de las cuentas en niveles requiere estimar las transferencias netas totales que realizan los individuos al gobierno y al sistema de pensiones y esto supone contar con información fiscal detallada a nivel individual, tanto presente como futura. La estimación del cambio en las transferencias permite centrar la atención exclusivamente en aquellos rubros que son modificados por la reforma del sistema de pensiones. Por otra parte, la estimación de la cuenta generacional en niveles es muy sensible a la forma en que se contabilizan algunas partidas del gasto del gobierno, mientras que la medida del cambio en las cuentas generacionales derivadas de la reforma previsional no es sensible a estas opciones contables relativamente arbitrarias<sup>7</sup>.

Un problema que usualmente se enfrenta en la estimación de cuentas generacionales es el de la elección de la tasa de descuento apropiada en un contexto de incertidumbre. En tal ambiente, las tasas de descuento deberían reflejar los riesgos involucrados evaluados desde la perspectiva del individuo cuya cuenta se está calculando. Como destaca Kotlikoff (1999), un flujo de impuestos contingentes puede significar tanto un aumento como una disminución del riesgo que enfrentan los contribuyentes, dependiendo de cómo covaría ese flujo con los ingresos del contribuyente. La variabilidad de los impuestos puede constituir un factor de riesgo para unos contribuyentes y un seguro para otros, de tal manera que la tasa de descuento apropiada para el primer grupo es inferior y para el segundo grupo superior a la tasa libre de riesgo. Lamentablemente, en la práctica no es fácil elegir adecuadamente la tasa de descuento en un contexto de incertidumbre. La práctica estándar es entonces realizar estimaciones para varias tasas de descuento y esto es lo que se hizo también en el marco del presente trabajo.

Se ha discutido en la literatura especializada si la contabilidad generacional recoge adecuadamente los efectos totales de la política fiscal sobre el bienestar de los individuos. Usualmente, las mediciones de contabilidad

<sup>6</sup> Los detalles se explican en el anexo 1.

<sup>7</sup> Raffelhuschen (1999) presenta estimaciones de las cuentas generacionales en Europa según las cuales contabilizar el gasto público en educación como una transferencia arroja una estimación del desbalance generacional casi dos veces mayor al que se obtiene contabilizando el gasto en educación como consumo público.

generacional no captan los efectos indirectos de los impuestos, como son el cambio en la retribución de los factores y el efecto de elusión impositiva<sup>8</sup>. Las estimaciones presentadas en este trabajo son relativamente inmunes a esa crítica, en la medida en que provienen de simulaciones realizadas con un modelo macroeconómico de una pequeña economía abierta. En particular, el modelo estima los cambios en la retribución de los factores inducidos por la reforma. Como se explica en Forteza (1999b, 2002), los supuestos de economía pequeña y abierta, individuos con horizonte de vida finito y ausencia de herencias permiten independizar los resultados económicos presentes de los desempeños en un futuro suficientemente lejano. Corresponde notar que esto no sería válido, por ejemplo, si se supusiera que los individuos son altruistas y dejan herencias.

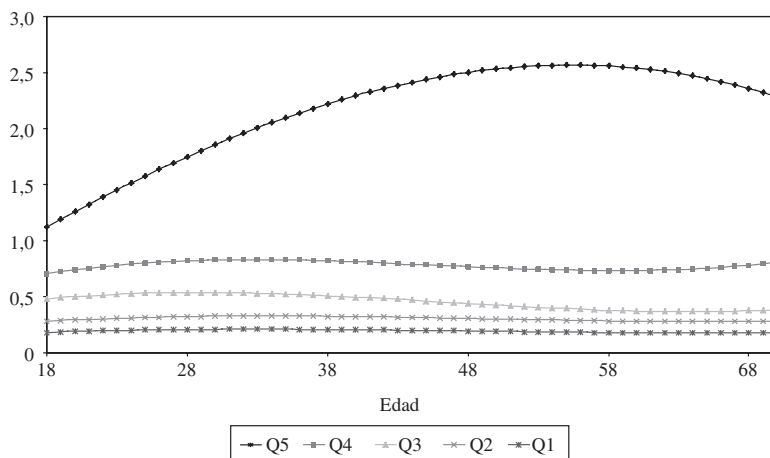
En el marco de esta investigación y dado que se dispone de un modelo de simulación calibrado para el país, se podrían evaluar los efectos de la reforma calculando los cambios en la utilidad de los individuos. Se prefirió presentar la cuenta generacional, por tratarse de una medida más intuitiva, que permite realizar comparaciones cuantitativas. Tiene sentido afirmar, por ejemplo, que la reforma “le cuesta” al trabajador de ingresos medios de la generación de 1980 el equivalente a ‘*x*’ salarios, pero no tiene sentido afirmar que la reforma “le cuesta” un ‘*y*’ por ciento de su utilidad, dado que la utilidad es una variable ordinal. En el marco del modelo se podría haber utilizado un indicador cardinal más “refinado” que la cuenta generacional, como la variación compensada, pero dada la correlación existente entre la cuenta generacional y la utilidad en diversos experimentos realizados con estos modelos (Fehr and Kotlikoff, 1999) se consideró preferible utilizar la cuenta generacional por ser un indicador más intuitivo y fácil de calcular.

El modelo utilizado reconoce cinco perfiles de ingresos laborales por edades para cada sexo. Un perfil de ingreso laboral es una función que va de la edad del trabajador al ingreso que obtiene cuando está en actividad. Los perfiles fueron estimados utilizando un panel de registros administrativos de historia laboral del BPS (ver Anexo 2). Se estimó un perfil diferente para cada quintil de la distribución del ingreso. Como se observa en la Figura 1, los trabajadores de altos ingresos tienen un perfil salarial de mayor pendiente que los trabajadores de menores ingresos. Este es un hecho que ha sido descrito también en otros países (ver, por ejemplo, Bosworth *et al.*, 1999).

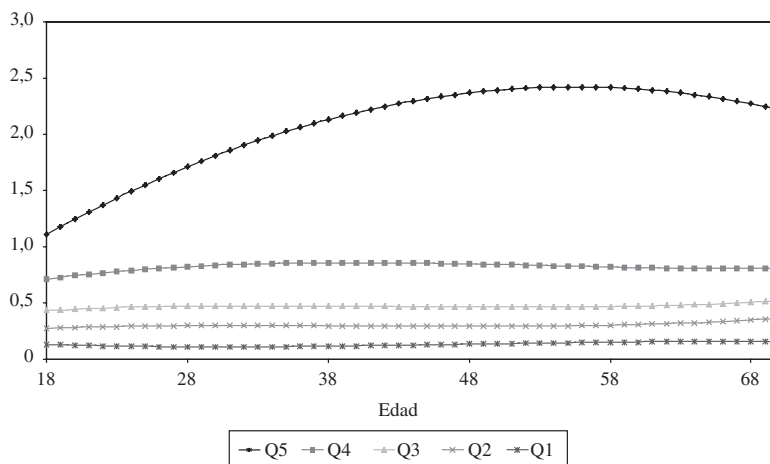
El quintil de mayores ingresos de la muestra de historia laboral tiene ingresos equivalentes a 2,8 y 2,7 veces el ingreso promedio de la muestra, para hombres y mujeres, respectivamente. El segundo quintil de mayores ingresos tiene ingresos iguales al promedio, tanto para hombres como para mujeres. El tercer quintil tiene ingresos equivalentes a 0,6 veces el promedio, tanto para hombres como para mujeres. El cuarto quintil tiene ingresos iguales a 0,4 veces

<sup>8</sup> Fehr and Kotlikoff (1999) analizan estas críticas y proponen varias correcciones que deben realizarse para atenuar el problema en diversos escenarios.

FIGURA 1  
PERFILES DE INGRESOS LABORALES  
HOMBRES  
(por quintiles de ingreso)



PERFILES DE INGRESOS LABORALES  
MUJERES  
(por quintiles de ingreso)



Fuente: Elaboración propia.

el promedio. Finalmente, el quintil de menores ingresos tiene ingresos equivalentes a 0,2 veces el ingreso promedio. Estos resultados, según los cuales la media es sensiblemente mayor a la mediana, reflejan la asimetría hacia la derecha que presenta la distribución.

#### 4. RESULTADOS

El presente trabajo centra su atención en los efectos de la reforma de la seguridad social sobre trabajadores que se rigen estrictamente por las normas del sistema. Por comodidad, se les llamará trabajadores formales. La pregunta que se intenta responder es cómo se verían afectados por la reforma distintos colectivos de trabajadores que hicieran aportes y recibieran beneficios del sistema de pensiones ateniéndose a las reglas formales del mismo. Sin embargo, el sistema sufre de evasión y esto afecta la carga de trabajadores futuros. Por esa razón se presentan primero brevemente estimaciones agregadas de las cuentas generacionales que incluyen a todos los miembros de la generación, es decir, que incluyen también a los evasores. Su consideración resulta fundamental para evaluar el efecto de la reforma en las generaciones futuras (2033 y siguientes). Luego se presenta un análisis más detallado de los efectos diferenciales de la reforma sobre distintos colectivos de trabajadores formales.

La distinción tajante que se realiza en estas simulaciones entre trabajadores formales e informales no refleja la diversidad de situaciones que se observan en la realidad, donde un número significativo de trabajadores pasa parte de su tiempo en actividades formales y parte en actividades informales. Las simulaciones realizadas tampoco reflejan el fenómeno de la subdeclaración de ingresos. El modelo permite simular estos casos (personas que oscilan entre empleos formales e informales y subdeclaración), pero como se señala en el párrafo anterior, el objetivo del estudio fue centrar la atención en el efecto que cabe esperar que tenga la reforma en los trabajadores que se atienen estrictamente a las reglas del sistema. No obstante ello y dada la importancia que tiene la evasión en Uruguay, se decidió representar este fenómeno en forma muy estilizada simulando la situación de trabajadores que evaden las contribuciones enteramente. En el contexto del presente trabajo, la consideración de estos trabajadores importa esencialmente para evaluar las cargas de las generaciones futuras. Probablemente esas cargas no serían muy diferentes de lo que se estima en este documento si en lugar de incorporar al modelo un grupo de trabajadores que evade todo el tiempo se incorporaran más trabajadores que evaden durante parte de su carrera laboral.

Los efectos de la reforma en los ingresos de los trabajadores dependen de un conjunto de variables cuyos valores futuros son difíciles de predecir. Resulta difícil establecer incluso la probabilidad de ocurrencia de distintos valores de esas variables. Por esa razón, es usual realizar proyecciones en un escenario base y luego analizar la sensibilidad de los resultados a cambios en las variables exógenas. En el presente trabajo se analizó la sensibilidad de las cuentas generacionales de los trabajadores formales frente a cambios en la edad de retiro de los trabajadores, en la tasa de interés y en la tasa de crecimiento de la productividad del trabajo.

En todos los escenarios se realizan estimaciones con y sin reforma, es decir, que se simulan los senderos esperados con las nuevas reglas de juego y con las anteriores. El efecto de la reforma se analiza comparando las transferencias netas con y sin reforma a lo largo de la vida de los trabajadores simulados. Se trata entonces de un análisis de dinámica comparada.

#### 4.1 Las cuentas generacionales “agregadas” (trabajadores formales e informales)

En el escenario base se supuso, en primer lugar, que la tasa de interés real es 3,8 por ciento anual. Esta tasa fue utilizada tanto para capitalizar los aportes realizados por los trabajadores a los fondos de ahorro previsional como para descontar los flujos de ingresos y transferencias reales. En segundo lugar se fijó la tasa de crecimiento de la productividad del trabajo en 1,1 por ciento anual. Los trabajadores simulados son heterogéneos en su productividad, lo cual permite recoger en el modelo una diversidad de salarios reales. Se supuso que la tasa de crecimiento de la productividad es la misma para todos. Los salarios reales crecen en el modelo a la misma tasa que la productividad del trabajo. Con estos supuestos se obtiene endógenamente una tasa de crecimiento del producto bruto interno que en promedio es algo menos de 1,7 por ciento anual, siendo la tasa de crecimiento de la población aproximadamente 0,6 por ciento anual. En tercer lugar, en el escenario base se supuso que no se modifica la edad de retiro como consecuencia de la reforma, salvo por el efecto esperable (pequeño) del aumento en la edad mínima de retiro de las mujeres. La edad promedio de retiro al momento de la reforma era de 64 años para los hombres y 59 para las mujeres. La ley de reforma fijó en 60 la edad mínima de retiro y, por lo tanto, en el escenario base se modificó la edad de retiro de las mujeres de 59 a 60.

Como cabía esperar, la reforma beneficia fundamentalmente a las generaciones “futuras”. En las simulaciones que se resumen en el Cuadro 1, la reforma de la seguridad social reduce fuertemente las transferencias netas al gobierno de las generaciones nacidas en 2033 y después<sup>9</sup>. Este resultado es consistente con proyecciones fiscales realizadas con supuestos similares a los del presente trabajo que indican que la reforma provoca una disminución de la deuda pública a mediano y largo plazo (Caristo y Forteza, 2003).

Ese alivio fiscal para las generaciones futuras tiene como contraparte un aumento de la carga que recae en casi todas las generaciones previas. Es inevitable que el alivio de unas generaciones esté acompañado de una mayor carga para las restantes, dada la naturaleza de suma cero de la contabilidad generacional (ver, por ejemplo, Kotlikoff, 1999), pero el Cuadro 1 aporta información sobre la distribución de esa mayor carga en el caso específico de la reforma uruguaya de la seguridad social.

Hay un grupo de generaciones nacidas antes de la reforma que no compartieron sus costos. Son las personas que al iniciarse la reforma estaban ya retiradas y recibiendo la pensión y las que sin estarlo habían adquirido el derecho pero no lo habían ejercido todavía (es decir, habían generado “causal jubilatoria” pero seguían en actividad). La ley de reforma (ley 16.713) respetó plenamente los derechos adquiridos por los primeros y otorgó nuevos beneficios a los segundos. En las

<sup>9</sup> La elección de la generación nacida en 2033 como la primera de las generaciones “futuras” es arbitraria. Como ya se indicó, se hizo esta elección por conveniencia analítica, para cubrir un período suficiente como para alcanzar la madurez del sistema de seguridad social reformado. Lo que no es arbitrario es el hecho fundamental de que el ajuste que no se realiza hoy debe realizarse en algún momento futuro.

simulaciones realizadas, pertenecen al primer grupo las generaciones nacidas en 1932 y antes y al segundo las generaciones nacidas en 1933 a 1936 inclusive<sup>10</sup>.

CUADRO 1  
VARIACIÓN DE LAS CUENTAS GENERACIONALES  
DEBIDO A LA REFORMA  
Trabajadores formales e informales

Generación	Variación de las cuentas generacionales
1932 y anteriores	0,0
1933	-1,1
1940	1,0
1950	3,3
1957	2,3
1960	2,1
1970	1,6
1980	1,2
1990	1,0
2000	0,8
2033 y posteriores	-2,5

Nota: Descontadas a 1995 y expresadas en puntos porcentuales del producto bruto interno de ese año. Un valor positivo (negativo) significa que el trabajador pierde (gana) con la reforma, ya que transfiere más (menos) después de la reforma de lo que hubiera transferido sin reforma.

Fuente: Elaboración propia.

El costo de la reforma recae en las generaciones intermedias (generaciones 1937 a 2032 en estas simulaciones). Todas ellas aumentan sus cuentas generacionales después de la reforma, pero los resultados resumidos en el Cuadro 1 muestran una considerable diversidad en la contribución de estas generaciones. Esta diversidad responde a su vez a las distintas formas en que la reforma afecta a diversos colectivos de trabajadores, tanto formales como informales. En las secciones que siguen se analiza esta diversidad, poniendo el énfasis en los efectos de la reforma sobre los trabajadores formales.

#### 4.2. Las cuentas generacionales de los trabajadores formales

##### *El escenario base*

Los cambios proyectados en las cuentas generacionales de los trabajadores *formales* indican que estos trabajadores, en promedio, se benefician con

<sup>10</sup> Por razones de espacio, los cuadros presentan resultados para algunas generaciones seleccionadas que son especialmente representativas de las situaciones que se comentan. Los resultados completos están disponibles para quien los solicite.



la reforma (Cuadro 2). Este resultado es consistente con proyecciones fiscales que indican que, en las condiciones de este escenario, la reforma puede provocar una mejora fiscal esencialmente si reduce las transferencias a los trabajadores informales (Forteza, 1999c; Caristo y Forteza, 2003). Se ha estimado que una parte significativa de los egresos del Banco de Previsión Social (BPS) al iniciarse la reforma eran prestaciones irregulares, obtenidas a través del otorgamiento de beneficios contributivos a individuos que no habían contribuido (Camacho, 1997, estima que aproximadamente un 20 por ciento de los egresos totales de invalidez, vejez y sobrevivencia del BPS en 1995 se explican por este concepto). Según los resultados que aquí se presentan, las proyecciones fiscales que indican mejoras fiscales en escenarios como el simulado aquí se apoyan en el supuesto de que a partir de la reforma dejan de otorgarse estos beneficios no legales<sup>11</sup> y no en que la reforma haya endurecido las condiciones que enfrentan los trabajadores que se rigen estrictamente por las normas del sistema. Según estos resultados y si se verifican las condiciones supuestas en este escenario, el peso del ajuste de la reforma uruguaya de la seguridad social recae fundamentalmente en los individuos que antes de la reforma obtenían estos beneficios en forma irregular. Si la reducción prevista en la evasión no se verificara, los trabajadores formales actuales no se verían *directamente* afectados<sup>12</sup>, pero los resultados fiscales serían peores a los proyectados y la carga de las generaciones futuras aumentaría.

El resultado promedio esconde una considerable variabilidad a través de género, niveles de ingreso y a lo largo de generaciones. Tanto los hombres como las mujeres que trabajan formalmente se verían beneficiados, pero mientras que los hombres tendrían una ganancia equivalente a 1,7 años de su salario de 55 años de edad, las mujeres tendrían una ganancia de sólo 0,1 año. Este resultado es consistente con estimaciones previas y con un objetivo expreso de la reforma en el sentido de reducir las diferencias que existían en el régimen anterior a favor de las mujeres (Bucheli, 1999, Noya *et al.*, 1999).

Los trabajadores que se benefician en mayor medida son los del quintil de menores ingresos. De acuerdo con las estimaciones presentadas en el Cuadro 2, estos trabajadores reducen sus transferencias netas a la seguridad social en montos considerables. Entre los factores que explican la mejora de estos trabajadores puede mencionarse el aumento de la prestación mínima y la boni-

<sup>11</sup> Existe un consenso generalizado entre los observadores y operadores del sistema de jubilaciones uruguayo en que el BPS endureció sensiblemente los controles de las condiciones de acceso a las jubilaciones en los noventa. Forteza (2003) argumenta que la introducción de los registros de historia laboral que se asocia a la reforma a raíz de las necesidades del pilar de ahorro individual es un determinante clave de este cambio de gestión. Murro (citado por Forteza 2003) entiende que el cambio se debe fundamentalmente a la incorporación de representantes sociales (de trabajadores, empresarios y pensionistas) en el directorio del BPS que puso límite a las prácticas clientelísticas de anteriores administraciones.

<sup>12</sup> Podrían verse afectados en forma indirecta, pero no por ello menos importante, si al no producirse la esperada reducción en la evasión el deterioro fiscal generara inestabilidad macroeconómica u obligara a un ajuste fiscal que hiciera recaer mayores cargas en los trabajadores formales actuales.

CUADRO 2  
VARIACIÓN DE LAS CUENTAS GENERACIONALES DE TRABAJADORES  
FORMALES DEBIDO A LA REFORMA. ESCENARIO BASE<sup>a/</sup>

Nivel de ingreso							Régimen aplicable (ley 16.713) <sup>c/</sup>
Generación	Alto	Alto-medio	Medio	Medio-bajo	Bajo	Promedio	
A) Hombres							
1933	-1,8	-0,8	-2,4	-4,0	-8,3	-2,0	titV
1940	-1,8	0,1	-2,3	-4,0	-8,3	-1,9	rt
1950	-0,3	-1,9	-4,1	-4,1	-8,4	-1,1	rt/mixto
1957	-0,9	-2,7	-4,7	-3,0	-7,4	-1,6	mixto
1960	-1,2	-2,2	-4,1	-3,4	-7,6	-1,7	mixto
1970	-1,9	-1,2	-3,3	-4,3	-8,1	-2,1	mixto
1980	-1,9	-1,2	-3,4	-4,4	-8,1	-2,2	mixto
1990	-1,8	-1,1	-3,4	-4,5	-8,5	-2,1	mixto
2000	-1,8	-1,1	-3,4	-4,5	-8,5	-2,1	mixto
Promedio <sup>b/</sup>	-1,4	-0,9	-2,8	-3,7	-7,2	-1,7	
B) Mujeres							
1933	0,0	0,0	0,0	0,0	0,0	0,0	art2
1940	-2,8	-0,3	-0,1	-1,0	-2,6	-2,1	titV
1950	1,2	0,7	0,8	-1,0	-2,7	0,9	rt/mixto
1957	0,7	-0,6	-0,5	-0,2	-0,9	0,4	mixto
1960	0,5	-0,2	0,0	-0,1	-1,0	0,3	mixto
1970	-0,1	1,0	1,2	-1,2	-1,4	0,1	mixto
1980	-0,4	1,1	1,2	-1,6	-1,6	-0,1	mixto
1990	-0,4	1,1	1,2	-1,6	-1,6	-0,1	mixto
2000	-0,4	1,1	1,2	-1,6	-1,6	-0,1	mixto
Promedio <sup>b/</sup>	-0,3	0,8	1,0	-1,1	-1,4	-0,1	

Notas:

<sup>a/</sup> Medido en salarios anuales del trabajador a los 55 años de edad. Un valor positivo (negativo) significa que el trabajador pierde (gana) con la reforma, ya que transfiere más (menos) después de la reforma de lo que hubiera transferido sin reforma.

<sup>b/</sup> Promedio de todas las generaciones simuladas (1916 a 2032).

<sup>c/</sup> Régimen aplicable: (i) art2 es el régimen que se aplica a personas que estaban ya jubiladas al entrar en vigencia la ley de reforma, (ii) titV es el régimen que se aplica a personas que sin estar jubiladas tenían generada la causal jubilatoria al iniciarse la reforma, (iii) rt es el llamado "régimen de transición" y se aplica a personas que tenían 40 años o más y no tenían generada la causal jubilatoria al entrar en vigencia la ley, (iv) mixto es el régimen que se aplica a quienes no habían cumplido 40 años al entrar en vigencia la ley.

Fuente: Elaboración propia.

ficación que se otorga a los trabajadores que optan por aportar al pilar de ahorro individual<sup>13</sup>. Pero la reforma benefició también a trabajadores de ingresos altos por varias vías, incluyendo el aumento en las prestaciones máximas, la

<sup>13</sup> En el sistema mixto uruguayo, los trabajadores de menores ingresos se mantienen, en principio, en el pilar de reparto exclusivamente. No obstante, el artículo 8 de la ley 16.713 concede a estos trabajadores la opción de asignar la mitad de sus aportes al pilar de ahorro individual, en cuyo caso se bonifican los aportes que quedan en el pilar de reparto a los efectos del cálculo de la prestación en un 50 por ciento.

sustitución parcial del pilar de reparto por el de ahorro individual y la reducción de los aportes patronales por encima de cierto umbral de ingreso. Estos resultados son consistentes con la intención declarada de los reformadores de aumentar la equidad actuarial del sistema, como forma de reducir las distorsiones en el mercado de trabajo.

### *Análisis de Sensibilidad*

#### La edad de retiro

Se ha señalado que los efectos macroeconómicos de la reforma dependen, entre otras cosas, de si las edades de retiro se modifican en las nuevas condiciones. Forteza (1999c) muestra que si los trabajadores postergan el retiro, los resultados fiscales mejoran considerablemente. Por otra parte, Durán, Poplavski y Vernengo (2003) presentan resultados de simulación de la decisión de retiro que sugieren fuertemente que la reforma aporta incentivos suficientes para que muchos trabajadores posterguen su retiro. Existe alguna evidencia informal de que este proceso ya se está verificando. Interesa entonces analizar cómo afecta un aumento de la edad de retiro a las cuentas generacionales de los trabajadores.

Los Cuadros 3 y 4 resumen resultados de simulaciones en que se supusieron condiciones análogas a las del escenario base, salvo por las edades de retiro después de la reforma que se supusieron uno y dos años mayores que en el escenario base, respectivamente. Los resultados de los Cuadros 2 a 4 muestran que el aumento de la edad de retiro aumenta las transferencias netas que hacen los trabajadores a la seguridad social. Si bien este resultado puede parecer esperable, existen dos efectos que operan en sentido opuesto y que, en principio, tornan el resultado ambiguo. Por un lado, un trabajador que se retira más tarde paga aportes durante más tiempo y cobra beneficios durante menos tiempo, lo cual deteriora la cuenta generacional. Pero, por otro lado, las tasas de reemplazo en el nuevo régimen son crecientes con la edad de retiro y los años de servicio. Los resultados de simulación indican que el primer efecto predomina sobre el segundo. Estos resultados de las cuentas generacionales son consistentes con las proyecciones fiscales que indican que las cuentas del gobierno mejoran si los trabajadores postergan el retiro.

La postergación del retiro aumenta las transferencias netas que hacen los trabajadores de casi todos los estratos de ingreso. La excepción más notoria es el caso de las mujeres de más bajos ingresos que pueden incrementar sus ganancias postergando el retiro.

Según los resultados comentados, la postergación del retiro reduce las transferencias netas que recibe el trabajador. Esto no significa, sin embargo, que el trabajador reduzca sus ingresos cuando se retira más tarde, ya que los salarios son mayores a la pensión y la pensión es una función creciente de los años trabajados. Por esta razón, es posible que los trabajadores prefieran retirarse más tarde, aun cuando ello signifique enfrentar una cuenta generacional menos favorable.

**CUADRO 3**  
**VARIACIÓN DE LAS CUENTAS GENERACIONALES**  
**DE TRABAJADORES FORMALES DEBIDO A LA REFORMA.**  
**ESCENARIO BASE Y AUMENTA UN AÑO EDAD DE RETIRO<sup>a/</sup>**

Nivel de ingreso							
Generación	Alto	Alto-medio	Medio	Medio-bajo	Bajo	Promedio	Régimen aplicable (ley 16.713) <sup>c/</sup>
<b>A) Hombres</b>							
1933	3,0	3,3	-2,0	-3,7	-7,8	2,2	titV
1940	-1,1	0,6	-2,0	-3,7	-7,8	-1,3	rt
1950	-0,1	-1,4	-3,6	-3,7	-7,9	-0,8	rt/mixto
1957	-0,7	-2,2	-4,2	-2,8	-7,1	-1,3	mixto
1960	-1,0	-1,7	-3,6	-3,1	-7,2	-1,5	mixto
1970	-1,6	-0,8	-3,1	-4,1	-7,7	-1,9	mixto
1980	-1,7	-0,9	-3,2	-4,2	-7,8	-1,9	mixto
1990	-1,6	-0,8	-3,2	-4,3	-8,2	-1,9	mixto
2000	-1,6	-0,8	-3,2	-4,3	-8,2	-1,9	mixto
Promedio <sup>b/</sup>	-1,0	-0,4	-2,6	-3,5	-6,9	-1,3	
<b>B) Mujeres</b>							
1933	0,0	0,0	0,0	0,0	0,0	0,0	art2
1940	4,9	5,8	2,6	-1,3	-3,6	4,3	titV
1950	1,4	1,0	0,9	-1,4	-3,7	1,0	rt/mixto
1957	0,9	-0,4	-0,3	0,0	-1,8	0,5	mixto
1960	0,7	0,0	0,2	-0,3	-1,9	0,4	mixto
1970	0,1	1,2	1,4	-1,4	-2,4	0,2	mixto
1980	-0,2	1,2	1,1	-1,9	-2,6	0,0	mixto
1990	-0,2	1,2	1,1	-1,9	-2,6	0,0	mixto
2000	-0,2	1,2	1,1	-1,9	-2,6	0,0	mixto
Promedio <sup>b/</sup>	0,2	1,2	1,0	-1,3	-2,3	0,3	

Notas:

<sup>a/</sup> Medido en salarios anuales del trabajador a los 55 años de edad. Un valor positivo (negativo) significa que el trabajador pierde (gana) con la reforma, ya que transfiere más (menos) después de la reforma de lo que hubiera transferido sin reforma.

<sup>b/</sup> Promedio de todas las generaciones simuladas (1916 a 2032).

<sup>c/</sup> Régimen aplicable: (i) art2 es el régimen que se aplica a personas que estaban ya jubiladas al entrar en vigencia la ley de reforma, (ii) titV es el régimen que se aplica a personas que sin estar jubiladas tenían generada la causal jubilatoria al iniciarse la reforma, (iii) rt es el llamado "régimen de transición" y se aplica a personas que tenían 40 años o más y no tenían generada la causal jubilatoria al entrar en vigencia la ley, (iv) mixto es el régimen que se aplica a quienes no habían cumplido 40 años al entrar en vigencia la ley.

Fuente: Elaboración propia.

**CUADRO 4**  
**VARIACIÓN DE LAS CUENTAS GENERACIONALES**  
**DE TRABAJADORES FORMALES DEBIDO A LA REFORMA.**  
**ESCENARIO BASE Y AUMENTA DOS AÑOS EDAD DE RETIRO<sup>a/</sup>**

Nivel de ingreso							
Generación	Alto	Alto-medio	Medio	Medio-bajo	Bajo	Promedio	Régimen aplicable (ley 16.713) <sup>c/</sup>
<b>A) Hombres</b>							
1933	3,2	3,6	-1,5	-3,2	-7,0	2,4	titV
1940	-0,5	1,2	-1,6	-3,2	-7,1	-0,7	rt
1950	0,2	-0,8	-3,1	-3,3	-7,1	-0,5	rt/mixto
1957	-0,4	-1,7	-3,6	-2,5	-6,5	-1,0	mixto
1960	-0,7	-1,2	-3,1	-2,8	-6,6	-1,2	mixto
1970	-1,4	-0,4	-2,8	-3,8	-7,1	-1,6	mixto
1980	-1,4	-0,5	-2,9	-3,8	-7,2	-1,7	mixto
1990	-1,4	-0,4	-3,0	-4,0	-7,8	-1,6	mixto
2000	-1,4	-0,4	-3,0	-4,0	-7,8	-1,6	mixto
Promedio <sup>b/</sup>	-0,8	-0,1	-2,3	-3,2	-6,4	-1,1	
<b>B) Mujeres</b>							
1933	0,0	0,0	0,0	0,0	0,0	0,0	art2
1940	5,0	5,9	2,6	-1,5	-4,3	4,4	titV
1950	1,6	1,2	1,0	-1,6	-4,4	1,2	rt/mixto
1957	1,1	-0,2	0,0	-0,2	-2,4	0,7	mixto
1960	0,9	0,2	0,4	-0,5	-2,5	0,6	mixto
1970	0,3	1,4	1,5	-1,6	-3,1	0,4	mixto
1980	0,0	1,4	1,0	-2,0	-3,3	0,1	mixto
1990	0,0	1,4	1,0	-2,0	-3,3	0,1	mixto
2000	0,0	1,4	1,0	-2,0	-3,3	0,1	mixto
Promedio <sup>b/</sup>	0,4	1,4	1,0	-1,4	-2,9	0,4	

## Notas:

<sup>a/</sup> Medido en salarios anuales del trabajador a los 55 años de edad. Un valor positivo (negativo) significa que el trabajador pierde (gana) con la reforma, ya que transfiere más (menos) después de la reforma de lo que hubiera transferido sin reforma.

<sup>b/</sup> Promedio de todas las generaciones simuladas (1916 a 2032).

<sup>c/</sup> Régimen aplicable: (i) art2 es el régimen que se aplica a personas que estaban ya jubiladas al entrar en vigencia la ley de reforma, (ii) titV es el régimen que se aplica a personas que sin estar jubiladas tenían generada la causal jubilatoria al iniciarse la reforma, (iii) rt es el llamado "régimen de transición" y se aplica a personas que tenían 40 años o más y no tenían generada la causal jubilatoria al entrar en vigencia la ley, (iv) mixto es el régimen que se aplica a quienes no habían cumplido 40 años al entrar en vigencia la ley.

Fuente: Elaboración propia.

## La tasa de interés

Se repitieron las simulaciones con dos tasas de interés adicionales, una menor y otra mayor a la del escenario base (Cuadros 5 y 6, respectivamente). En

CUADRO 5  
VARIACIÓN DE LAS CUENTAS GENERACIONALES  
DE TRABAJADORES FORMALES DEBIDO A LA REFORMA. ESCENARIO  
BASE Y TASA DE INTERÉS REAL IGUAL A 2,8 POR CIENTO ANUAL<sup>a/</sup>

Nivel de ingreso							
Generación	Alto	Alto-medio	Medio	Medio-bajo	Bajo	Promedio	Régimen aplicable (ley 16.713) <sup>c/</sup>
A) Hombres							
1933	-2,1	-0,9	-2,8	-4,6	-9,6	-2,3	titV
1940	-2,1	0,1	-2,7	-4,6	-9,6	-2,1	rt
1950	-0,2	-2,1	-4,6	-4,7	-9,7	-1,1	rt/mixto
1957	-0,7	-2,8	-5,1	-3,2	-8,4	-1,5	mixto
1960	-0,9	-2,1	-4,2	-3,4	-8,6	-1,5	mixto
1970	-1,4	-0,5	-3,0	-4,2	-9,0	-1,7	mixto
1980	-1,5	-0,6	-3,0	-4,3	-9,0	-1,8	mixto
1990	-1,4	-0,5	-3,0	-4,4	-9,5	-1,7	mixto
2000	-1,4	-0,5	-3,0	-4,4	-9,5	-1,7	mixto
Promedio <sup>b/</sup>	-1,2	-0,5	-2,7	-3,7	-8,1	-1,5	
B) Mujeres							
1933	0,0	0,0	0,0	0,0	0,0	0,0	art2
1940	-3,2	-0,4	-0,2	-1,1	-3,0	-2,4	titV
1950	1,5	0,9	1,0	-1,2	-3,1	1,2	rt/mixto
1957	1,1	-0,4	-0,2	0,0	-0,9	0,7	mixto
1960	0,9	0,2	0,3	0,3	-1,0	0,7	mixto
1970	0,4	2,0	2,2	-0,5	-1,3	0,7	mixto
1980	0,3	2,2	2,4	-0,9	-1,5	0,6	mixto
1990	0,3	2,2	2,4	-0,9	-1,5	0,6	mixto
2000	0,3	2,2	2,4	-0,9	-1,5	0,6	mixto
Promedio <sup>b/</sup>	0,1	1,6	1,7	-0,6	-1,4	0,4	

Notas:

<sup>a/</sup> Medido en salarios anuales del trabajador a los 55 años de edad. Un valor positivo (negativo) significa que el trabajador pierde (gana) con la reforma, ya que transfiere más (menos) después de la reforma de lo que hubiera transferido sin reforma.

<sup>b/</sup> Promedio de todas las generaciones simuladas (1916 a 2032).

<sup>c/</sup> Régimen aplicable: (i) art2 es el régimen que se aplica a personas que estaban ya jubiladas al entrar en vigencia la ley de reforma, (ii) titV es el régimen que se aplica a personas que sin estar jubiladas tenían generada la causal jubilatoria al iniciarse la reforma, (iii) rt es el llamado "régimen de transición" y se aplica a personas que tenían 40 años o más y no tenían generada la causal jubilatoria al entrar en vigencia la ley, (iv) mixto es el régimen que se aplica a quienes no habían cumplido 40 años al entrar en vigencia la ley.

Fuente: Elaboración propia.

CUADRO 6  
 VARIACIÓN DE LAS CUENTAS GENERACIONALES DE TRABAJADORES  
 FORMALES DEBIDO A LA REFORMA. ESCENARIO BASE Y TASA  
 DE INTERÉS REAL IGUAL A 4,8 POR CIENTO ANUAL<sup>a/</sup>

Nivel de ingreso							
Generación	Alto	Alto-medio	Medio	Medio-bajo	Bajo	Promedio	Régimen aplicable (ley 16.713) <sup>c/</sup>
A) Hombres							
1933	-1,6	-0,7	-2,1	-3,5	-7,2	-1,7	titV
1940	-1,6	0,1	-2,0	-3,5	-7,2	-1,6	rt
1950	-0,5	-1,7	-3,6	-3,6	-7,3	-1,1	rt/mixto
1957	-1,1	-2,7	-4,4	-2,9	-6,6	-1,7	mixto
1960	-1,5	-2,3	-4,0	-3,3	-6,7	-1,9	mixto
1970	-2,3	-1,8	-3,8	-4,5	-7,3	-2,6	mixto
1980	-2,4	-1,9	-3,9	-4,6	-7,4	-2,6	mixto
1990	-2,3	-1,8	-3,9	-4,7	-7,7	-2,6	mixto
2000	-2,3	-1,8	-3,9	-4,7	-7,7	-2,6	mixto
Promedio <sup>b/</sup>	-1,9	-1,3	-3,0	-3,7	-6,5	-1,9	
B) Mujeres							
1933	0,0	0,0	0,0	0,0	0,0	0,0	art2
1940	-2,4	-0,3	-0,1	-0,8	-2,3	-1,8	titV
1950	1,0	0,6	0,7	-0,9	-2,3	0,7	rt/mixto
1957	0,4	-0,8	-0,7	-0,5	-0,9	0,1	mixto
1960	0,2	-0,5	-0,4	-0,4	-1,0	-0,1	mixto
1970	-0,7	0,1	0,3	-1,8	-1,5	-0,6	mixto
1980	-1,0	-0,1	0,0	-2,4	-1,8	-0,9	mixto
1990	-1,0	-0,1	0,0	-2,4	-1,8	-0,9	mixto
2000	-1,0	-0,1	0,0	-2,4	-1,8	-0,9	mixto
Promedio <sup>b/</sup>	-0,6	0,1	0,2	-1,5	-1,5	-0,5	

Notas:

<sup>a/</sup> Medido en salarios anuales del trabajador a los 55 años de edad. Un valor positivo (negativo) significa que el trabajador pierde (gana) con la reforma, ya que transfiere más (menos) después de la reforma de lo que hubiera transferido sin reforma.

<sup>b/</sup> Promedio de todas las generaciones simuladas (1916 a 2032).

<sup>c/</sup> Régimen aplicable: (i) art2 es el régimen que se aplica a personas que estaban ya jubiladas al entrar en vigencia la ley de reforma, (ii) titV es el régimen que se aplica a personas que sin estar jubiladas tenían generada la causal jubilatoria al iniciarse la reforma, (iii) rt es el llamado "régimen de transición" y se aplica a personas que tenían 40 años o más y no tenían generada la causal jubilatoria al entrar en vigencia la ley, (iv) mixto es el régimen que se aplica a quienes no habían cumplido 40 años al entrar en vigencia la ley.

Fuente: Elaboración propia.

términos generales, la reforma genera mayores ganancias y menores pérdidas para los trabajadores formales cuanto mayor es la tasa de interés. Sólo las primeras generaciones afectadas por la reforma ganan menos si la tasa de interés aumenta. Se trata de generaciones que se mantienen enteramente en el sistema de reparto<sup>14</sup>. Las restantes generaciones se benefician con mayores tasas de interés. El mismo patrón se observa en ambos géneros y en los cinco niveles de ingreso.

Estos resultados se explican fundamentalmente por el efecto positivo que una mayor tasa de interés tiene en el rendimiento de los fondos de ahorro individual. Hay un efecto de segundo orden que puede operar en el sentido opuesto: el factor de descuento utilizado para calcular las cuentas generacionales es menor cuanto mayor es la tasa de interés. Este efecto explica los resultados comentados para las generaciones que se mantienen en el sistema de reparto.

#### La tasa de crecimiento de la productividad del trabajo

Se simularon dos escenarios adicionales con tasas de crecimiento de la productividad del trabajo menor y mayor a la del escenario base (Cuadros 7 y 8, respectivamente). Por lo demás, se mantuvieron los supuestos del escenario base. En la mayor parte de los casos, se obtuvieron menores ganancias y mayores pérdidas para los trabajadores formales cuanto mayor fue la tasa de crecimiento de la productividad del trabajo y del salario real. Sólo las primeras generaciones afectadas por la reforma obtuvieron ganancias ligeramente mayores cuando la tasa de crecimiento del salario fue mayor. Se trata de generaciones que se mantienen enteramente en el sistema de reparto. Las restantes generaciones se benefician menos con la reforma cuanto mayor es el crecimiento de la productividad. Se observa una tendencia similar en ambos géneros y en los distintos niveles de ingreso, salvo en los hombres pertenecientes al quintil de menores ingresos que se benefician más con la reforma cuanto mayor es la tasa de crecimiento del salario real.

La sustitución parcial del pilar de reparto por un pilar de ahorro individual explica estos resultados. A diferencia del sistema de ahorro individual, el sistema de reparto otorga una mejor relación de beneficios a aportes cuanto mayor es la tasa de crecimiento del salario. El sistema no reformado era, por lo tanto, más sensible a esta variable que el nuevo sistema mixto. Como consecuencia, las ganancias que se producen con la reforma en relación con el régimen previo son menores en los escenarios favorables, en que la productividad del trabajo y el salario real crecen a buen ritmo. Como contraparte, la reforma se vuelve relativamente más atractiva para los trabajadores en escenarios de escaso crecimiento del salario. Los hombres del quintil de menores ingresos constituyen una excepción debido a que su jubilación está determinada por la jubilación mínima. La jubilación mínima

<sup>14</sup> Son las generaciones regidas por el título V de la ley 16.713, aplicable a los trabajadores que tuvieran generada la causal jubilatoria y no estuvieran retirados al iniciarse el nuevo régimen. En las simulaciones realizadas, están en esta situación entre cuatro y cinco generaciones. Por razones de presentación, en los cuadros se incluyeron sólo dos generaciones representativas de esta situación, la generación de 1933 en hombres y la de 1940 en mujeres.



CUADRO 7  
 VARIACIÓN DE LAS CUENTAS GENERACIONALES DE TRABAJADORES  
 FORMALES DEBIDO A LA REFORMA. ESCENARIO BASE + TASA DE  
 CRECIMIENTO DE LA PRODUCTIVIDAD IGUAL A 0,4 POR CIENTO ANUAL<sup>a/</sup>

Nivel de ingreso							
Generación	Alto	Alto-medio	Medio	Medio-bajo	Bajo	Promedio	Régimen aplicable (ley 16.713) <sup>c/</sup>
A) Hombres							
1933	-1,6	-0,7	-2,2	-3,6	-7,5	-1,8	titV
1940	-1,6	0,1	-2,1	-3,6	-7,5	-1,7	rt
1950	-0,4	-1,8	-3,8	-3,7	-7,6	-1,1	rt/mixto
1957	-1,1	-2,7	-4,5	-3,0	-6,8	-1,7	mixto
1960	-1,4	-2,3	-4,0	-3,3	-7,0	-1,9	mixto
1970	-2,2	-1,6	-3,6	-4,5	-7,5	-2,4	mixto
1980	-2,2	-1,7	-3,7	-4,5	-7,6	-2,5	mixto
1990	-2,2	-1,6	-3,7	-4,6	-7,9	-2,5	mixto
2000	-2,2	-1,6	-3,7	-4,6	-7,9	-2,5	mixto
Promedio <sup>b/</sup>	-1,6	-1,2	-2,9	-3,6	-6,6	-1,9	
B) Mujeres							
1933	0,0	0,0	0,0	0,0	0,0	0,0	art2
1940	-2,5	-0,3	-0,1	-0,9	-2,4	-1,9	titV
1950	1,0	0,6	0,7	-0,9	-2,4	0,8	rt/mixto
1957	0,5	-0,7	-0,6	-0,4	-0,9	0,1	mixto
1960	0,2	-0,4	-0,3	-0,3	-1,0	0,1	mixto
1970	-0,5	0,4	0,5	-1,6	-1,5	-0,4	mixto
1980	-0,8	0,2	0,4	-2,2	-1,7	-0,7	mixto
1990	-0,8	0,2	0,4	-2,2	-1,7	-0,7	mixto
2000	-0,8	0,2	0,4	-2,2	-1,7	-0,7	mixto
Promedio <sup>b/</sup>	-0,5	0,3	0,4	-1,4	-1,5	-0,4	

Notas:

<sup>a/</sup> Medido en salarios anuales del trabajador a los 55 años de edad. Un valor positivo (negativo) significa que el trabajador pierde (gana) con la reforma, ya que transfiere más (menos) después de la reforma de lo que hubiera transferido sin reforma.

<sup>b/</sup> Promedio de todas las generaciones simuladas (1916 a 2032).

<sup>c/</sup> Régimen aplicable: (i) art2 es el régimen que se aplica a personas que estaban ya jubiladas al entrar en vigencia la ley de reforma, (ii) titV es el régimen que se aplica a personas que sin estar jubiladas tenían generada la causal jubilatoria al iniciarse la reforma, (iii) rt es el llamado "régimen de transición" y se aplica a personas que tenían 40 años o más y no tenían generada la causal jubilatoria al entrar en vigencia la ley, (iv) mixto es el régimen que se aplica a quienes no habían cumplido 40 años al entrar en vigencia la ley.

Fuente: Elaboración propia.

CUADRO 8  
 VARIACIÓN DE LAS CUENTAS GENERACIONALES DE TRABAJADORES  
 FORMALES DEBIDO A LA REFORMA. ESCENARIO BASE + TASA DE  
 CRECIMIENTO DE LA PRODUCTIVIDAD IGUAL A 1,8 POR CIENTO ANUAL<sup>a/</sup>

Nivel de ingreso							
Generación	Alto	Alto-medio	Medio	Medio-bajo	Bajo	Promedio	Régimen aplicable (ley 16.713) <sup>c/</sup>
A) Hombres							
1933	-2,0	-0,8	-2,7	-4,4	-9,2	-2,2	titV
1940	-2,0	0,1	-2,5	-4,4	-9,2	-2,1	rt
1950	-0,2	-2,0	-4,5	-4,5	-9,3	-1,1	rt/mixto
1957	-0,7	-2,8	-4,9	-3,1	-8,1	-1,5	mixto
1960	-1,0	-2,1	-4,2	-3,4	-8,3	-1,6	mixto
1970	-1,6	-0,7	-3,0	-4,2	-8,7	-1,9	mixto
1980	-1,6	-0,8	-3,1	-4,3	-8,7	-1,9	mixto
1990	-1,5	-0,7	-3,1	-4,4	-9,2	-1,8	mixto
2000	-1,5	-0,7	-3,1	-4,4	-9,2	-1,8	mixto
Promedio <sup>b/</sup>	-1,3	-0,6	-2,7	-3,7	-7,8	-1,5	
B) Mujeres							
1933	0,0	0,0	0,0	0,0	0,0	0,0	art2
1940	-3,0	-0,4	-0,1	-1,1	-2,9	-2,3	titV
1950	1,4	0,9	1,0	-1,1	-2,9	1,1	rt/mixto
1957	1,0	-0,4	-0,3	0,0	-0,9	0,6	mixto
1960	0,8	0,1	0,2	0,2	-1,0	0,6	mixto
1970	0,3	1,7	2,0	-0,7	-1,3	0,5	mixto
1980	0,1	1,9	2,0	-1,1	-1,5	0,4	mixto
1990	0,1	1,9	2,0	-1,1	-1,5	0,4	mixto
2000	0,1	1,9	2,0	-1,1	-1,5	0,4	mixto
Promedio <sup>b/</sup>	0,0	1,4	1,5	-0,8	-1,4	0,2	

## Notas:

<sup>a/</sup> Medido en salarios anuales del trabajador a los 55 años de edad. Un valor positivo (negativo) significa que el trabajador pierde (gana) con la reforma, ya que transfiere más (menos) después de la reforma de lo que hubiera transferido sin reforma.

<sup>b/</sup> Promedio de todas las generaciones simuladas (1916 a 2032).

<sup>c/</sup> Régimen aplicable: (i) art2 es el régimen que se aplica a personas que estaban ya jubiladas al entrar en vigencia la ley de reforma, (ii) titV es el régimen que se aplica a personas que sin estar jubiladas tenían generada la causal jubilatoria al iniciarse la reforma, (iii) rt es el llamado "régimen de transición" y se aplica a personas que tenían 40 años o más y no tenían generada la causal jubilatoria al entrar en vigencia la ley, (iv) mixto es el régimen que se aplica a quienes no habían cumplido 40 años al entrar en vigencia la ley.

aumentó con la reforma y debido a su indexación al índice medio de salarios crece más rápido cuanto más rápido crecen los salarios.

## 5. COMENTARIOS FINALES

Los resultados que se presentan en este trabajo indican que la reforma de la seguridad social iniciada en Uruguay en 1996 podría reducir sensiblemente la carga fiscal que recae en generaciones futuras, sin aumentar considerablemente la carga de los trabajadores formales actuales. La clave radica en la eliminación de las importantes transferencias que, al iniciarse la reforma, realizaba el sistema a personas que no habían contribuido. Corresponde advertir que la reducción proyectada en las transferencias irregulares depende de estimaciones propias y ajenas que, si bien han tendido a ser bastante coincidentes, es posible que involucren un considerable margen de error, por la propia naturaleza del tema. El error involucrado en la estimación de los efectos de la reforma sobre los trabajadores formales es probablemente mucho menor.

Si bien el efecto medio de la reforma en las cuentas generacionales de los trabajadores formales es moderado, hay colectivos de trabajadores que sentirán efectos considerables. En el presente trabajo se analizaron los efectos diferenciales según género, nivel de ingreso y generación. Se observó que la reforma beneficia en general a los hombres y a las mujeres de ingresos bajos, pero perjudica a las mujeres de ingresos medios. Este resultado es consistente con un objetivo explícito de los reformadores en el sentido de reducir las diferencias que había en el sistema previo a favor de las mujeres.

Las simulaciones realizadas indican que cabe esperar una mejora de los trabajadores formales de menores y mayores ingresos, con un deterioro relativo de los sectores medios. Los trabajadores de altos ingresos se benefician por la introducción del pilar de ahorro individual, el aumento de las pensiones máximas y la reducción de aportes por encima de cierto nivel. En la otra punta de la distribución, los trabajadores de menores ingresos se benefician por el aumento de las pensiones mínimas y por la bonificación que se otorga a quienes hacen la opción por el pilar de ahorro individual.

Los análisis de sensibilidad indican, en primer lugar, que la reforma resulta tanto más beneficiosa para los trabajadores formales cuanto mayor es la tasa de interés. Este resultado se explica por el rendimiento de los fondos de ahorro individual. En segundo lugar, la reforma es relativamente menos beneficiosa para los trabajadores formales cuanto mayor es la tasa de crecimiento de la productividad del trabajo y del salario. En una economía pujante, la reforma resulta menos imprescindible y sus beneficios son menos significativos. Por último, los beneficios se reducen y los costos aumentan si los trabajadores postergan la edad de retiro después de la reforma. Si bien se redujeron sensiblemente los incentivos al retiro temprano, el pilar de reparto del sistema reformado sigue castigando el retiro tardío.

El análisis presentado en este documento recoge básicamente los impactos distributivos *directos* de la reforma. No considera en cambio algunas de las externalidades positivas que se esperan de una reforma previsional, como ser el mayor desarrollo del mercado de capitales que permite mayor eficiencia en la asignación de la inversión, la mayor eficiencia en el mercado laboral, las oportunidades para rebalancear los distintos tipos de impuestos generales y el efecto de reducir el riesgo de insolvencia fiscal que provocaría una reducción de la prima por riesgo país que se paga sobre los intereses de la deuda pública. De todos modos, la relevancia empírica de tales externalidades está todavía sujeta a controversia y no parece en principio probable que las externalidades alteren en forma significativa los efectos distributivos de la reforma.

ANEXO 1  
LAS CUENTAS GENERACIONALES

Siguiendo a Auerbach y Kotlikoff (1999), se define la cuenta de la generación  $k$  desde el período  $t$  en adelante como:

$$(A1) \quad N_{t,k} = \sum_{s=\kappa}^{k+D} T_{s,k} P_{s,k} (1+r)^{k-s}, \quad \kappa = \max(t, k)$$

$T_{s,k}$  es la transferencia neta (impuestos pagados menos transferencias recibidas) que realiza en el año  $s$  un miembro de la generación nacida en  $k$ .  $P_{s,k}$  es la cantidad de sobrevivientes de la generación  $k$  en el año  $s$ . La tasa de interés es  $r$  y  $D$  es la edad de muerte menos 1. Así definida, la cuenta generacional  $N_{t,k}$  suma las transferencias netas de la generación  $k$  a partir de su nacimiento (o a partir de  $t$ , si la generación nace antes de  $t$ ), descontadas al nacimiento (o a  $t$ , si nace antes de esa fecha).

La restricción presupuestal intertemporal del gobierno puede escribirse en términos de las cuentas generacionales del siguiente modo:

$$(A2) \quad \sum_{k=t-D}^t N_{t,k} + \sum_{k=t+1}^{\infty} (1+r)^{t-k} N_{t,k} = \sum_{s=t}^{\infty} G_s (1+r)^{t-s} - W_t^g$$

El primer sumando en el lado izquierdo de la ecuación agrega las cuentas generacionales de las generaciones existentes en  $t$ . El segundo sumando agrega las cuentas de las generaciones “futuras”.  $G_s$  es el consumo del gobierno en  $s$  y  $W_t^g$  es la riqueza neta del gobierno en  $t$ .

Usualmente se estiman las cuentas de las generaciones “futuras” por residuo, como diferencia entre las obligaciones del gobierno (lado derecho de la ecuación) y las cuentas de las generaciones existentes en el momento inicial (primer sumando del lado izquierdo de la ecuación). Este cálculo permite estimar la carga fiscal de las generaciones futuras implícita en las obligaciones del gobierno y en los pagos netos que realizan las generaciones actuales.

En el presente estudio se evaluó el efecto de la reforma en base a la variación de las cuentas generacionales debida a la reforma. Sea  $DX_t$  la diferencia entre el valor que adopta la variable  $X_t$  dado que se hizo la reforma y el valor que habría adoptado esa misma variable si la reforma no se hubiera implementado. En una economía pequeña y abierta y suponiendo que rige la paridad de tasas de interés, los efectos de la reforma en la restricción presupuestal del gobierno resultan:

$$(A3) \quad \sum_{k=t-D}^t \Delta N_{t,k} + \sum_{k=t+1}^{\infty} (1+r)^{t-k} \Delta N_{t,k} = \sum_{s=t}^{\infty} \Delta G_s (1+r)^{t-s} - \Delta W_t^g = 0$$

En la medida en que la reforma no modificó el consumo público ni la riqueza neta inicial del gobierno, el lado derecho de esta ecuación es cero. Esto implica que la suma de los cambios en las cuentas de las generaciones presentes en  $t$ , cuando se inicia la reforma, es necesariamente igual a menos la suma de los cambios de las cuentas generacionales de las generaciones posteriores.

La reforma de la seguridad social afecta en forma distinta a las generaciones según la edad que los individuos tienen cuando se inicia la reforma. En consecuencia, para evaluar plenamente sus efectos, considerando tanto a las generaciones de transición como a las del sistema maduro, es necesario simular el sistema reformado durante varias décadas. En el caso de la reforma uruguaya, el nuevo sistema estará maduro hacia mediados del siglo. Se consideraron entonces tres conjuntos de generaciones: las que inician su actividad antes de la reforma (1995 y antes), las que inician su actividad entre el primer año de funcionamiento del sistema reformado (1996) y un año en que el nuevo sistema está ya maduro (2050) y las generaciones “futuras” que inician su actividad en el 2051 o después.

Se estimó el cambio de las cuentas de los dos primeros conjuntos de generaciones simulando el modelo MISS con y sin reforma de la seguridad social. Luego, siguiendo el procedimiento usual de la contabilidad generacional, se procedió a calcular el cambio implícito en la cuenta de las generaciones futuras por diferencia:

$$(A4) \quad \sum_{k=2051}^{\infty} (1+r)^{1995-k} \Delta N_{1995,k} = -\sum_{k=1995-D}^{1995} \Delta N_{1995,k} - \sum_{k=1996}^{2050} (1+r)^{1995-k} \Delta N_{1995,k}$$

Finalmente, siguiendo también la metodología estándar (Auerbach y Kottlikoff, 1999), se supuso que las cargas se distribuyen entre las generaciones futuras en proporción al crecimiento del producto ( $\hat{y}$ ):

$$(A5) \quad \Delta N_{t,k} = (1+\hat{y})^{k-2051} \Delta N_{t,2051}, \quad k \geq 2051$$

Operando a partir de estas dos últimas ecuaciones y considerando que la economía es dinámicamente eficiente ( $\hat{y} < r$ ) se obtiene:

$$(A6) \quad \Delta N_{1995,2051} = -\left[ \sum_{k=1995-D}^{1995} \Delta N_{1995,k} + \sum_{k=1996}^{2050} (1+r)^{1995-k} \Delta N_{1995,k} \right] \frac{(r-\hat{y})(1+r)^{2050-1995}}{(1+\hat{y})}$$

donde  $\Delta N_{1995,2051}$  es el cambio provocado por la reforma aprobada en 1995 en la cuenta de la generación que inicia su actividad en 2051, es decir, la generación que nace en 2033.

En el Cuadro 1 se presentan las variaciones en las cuentas generacionales descontadas al año 1995 y expresadas como porcentaje del producto bruto interno de ese año:  $\Delta N_{1995,k} (1+r)^{1995-k} / PBI_{1995}$ ,  $\kappa = \max(1995, k)$ . En los Cuadros 2 a 8 se expresan las variaciones en las cuentas generacionales como proporción del salario (antes de impuestos) a los 55 años de edad del individuo correspondiente.

## ANEXO 2 LOS PERFILES DE INGRESOS LABORALES

Para la estimación de los perfiles de ingresos laborales se utilizó una muestra de los registros administrativos de historia laboral del Banco de Previsión Social. La muestra comprendió a 70.000 trabajadores, con registros mensuales entre abril de 1996 y diciembre de 2004. Bucheli *et al.* (2005) presentan una descripción de la base de datos.

Se estimaron los perfiles de ingresos laborales utilizando una variante de la metodología indicada en Bosworth *et al.* (1999) y Peña (2005). Se estimaron regresiones del siguiente tipo:

$$(A7) \quad \frac{w_{it}}{W_t} = \beta_0 + \beta_1 * edad_{it} + \beta_1 * edad_{it}^2 + \beta_2 * edad_{it}^3 + u_i + \varepsilon_{it}$$

donde  $w_{it}$  es el salario nominal para el individuo  $i$  en el período  $t$ ;  $W_t$  es el salario nominal promedio para el período  $t$  de todos los individuos pertenecientes a la categoría de análisis;  $edad_{it}$  es la edad en años del individuo  $i$  en el período  $t$ ;  $u_i$  es un efecto fijo individual inobservable para el individuo  $i$  (potencialmente correlacionado con la edad) y  $\varepsilon_{it}$  es un término de error *iid*.

Se hicieron estimaciones por efectos fijos y aleatorios. Se evaluó en cada caso la posibilidad de aprovechar la ganancia de eficiencia asociada a la estimación por efectos aleatorios utilizando el test de Hausman. En la mayor parte de las corridas se rechazó la hipótesis nula de que los efectos individuales están incorrelacionados con los regresores, por lo cual se utilizó predominantemente efectos fijos. Por razones de espacio, no se presentan aquí las 20 estimaciones realizadas, pero se encuentran disponibles para quien las solicite.

Una vez hechas las estimaciones se calcularon los salarios por edad, multiplicando los salarios relativos predichos por el salario medio en cada período.

## REFERENCIAS

- Auerbach, A. y L. Kotlikoff (1987), *Dynamic Fiscal Policy*, Cambridge University Press, 196 p.
- Auerbach, A. y L. Kotlikoff (1999), "The Methodology of Generational Accounting". In: *Generational Accounting around the World* edited by A. Auerbach, L. Kotlikoff and W. Leibfritz. National Bureau of Economic Research. The University of Chicago Press.
- Auerbach, A.; L. Kotlikoff y W. Leibfritz (1999), *Generational Accounting around the World*. National Bureau of Economic Research. The University of Chicago Press.
- Auerbach, A.; J. Gokhale y L. Kotlikoff (1994), "Generational Accounting: A Meaningful Way to Evaluate Fiscal Policy". *Journal of Economic Perspectives*. Winter; 8(1): 73-94.

- Bosworth, B.; G. Burtless y E. Steuerle (1999), "Lifetime Earnings Patterns, The Distribution Of Future Social Security Benefits, And The Impact Of Pension Reform". CRR WP 1999-06, *Center for Retirement Research* at Boston College.
- Bucheli, M. (1999), "Los efectos de la reforma en las cuentas generacionales". En: *La reforma de la seguridad social en Uruguay: efectos macroeconómicos y mercados de capitales* editado por Forteza. Universidad de la República. Uruguay.
- Bucheli, M.; N. Ferreira-Coimbra; A. Forteza y I. Rossi (2005), "El acceso a la jubilación o pensión en Uruguay: ¿cuántos y quiénes lo lograrían?". CEPAL y Departamento de Economía, FCS, Universidad de la República. Uruguay.
- Camacho, L. (1997), *Elementos generales del financiamiento del nuevo sistema previsional uruguayo*, BPS, Uruguay.
- Caristo, A. y A. Forteza (2003), *El déficit del Banco de Previsión Social y su impacto en las finanzas del gobierno*. Facultad de Ciencias Sociales. Universidad de la República, Uruguay.
- Durán, V.; G. Poplavski y A. Vernengo (2003), "Impacto de la reforma de la seguridad social en la edad de retiro y la evasión". Trabajo monográfico de la Facultad de Ciencias Económicas y de Administración, Universidad de la República, Uruguay.
- Fehr, H. y L. Kotlikoff (1999), "Generational Accounting in General Equilibrium". In: *Generational Accounting around the World*. A. Auerbach; L. Kotlikoff y W. Leibfritz, National Bureau of Economic Research. The University of Chicago Press.
- Forteza, A. (1999a), "Los efectos de la reforma en el ahorro". En: *La reforma de la seguridad social en Uruguay: efectos macroeconómicos y mercados de capitales*, editado por A. Forteza. Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República. Uruguay.
- Forteza, A. (1999b) "Un modelo de simulación de la reforma de la seguridad social en Uruguay". En *La reforma de la seguridad social en Uruguay: efectos macroeconómicos y mercados de capitales*, editado por A. Forteza. Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República. Uruguay.
- Forteza, A. (1999c), "Los efectos fiscales de la reforma". En *La reforma de la seguridad social en Uruguay: efectos macroeconómicos y mercados de capitales*, editado por A. Forteza. Departamento de Economía, Facultad de Ciencias Sociales, Universidad de la República. Uruguay.
- Forteza, A. (2002), "Un modelo macroeconómico de simulación para el Banco de Previsión Social". Facultad de Ciencias Sociales. Universidad de la República, Uruguay. <http://decon.edu.uy/publica/2002/Doc1202.pdf>.
- Forteza, A. (2003), "Seguridad Social y competencia política". En: *Economía Política en Uruguay. Instituciones y actores políticos en el proceso económico*, editado por Aboal y Moraes. Cinve-UDELAR-Trilce, Uruguay.
- Kotlikoff, L. y B. Raffelhuschen (1999), "Generational Accounting around the Globe". *American Economic Review*. May; 89(2): 161-66.
- Noya, N.; A. Fernández; F. Filgueira; S. Laens; J. A. Moraes y C. Moreira (1999), "Efectos económicos y políticos de la reforma de la seguridad social en Uruguay". *Cinve*. Uruguay.



- Peña, G. (2005), "Estimación de perfiles salariales: Una aproximación a partir de registros administrativos del Sistema de Seguridad Social". *dECON*, Universidad de la República, Uruguay.
- Raffelhuschen, B. (1999), "Generational Accounting in Europe". *American Economic Review* 89(2): 167-70, May.

## ¿PUEDE EL DISEÑO DE UN TORNEO DEPORTIVO AFECTAR SU ASISTENCIA?\*

GIORGO SERTSIOS  
*Universidad de los Andes*

*Tournament design variables -such as the number of teams and matches - affect attendance. This paper provides evidence of how design variables have affected attendance in the Chilean football league. We estimate the annual demand for stadium tickets with tournament design variables among the regressors. The results suggest that moving towards the optimal tournament structure implies fewer teams and more matches. For example, a tournament with 14 teams and 3 rounds of matches (totaling 273 games) -as opposed to the current 16-team, 2-round league structure (240 games)- would increase attendance by 28%.*

*JEL: L83, C22*

*Keywords: Football demand, annual attendance, tournament design.*

### 1 INTRODUCCIÓN

Embarcarse en grandes cambios institucionales requiere visión o desesperación por parte de los tomadores de decisiones. Las autoridades del fútbol chileno han instituido tantos cambios en el diseño de la primera división A en los últimos 70 años que parece probable que ambas motivaciones hayan estado presentes en sus decisiones. De hecho, entre 1935 y 2005 la cantidad de equipos participantes y partidos jugados correspondientes a la primera división A cambiaron en 16 y 27 ocasiones, respectivamente.

Este comportamiento revela que los organizadores han tenido una insatisfacción constante con el diseño del torneo: su diseño óptimo les ha resultado esquivo. Esto se puede deber a que las nuevas propuestas de torneos son copias de campeonatos de fútbol que han sido exitosos en otros países. Lamentablemente, a la hora de implementar esas propuestas, no se considera que el mercado

\* Agradezco a Alexander Galetovic, Manuel Hermosilla, Iván Marinovic, William Mullins, Carlos Rodríguez y al árbitro anónimo sus valiosos aportes.  
E-mail: gsertsios@uandes.cl

chileno es muy distinto en cuanto a tamaño, ingreso y otras variables, al de los países en donde esos formatos de torneos resultaron exitosos.

La iteración entre diferentes diseños del torneo de fútbol resalta el hecho de que éstos han sido impuestos exógenamente. De hecho, si el fútbol fuese una industria no regulada, sin barreras a la entrada o a la salida, tanto el número de empresas (equipos) como la cantidad producida (partidos) tenderían a su nivel óptimo. Sin embargo, como en Chile los clubes deportivos en su mayoría no son empresas privadas y el formato del torneo es impuesto exógenamente, mediante el sistema de prueba y error, no necesariamente se tiende hacia su diseño óptimo.

El aporte de este trabajo es proveer evidencia acerca de cómo el diseño del torneo de fútbol en Chile ha afectado su atractivo, medido como asistencia anual a los estadios, permitiendo inferir cuándo el diseño del torneo ha estado más cerca de su óptimo<sup>1</sup>. Para este propósito, la demanda anual a los estadios de la primera división A es estimada, ocupando entre los regresores las variables de diseño del torneo.

La literatura de reestructuración de torneos deportivos es la más cercana al tema de estudio del presente trabajo y además provee la adecuada contextualización, porque trata acerca de cómo un nuevo diseño de torneo afecta la demanda a los estadios. Cairns (1987) estima la demanda por partido para tres equipos en la liga de fútbol escocesa para las temporadas 1971-80, controlando por el cambio de diseño que tuvo esa liga en 1975. Burkitt y Cameron (1992) realizaron un análisis similar para la liga de rugby inglesa, estimando la demanda por partido entre 1966-90 usando datos de panel y controlando por el cambio de diseño de la liga de 1973. Finalmente Dobson *et al.* (2001) extendieron el estudio de Burkitt y Cameron generando series simuladas de asistencias ante diferentes escenarios de estructuras de la liga de rugby inglesa. Sin embargo, ninguno de estos trabajos estima los efectos marginales de las variables de diseño sobre la asistencia. Sólo estiman los efectos de reestructuraciones completas.

Este trabajo es único por dos motivos. El primero es que los datos del torneo de fútbol chileno presentan un alto grado de variación en el número de equipos y número de partidos, permitiendo la inclusión directa de ambas variables de diseño como variables explicativas en la estimación de tipo series de tiempo<sup>2</sup>. De hecho, en comparación con la liga inglesa, escocesa, italiana y española, la primera división A del fútbol chileno es la que ha cambiado más veces el número de equipos y es la segunda que más ha cambiado la cantidad de partidos jugados, después de la escocesa. Es más, la mayor parte de esta variación ha sido dentro del mismo tipo de torneo (todos contra todos)<sup>3</sup>, lo que provee un experi-

<sup>1</sup> Palomino y Rigotti (2000) señalan que la organización institucional de los equipos de fútbol profesionales en Europa tienen como objetivo maximizar la demanda por el deporte. En Chile, los equipos de fútbol se fundaron bajo las mismas bases que sus pares de Europa, por lo que es razonable pensar que los objetivos institucionales sean los mismos.

<sup>2</sup> Bird (1982) también estima la demanda anual por tickets al estadio mediante series de tiempo, pero él analiza sólo el efecto de variables económicas (como precio de los tickets e ingreso de los hinchas) sobre la asistencia y no considera los efectos de las variables de diseño sobre la asistencia.

<sup>3</sup> Otro tipo de torneo es jugar por grupos en una primera fase, o que haya más de un torneo en el año.

mento natural que permite medir el impacto marginal de los cambios de estas variables sobre la asistencia.

La segunda diferencia es metodológica: a diferencia del resto de la literatura de economía deportiva, que ha estimado la demanda por partidos y ha intentado resolver el problema de la simultaneidad entre los precios de las entradas y la asistencia vía variables instrumentales u omisión de la variable precio, este estudio estima la demanda anual de tickets mediante una ecuación reducida que elude el problema de la simultaneidad y permite obtener los parámetros estructurales.

Este artículo se estructura de la siguiente manera: en la sección 2 se describe la ecuación de demanda y se realizan pruebas empíricas a su forma funcional. En la sección 3 se exponen los resultados. Finalmente, en la sección 4 se concluye.

## 2 ESPECIFICACIÓN DEL MODELO

### 2.1 Modelo, variables y datos

La demanda anual del torneo de la primera división A del fútbol chileno es estimada tomando como variable dependiente la asistencia anual ( $A_t$ ) y como variables explicativas el precio promedio de las entradas ( $P_t$ ) y un vector de variables exógenas ( $Z_t$ ), tal como se plantea en la ecuación 1.

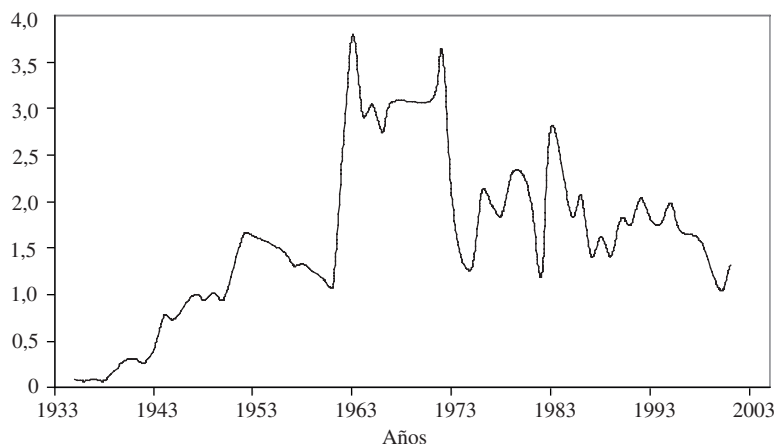
$$(1) \quad A_t = A_t(P_t, Z_t)$$

Las variables exógenas  $Z_t$  se pueden separar en tres categorías: variables económicas, variables de diseño del torneo y variables de control.

Dentro de las variables económicas se incluyen el ingreso promedio de la población expresado en millones de pesos de 1996, la población total de Chile y dos variables que tratan de enfocar los efectos de los sustitutos de asistir al estadio. Estas variables son D75 y DTV. D75 es una variable dicotómica que toma valor 0 hasta 1974 y 1 de 1975 en adelante. Esta variable intenta capturar los efectos de la apertura económica de Chile a partir de ese año que se dio como consecuencia de cambios políticos, ya que junto con la mayor apertura, llegaron también gran parte de nuevas actividades recreacionales que sustituyeron al fútbol. DTV es una variable que toma valor 1 cuando el torneo de fútbol fue televisado y 0 en caso contrario.

No se considera ninguna tendencia para explicar la trayectoria de la asistencia, ya que no existe soporte teórico para hacerlo y porque la inspección visual de los datos de asistencia no determinan ningún patrón inidentificable que relacione la asistencia anual al estadio con el paso del tiempo como se observa en el Gráfico 1.

GRAFICO 1  
ASISTENCIA ANUAL AL ESTADIO A LO LARGO DEL TIEMPO  
(en millones de personas)



Las variables de diseño son el número de equipos, la cantidad de partidos jugados en el torneo y el número de equipos que descendieron a una división inferior por su bajo desempeño.

El efecto de estas variables sobre la calidad del torneo es derivado en el Apéndice A ocupando un modelo de teoría de juegos con equipos heterogéneos. La intuición es que si la calidad del torneo aumenta, la asistencia a los estadios también lo hará, *ceteris paribus*. El modelo predice que aumentar el número de equipos en el torneo, manteniendo constante la cantidad de partidos jugados, la inversión de los equipos en calidad –medida como el gasto planilla– se reduce, ya que la probabilidad de descender disminuye. A su vez, si el número de partidos aumenta, manteniendo el número de equipos, la inversión en calidad de los equipos aumenta (contratan jugadores de mejor calidad), porque más partidos permite hacer un uso más intensivo de la inversión, rentabilizándola. Por último, en el modelo también se encuentra que la posibilidad de descender induce una mayor inversión en calidad.

Sin embargo, los resultados del modelo no implican que el diseño óptimo del torneo deba tener dos equipos e infinitos partidos, ni que un gran número de equipos deba ser relegado cada temporada. Esto es porque el modelo no incluye algunos factores que influncian la demanda, al no poder ser modelados adecuadamente. De hecho, un torneo con dos equipos atraería muy poco público porque rápidamente se tornaría repetitivo y monótono. A su vez, un torneo con varios cientos de partidos por temporada saturaría rápidamente el mercado.

Un torneo en el cual descienda una gran proporción de equipos enfrentaría un problema diferente: los equipos con grandes bases de apoyo descenderían más seguido, privando a la primera división A de una fracción significativa de sus consumidores naturales cada temporada. Esto también reduce el atractivo del torneo indirectamente, ya que la oportunidad de generar partidos entre clásicos

cos rivales sería imposible en algunos años porque uno de esos equipos puede haber descendido la temporada anterior. A este inconveniente le llamaremos rotación ineficiente de equipos.

Por ende, los resultados a presentar en este trabajo deben considerar tanto los efectos de calidad, desarrollados en el Apéndice A, como los factores recién mencionados a la hora de ser interpretados: un torneo tendría el número de equipos óptimo si la reducción de un equipo aumenta el atractivo del torneo, mediante el aumento en calidad, por el mismo monto que la disminuiría la reducción en variedad que produciría jugar con un equipo menos. Naturalmente, este argumento se sostiene para las otras dos variables de diseño.

Finalmente, el último set de variables del vector  $Z_t$ , son las variables de control, dentro de las cuales se incluye la dicotómica D7374, que toma valor 1 en los años 73 y 74 y 0 en otro caso. Esta variable trata de aislar el efecto de la inestabilidad política y social de Chile en esos años. También se considera la *dummy* tres puntos (D3P) que quiere probar si la norma internacional de asignar tres puntos al ganar en vez de dos contribuyó en algo al atractivo y, por ende, a la asistencia del fútbol de Chile. Por último, se considera como variable de control un indicador de daños a la propiedad privada denominado violencia. Este indicador se ocupa como instrumento (*proxy*) de los desmanes ocurridos en los estadios e intenta capturar el mayor precio implícito que pagan los individuos al asistir a los estadios de Chile a partir de los años noventa, por los desmanes frecuentes que allí ocurren. La fuente de todas las variables se detallan en el Apéndice B.

La base de datos tiene 71 observaciones, pero 15 torneos se han jugado con una estructura distinta a todos-contra-todos y a su vez muy distinta entre sí. Por esto, para aislar el efecto de las variables de diseño, número de equipos, partidos jugados y descenso, de otros efectos de diseño de campeonato, se consideran solamente los 56 torneos que ha tenido el formato “todos-contra-todos”, ya sea dos o más ruedas<sup>4</sup>.

## 2.2 Especificación econométrica

Para estimar la ecuación (1), hay que resolver el problema de la endogeneidad de los precios con la cantidad de personas que asisten al estadio, ya que ambas variables se determinan simultáneamente.

En la literatura de torneos deportivos este problema se repite con frecuencia y se han tomado básicamente tres posiciones: la primera de ellas ha sido ignorar el problema y considerar la variable precio como exógena conllevando a estimadores de parámetros inconsistentes; la segunda alternativa ha sido omitir la variable precio, lo que implica sesgar los parámetros que estuvieran relacionados con dicha variable; la tercera alternativa ha sido ocupar variables instrumentales, metodología correcta, pero en la mayoría de las ocasiones los instrumentos adecuados escasean, haciendo difícil la implementación de esta técnica.

<sup>4</sup> En esta muestra no hay observaciones de una sola rueda.

En este trabajo se intenta solucionar la endogeneidad de los precios mediante la estimación reducida de un modelo estructural. Simultáneamente con la elección del modelo reducido se realizará la elección de la forma funcional que debiese tener la ecuación 1. La técnica se describe a continuación.

Si se elucubra acerca de la forma en que se fijan los precios promedios anuales de las entradas ( $P_t$ ) podemos pensar que los clubes, en su conjunto, actúan con algún grado de colusión  $\theta$ . Si  $\theta = 1$  la industria se comporta de manera monopólica, si  $\theta = 0$  se comporta de manera competitiva y si  $0 < \theta < 1$  lo hace de alguna manera intermedia.

De esta manera podemos obtener la regla de fijación de precios según la técnica descrita por Bresnahan (1982).

$$(2) \quad P_t + \frac{dP_t}{dA_t} * A_t * \theta = C$$

Donde  $C$  es el costo marginal de que una persona presencie un encuentro.  $C$  puede ser el costo del papel de la entrada, la limpieza necesaria por asistente, etc., ya que todos los grandes costos de los torneos de fútbol no dependen del número de personas que asisten al estadio.

Se esgrimen tres hipótesis acerca de la forma funcional de la ecuación (1).

$$H_0: \log(A_t) = \alpha_0 + \beta_0 \log(P_t) + \gamma_0 \log(Z_t)$$

$$H_1: \log(A_t) = \alpha_1 + \beta_1 P_t + \gamma_1 Z_t$$

$$H_2: A_t = \alpha_2 + \beta_2 P_t + \gamma_2 Z_t$$

Donde  $\gamma_0, \gamma_1, \gamma_2$  son vectores de parámetros que acompañan a las variables exógenas.

La fijación de precios para las especificaciones  $H_0, H_1$  y  $H_2$  se derivan de la ecuación (2), quedando:

$$(3) \quad P_t = \frac{C}{\left[1 + \frac{\theta}{\beta_0}\right]}$$

para la especificación  $H_0$

$$(4) \quad P_t = C - \frac{\theta}{\beta_1}$$

para la especificación  $H_1$

$$(5) \quad P_t = C - \frac{\theta * A_t}{\beta_2}$$

para la especificación  $H_2$ .

Reemplazando las ecuaciones de fijación de precios en las especificaciones de demandas originales se llega a los siguientes modelos reducidos:

$$H_0': \log(A_t) = \phi_0 + \gamma_0 \log(Z_t)$$

$$H_1': \log(A_t) = \phi_1 + \gamma_1 Z_t$$

$$H_2': A_t = \phi_2 + \gamma_3 Z_t$$

Donde,

$$\phi_0 = \alpha_0 + \beta_0 \log\left(\frac{C}{1 + \frac{\theta}{\beta_1}}\right), \phi_1 = \alpha_1 + \beta_1 * C - \theta, \phi_2 = \alpha_2 + \frac{\beta_2 * C}{1 + \theta}, \gamma_3 = \frac{\gamma_2}{1 + \theta}$$

Por lo tanto si se estima por *OLS*  $H_0'$  y  $H_1'$ , se obtienen los parámetros de interés  $\gamma_0$  y  $\gamma_1$  de manera directa, ya que cualquiera sea el grado de colusión de la industria es absorbido por la constante. Sin embargo, si la forma funcional de la demanda es  $H_2'$ , no se pueden obtener los verdaderos parámetros estimando  $H_2'$ , ya que los parámetros obtenidos por *OLS* habría que multiplicarlos por  $(1 + \theta)$  y el valor de  $\theta$  es desconocido.

Por este motivo, primero se contrasta la especificación  $H_0'$  con  $H_1'$ , ya que son las únicas formas funcionales bajo las cuales se puede obtener los parámetros de interés insesadamente. El contraste de estas formas funcionales se realiza con el *test J* de Mackinnon et al. (1983) y consiste en comparar qué forma funcional es la que mejor describe los datos.

Primero se corren las regresiones  $H_0'$  y  $H_1'$  incluyendo todas las variables explicativas que han sido descritas<sup>5</sup>, obteniendo los *fitted values* para ambas especificaciones. A la diferencia de los valores *fitted* en cada momento del tiempo se le llama  $W_t$ . Posteriormente se vuelve a correr la regresión  $H_0'$ , pero incluyendo como variable explicativa a  $W_t$ . La idea es ver si esta variable es estadísticamente significativa. Si lo es, quiere decir que la especificación  $H_1'$  tiene un aporte significativo a la estimación  $H_0'$  y por lo tanto no se puede rechazar la hipótesis de que  $H_1'$  es la verdadera forma funcional. Si  $W_t$  no es significativa se rechaza la hipótesis de que  $H_1'$  sea la verdadera forma funcional. Posteriormente se realiza el mismo análisis corriendo  $H_1'$  en función de sus variables explicativas y  $W_t$ .

El *test J* de Mackinnon rechaza la hipótesis nula de que  $H_1'$  sea la mejor especificación, pero no rechaza que  $H_0'$  lo sea. De hecho, cuando se corre  $H_0'$  incluyendo  $W_t$  como variable explicativa, ésta tiene un *test t* de  $-1,25$ , lo que permite rechazar la hipótesis nula de  $H_1'$  sea la forma funcional verdadera. Sin embargo, cuando se corre  $H_1'$  incluyendo  $W_t$  como variable explicativa, ésta tiene un *test t* de  $6,63$ , lo que implica hay algo de  $H_0'$  que ayuda a explicar  $H_1'$ .

Para no descartar al modelo lineal sólo porque no se pueden recuperar sus parámetros estimando su forma reducida, también se contrasta  $H_0'$  con  $H_2'$ . A la diferencia entre estos valores predichos (*fitted*) se le denomina  $V_t$ <sup>6</sup>. Cuando  $V_t$  es

<sup>5</sup> No se incluye población urbana, porque es intercambiable con población.

<sup>6</sup> *Vt* es la diferencia de valores *fitted* expresados en la misma escala, es decir, en logaritmos cuando se vuelve a correr  $H_0'$  y en niveles cuando se vuelve a correr  $H_2'$ .



incluida en la estimación de  $H_0'$  tiene un *test t* de  $-0,17$ , mientras que si es incluida en la estimación de  $H_2'$  tiene un *test t* de  $2,57$ , encontrando mayor evidencia a favor de que  $H_0'$  es la verdadera forma funcional de la demanda, ya que aporta significativamente sobre la estimación de  $H_2'$ , pero no sucede lo mismo al contrario.

En conclusión, se estima la demanda asumiendo que la verdadera forma funcional es del tipo doble logarítmico (*log-log*), ya que es la que mejor se ajusta a los datos y además permite obtener de manera directa los verdaderos parámetros que acompañan a las variables independientes<sup>7</sup>.

Antes de finalizar esta sección, un último problema debe ser considerado: en toda estimación de demanda de torneos deportivos, puede haber partidos en donde la demanda observada no es la real, ya que puede existir restricción de capacidad en los recintos deportivos. Sin embargo, este fenómeno es poco común en Chile, por lo que no se considera como relevante en la estimación.

### 3 RESULTADOS

#### 3.1 Resultados generales

Antes de exponer los resultados, dos puntos deben ser aclarados. El primero, es que no existe correlación espuria en la regresión; el segundo, es la ausencia de endogeneidad en las variables de diseño.

Cuando se corre una regresión de serie de tiempo, una de las preocupaciones principales es que la relación entre las variables no sea espuria. Para este propósito, el primer paso es analizar si las variables a ocupar son estacionarias. En el presente trabajo se encuentra que las variables asistencia, ingreso, población, equipos y partidos son no estacionarias. En consecuencia, se aplicó el *test* de cointegración de Engle-Granger. La hipótesis nula de no cointegración fue rechazada de acuerdo a los valores críticos reportados por Mac Kinnon (1991)<sup>8</sup> con un nivel de significancia del 5%. Por ende, a pesar de que las variables ocupadas no son estacionarias, no existe correlación espuria.

Habiendo resuelto el tema de la presunta relación espuria entre las variables, queda otra preocupación importante: el hecho que la calidad del torneo pueda ser pensada como endógena con la asistencia. Esto implicaría que las variables de diseño, que son *proxies* de la calidad, también pueden tener el riesgo de ser endógenas. Sin embargo, es sabido que estas variables son im-

<sup>7</sup> Estos resultados son válidos, porque como se verá más adelante, no hay problemas de correlación espuria ni de endogeneidad entre las variables. Además, se ocupó el método de Newey-West para obtener estimadores más eficientes.

<sup>8</sup> Se corrió una regresión simple incluyendo todas las variables explicativas y se aplicó el *test* Dickey-Fuller a sus residuos. Este *test* no incluyó diferencias rezagadas de los residuos, porque no había evidencia de autocorrelación en su proceso. El *test t* encontrado para el error rezagado fue de  $-6.5$ , el cual es mayor, en valor absoluto, que el valor de rechazo:  $-5.4$ .

puestas exógenamente por los organizadores del torneo y que no sólo contienen efectos de calidad, sino también de valoración de los consumidores por la variedad, la saturación de la demanda y la rotación ineficiente. Estos factores disminuyen la posibilidad de endogeneidad en las variables de diseño, al menos teóricamente.

Teniendo claro que teóricamente las variables de diseño pueden o no ser endógenas con la variable dependiente, se realiza una versión del *test* de Hausman y se encuentra evidencia a favor de la no endogeneidad. El *test* es como sigue. Primero, la asistencia se regresa contra todas las variables de las cuales no se sospecha endogeneidad, es decir, todas menos las de diseño. De esta regresión obtenemos la serie de errores  $U_t$ . Después, la variable equipos se regresa contra asistencia y  $U_t$ . Bajo la hipótesis nula de no endogeneidad entre asistencia y equipos, el parámetro que acompaña a  $U_t$  debiese ser cero, ya que los estimadores serían consistentes. El resultado es que el *test t* encontrado para el parámetro que acompaña a  $U_t$  es  $-1,1$ , cayendo por debajo, en valor absoluto, del valor de rechazo de significancia estadística al 5%,  $-2$ . Esto implica que la no endogeneidad no puede ser rechazada. Este mismo procedimiento se repite ocupando como variables dependientes las variables partidos y descenso con similares resultados. Los *test t* encontrados fueron de  $-1,29$  y  $-1,78$ , respectivamente. En consecuencia, la estimación de demanda por mínimos cuadrados ordinarios del modelo reducido ocupando como variables explicativas las variables de diseño produce estimadores consistentes.

Ahora que se ha demostrado que los resultados mediante mínimos cuadrados ordinarios de  $H_0'$  son consistentes y no espurios, los resultados de  $H_0'$  se exponen en la Tabla 1. En la columna 1 se detallan los parámetros de la estimación de demanda ocupando todas las variables explicativas mencionadas en la sección 2; en la columna 2 se muestran los *test t* correspondientes. La columna 3 corresponde a los parámetros estimados de la mejor especificación de la demanda, es decir, eliminando las variables que no son individual o colectivamente significativas al 10%; la columna 4 muestra los *test t* de estos parámetros.

### 3.2 Resultados de las variables económicas

El coeficiente de la variable ingreso es estadísticamente significativo y levemente mayor que uno, lo que sugiere que la asistencia al estadio es un bien normal. Aunque este resultado se contrapone a la elasticidad ingreso negativa encontrada por Bird (1982), la evidencia no es necesariamente contradictoria. Bird estimó la demanda por partidos de fútbol para Inglaterra, un país con mucho mayor ingreso per cápita que Chile. Por ende, es posible que en Inglaterra el fútbol haya pasado de ser un bien normal a uno inferior a medida que el ingreso fue aumentando y otras actividades de ocio más caras desplazaron al fútbol a altos niveles de ingreso.

Si nos enfocamos en la variable población, nos percatamos que su crecimiento ha ocasionado aumentos en demanda que no son estadísticamente signi-

TABLA 1  
RESULTADOS DE LAS ESTIMACIONES

Variable dependiente: Asistencia	1	2	3	4
<b>Variab</b>	<b>C</b>	<b>T</b>	<b>C</b>	<b>T</b>
<i>Variab</i>				
Log(Ingreso)	1,05**	2,06	1,37***	5,80
Log(Población)	0,56	0,85		
D75	-0,47***	-2,88	-0,27***	-4,98
DTV	0,05	0,56		
<i>Variab</i>				
Log(Equipos)	-0,75*	-1,93	-0,33*	
Log(Partidos)	1,86***	12,12	1,76***	13,55
Descenso	0,04	1,01		
<i>Variab</i>				
Log(Violencia)	-0,98***	-5,50	-0,97***	-4,90
D3P	-0,05	-0,60		
D7374	-0,60***	-4,90	-0,50***	-6,20
C	7,02	0,71	15,30***	8,10
R <sup>2</sup>	0,968		0,966	
R <sup>2</sup> ajustado	0,961		0,962	

\*, \*\*, \*\*\* Variables estadísticamente significativas al 10%, 5% y 1% respectivamente.

Fuente: Elaboración propia.

ficativos. La correlación alta entre el índice de violencia y la variable población podría ser la causa de la no significancia de esta última, sin embargo, esto fue descartado mediante un test de significancia conjunta. También se pensó que la población relevante para la estimación podría estar mal definida y se reemplazó la población total por la población urbana, ya que los equipos de fútbol sólo pertenecen a sectores altamente urbanizados. Sin embargo, no se encontraron cambios importantes en el signo ni en la significancia, por lo que también se descartó esta hipótesis. Por último, se pensó que la falta de significancia estadística de la población se debe a que ésta crece a una tasa similar al costo de oportunidad de ir al estadio, y como este último fenómeno no puede ser cuantificado, la población estaría enfrascando ambos. Lamentablemente, no se puede verificar esta hipótesis, pero parece ser la explicación más razonable.

La variable dicotómica D75 es estadísticamente significativa y tiene un gran efecto negativo sobre la asistencia, como se esperaba. Esto sugiere que la apertura internacional de Chile a mediados de los setentas redujo la asistencia, mediante un influjo de bienes sustitutos del fútbol.

La magnitud del efecto de D75 sobre asistencia puede ser obtenida de la columna 3 de la Tabla 1. A la hora de la interpretación, hay que considerar que un cambio de D75 de 0 a 1 no es marginal:

$$(6) \quad \frac{Asistencia(D75 = 1) - Asistencia(D75 = 0)}{Asistencia(D75 = 0)} = (e^{-0,27} - 1) = -0,24$$

Ocupando la asistencia promedio del período 1935-1974 en la ecuación (6) (1,45 millones anuales) se concluye que la apertura internacional de mediados de los 70 redujo la asistencia aproximadamente en 348.000 espectadores, capturando gran parte del declive en asistencia que se observa en el Cuadro 1.

Finalmente, la variable dicotómica DTV, que controla aquellos años en que el torneo de fútbol fue televisado, no es estadísticamente significativa. Esto no debe sorprender, ya que en el 2000 sólo el 25% de los hogares chilenos tenía acceso a televisión por cable<sup>9</sup> y no todos los planes de suscripción al cable permiten el acceso al torneo de fútbol. En consecuencia, el restringido acceso a la transmisión de partidos conlleva a efectos limitados sobre la asistencia.

### 3.3 Resultados de variables de diseño

En la columna 2 de la tabla 1 se observa que el coeficiente de la variable equipos es negativo y estadísticamente significativo al 10%. Más aún, su *valor-p* es levemente mayor que 5% (5,5%). Como se argumentó en la sección 2.1, una reducción en el número de equipos conlleva a dos efectos contrapuestos sobre la asistencia: uno positivo (porque menos equipos llevan a mayor inversión en calidad) y otro negativo (porque se reduce la variedad de partidos). El coeficiente estimado indica que, en la muestra, el efecto positivo fue más fuerte que el negativo: una reducción en 1% en el número de equipos aumenta la asistencia en 0,33%.

La variable partidos es positiva y estadísticamente significativa a un nivel de 1%. Esto también implica que el efecto positivo sobre la asistencia propiciado por la mayor calidad sobrepasa el efecto negativo de la saturación del mercado al aumentar la cantidad de partidos. Un aumento de 1% en el número de partidos lleva a un incremento de 1,76% en la asistencia.

No obstante, esto no implica que el número de equipos óptimo deba ser dos y que se jueguen infinitos partidos, como se mencionó en la sección 2.1. De hecho, tales predicciones están significativamente fuera del rango de la muestra que se ocupó para estimar los parámetros, por lo que no se puede hacer ese tipo de inferencias. Sin embargo, no todas las predicciones son imposibles: si se analizan cambios sobre las variables de diseño en torno a sus medias muestrales, se pueden obtener los efectos de esos cambios sobre la asistencia con mayor confianza<sup>10</sup>.

<sup>9</sup> De acuerdo al Instituto Libertad y Desarrollo (2001).

<sup>10</sup> Otro aporte sería estimar el número óptimo de partidos y equipos incorporando términos de segundo o más orden de estas variables. Sin embargo, cuando se intentó realizar esto se llegó a que esos términos de mayor orden no eran estadísticamente significativos, por lo que no se siguió con esta metodología.

En esta línea, los resultados sugieren que un cambio en el diseño del torneo de 16 equipos, 2 ruedas y 240 partidos –el diseño más común en los noventa– a un diseño con 14 equipos, 3 ruedas y 273 partidos aumentaría la asistencia anual en 28%<sup>11</sup>. Un aumento potencial de esta magnitud merece especial atención y por ende se detalla a continuación.

Este cambio reduciría el número de equipos en dos (12,5%) y añadiría 33 partidos (13,75%) por temporada. El cambio en asistencia es calculado ocupando los coeficientes estimados:

$$\% \Delta \text{Asistencia} = -0,33(-12,5\%) + 1,76(13,75\%) = 28\%$$

Este experimento ocupa números de equipos y partidos que no se desvían mucho de sus medias muestrales. El reducir el número de equipos de 16 a 14 implica un cambio de  $(\bar{X} + 0,29\hat{\sigma})$  a  $(\bar{X} - 0,28\hat{\sigma})$ , mientras que el aumento de 240 a 273 partidos implica un cambio de  $(\bar{X} + 0,19\hat{\sigma})$  a  $(\bar{X} + 0,55\hat{\sigma})$ .

Aunque el número de equipos óptimo puede ser menor a 14 y el número de ruedas mayor a 3 (para que haya más partidos) el diseño óptimo es imposible de determinar con esta muestra. Lo único que se puede decir es que el diseño con 14 equipos, 3 ruedas y 273 partidos parece apuntar en la dirección adecuada en pos de aumentar la asistencia. La intuición es que al haber menos equipos, cada partido se torna más importante, ya que cada rival es más directo. Esto aumenta la cantidad de partidos denominados *six-pointers* en donde es tan importante ganar como evitar que el otro equipo gane. Al haber más de este tipo de partidos aumenta el drama y por ende los espectadores.

La última variable de diseño a analizar es el descenso. Como se observa en la columna 1 del cuadro 1, tiene efecto positivo sobre la asistencia, pero no estadísticamente significativo. Esto se interpreta como que el aumento en asistencia provocado por la mayor inversión en calidad que propicia el descenso es compensado por la menor asistencia causada por la rotación ineficiente de equipos. En consecuencia, el efecto neto sobre la asistencia no es estadísticamente distinto de cero para todo el rango de valores que tomó la variable descenso en la muestra: entre 0 y 4.

El efecto no significativo del descenso sobre la asistencia no es nuevo en la literatura. De hecho, Burkitt y Cameron (1992) también reportan un coeficiente no significativo para la liga de rugby inglesa.

### 3.4 Resultados de las variables de control

El coeficiente de la variable dicotómica D7374 es negativo y estadísticamente significativo, como se esperaba. Esto se puede interpretar como que la inestabilidad social en Chile en el período 1973-74 provocó una fuerte caída en asistencia.

<sup>11</sup> Se escogió un torneo con 14 equipos y 3 ruedas (los equipos juegan todos contra todos 3 veces) porque es uno de los pocos diseños de torneos que no se desvía demasiado de las medias muestrales, proveyendo predicciones más confiables.

La *dummy* D3P no es estadísticamente significativa, implicando que la medida de la FIFA de dar tres puntos al equipo ganador en vez de dos no alteró el atractivo de los partidos. Este resultado es consistente con el modelo teórico desarrollado por Brocas y Carrillo (2004), quienes concluyeron que el efecto de dar tres puntos al equipo ganador tiene implicancias ambiguas sobre las estrategias ofensivas de los equipos, y en consecuencia, también tiene implicancias ambiguas sobre el atractivo de los partidos.

#### 4 CONCLUSIÓN

Este trabajo considera la pregunta ¿Puede el diseño del torneo de fútbol de Chile afectar su asistencia? Para responder esta pregunta se estima la demanda anual por asistencias al estadio entre 1935 y 2005 en función de variables económicas, de diseño de torneo y de control.

Para solucionar el problema de simultaneidad entre los precios de los tickets y la asistencia a los estadios, este trabajo estima la demanda en forma reducida de forma tal que los parámetros obtenidos son consistentes e iguales a los estructurales. Esto es posible gracias a que la forma funcional que mejor se ajusta a los datos es una *log-log*. Además, se encuentra que las variables no estacionarias cointegran y que las variables de diseño no provocan problemas de endogeneidad.

Los resultados encontrados en las variables económicas sugieren que la asistencia a los estadios es un bien normal, y que el influjo de sustitutos producido por la apertura económica de mediados de los años 1970 produjo una caída en la asistencia anual mayor a 300.000 espectadores.

Los resultados de las variables de diseño de torneo sobre la asistencia son decisivos. Una reducción de 1% en el número de equipos aumentaría la asistencia anual en 0,33% y un aumento de 1% en el número de partidos la incrementaría en 1,76%.

Sin embargo, estas elasticidades diseño-asistencia no implican que el diseño óptimo del torneo deba ser con 2 equipos e infinitos partidos por temporada. Esto último conllevaría a un torneo extremadamente monótono y que terminaría saturando el mercado por su excesiva oferta. Dada la muestra que se tiene no se puede decir cuál es el diseño óptimo del torneo. Sin embargo, sí se puede concluir en qué dirección se encuentra el óptimo: hacia menos equipos y más partidos, lo que implicaría mayor número de ruedas.

Por ejemplo, los resultados sugieren que un torneo con 14 equipos y 273 partidos (3 ruedas) llevaría un 28% más de espectadores al estadio anualmente que un torneo con 16 equipos y 240 partidos (2 ruedas).

En consecuencia, se concluye que el diseño del torneo de fútbol de Chile puede afectar la asistencia. De hecho, la reducción en el número de equipos acompañada de mayor cantidad de ruedas para generar más partidos no es nueva en el fútbol. En Escocia, en 1975, se cambió el diseño de la división de honor de 18 equipos y 2 ruedas a 10 equipos y 4 ruedas revirtiendo la tendencia negativa en asistencia que se venía manifestando en la última década.

La metodología descrita recientemente puede ser aplicada para evaluar una gran variedad de torneos deportivos, proveyendo un criterio más objetivo acerca de si éstos necesitan o no una modificación. También puede servir para evaluar industrias en donde algunas condiciones se imponen exógenamente.

#### REFERENCIAS

- Bird, P. (1982), "The Demand for League Football". *Applied Economics* 14: 637-649.
- Bresnahan, T. (1982), "The Oligopoly Solution Concept is Identified". *Economics Letters* 10: 87-92.
- Brocas, I. y J. Carrillo (2004), "Do the "Three-Point Victory" and the "Golden Goal" Rules Make Soccer More Exciting?" *Journal of Sports Economics* 5(2): 169-185.
- Burkitt, B. y S. Cameron (1992), "Impact of League Restructuring on Team Sport Attendances: the case of rugby league". *Applied Economics* 24: 265-271.
- Cairns, J. (1987), "Evaluating changes in league structure: the reorganization of the Scottish Football League". *Applied Economics* 19: 259-275.
- Díaz, J., R. Luders y G. Wagner (2004), "Chile 1810-2000: La República en Cifras". Libro en preparación para la publicación.
- Dobson, S., J. Goddard y J. Wilson (2001), "League Structure and Match Attendances in English Rugby League". *International Review of Applied Economics* 15(3): 335-351.
- Fundación Paz Ciudadana (2005), Anuarios de Estadísticas Criminales 1999-2001. [www.pazciudadana.cl/estadisticas.php](http://www.pazciudadana.cl/estadisticas.php)
- Instituto Libertad y Desarrollo (2001), "El Bienestar de los Pobres y la Importancia de Crecer al 7%". Temas públicos N° 51.
- Instituto Nacional de Estadísticas (INE), Censos de Chile, para los años 1930, 1940, 1952, 1960, 1970, 1982, 1992, 2002.
- Marín, E. (1995), "Historia Total del Fútbol Chileno: 1895-1995". Editores y Consultores REI limitada.
- Mackinnon, J. (1991), "Critical Values for Cointegration Tests". In *Long-run relationships* edited by R.F. Engle and C.W.J. Granger, Oxford University Press.
- MacKinnon, J., H. White y R. Davidson (1983), "Tests for Model Specification in the Presence of Alternative Hypotheses; some further results". *Journal of Econometrics* 21: 53-70.
- Noll, R. (2002), "The economics of promotion and relegation in sports leagues: the case of English Football". *Journal of Sports Economics* 3(2): 169-203.
- Nti, K. (1997), "Comparative Statics of Contests and Rent-seeking Games". *International Economic Review* 38(1) 43-59.
- Palomino, F. y L. Rigotti (2000), "The Sports League Dilemma: Competitive Balance versus Incentives to Win", paper E00'292. Institute of Business and Economic Research.
- Pyndick, R. y D. Rubinfeld (2000), *Econometric Models and Forecasts*. Forth edition, McGraw-Hill.
- Szymanski, S. y T. Valletti (2003), "Promotion and Relegation in Sporting Contests", working paper, Imperial College London. [www.nhh.no/sam/stabssem/2003/szymanski.pdf](http://www.nhh.no/sam/stabssem/2003/szymanski.pdf).

## A. APÉNDICE A: EL MODELO

### A. 1 Introducción al modelo

Este modelo se enmarca dentro de la literatura de torneos deportivos. En particular, se inserta dentro de los modelos que intentan explicar cómo cambia la calidad que los equipos eligen para enfrentar un torneo ante cambios en el diseño de dicho torneo, entendiendo la calidad como el gasto en planilla que realizan los equipos. Los modelos que están más en la línea del presente son los de Noll (2002) y Szymanski y Valletti (2003). El modelo de Noll examina cuál es el efecto del descenso sobre la elección de calidad de los equipos en un torneo deportivo que inicialmente estaba cerrado (no existía la posibilidad de descender). Él asume que los clubes maximizan utilidades esperadas. Sin embargo, no explicita ninguna forma funcional para la función generadora de ingresos de los clubes ni para las probabilidades de descenso. Sólo señala que los ingresos de los clubes son cóncavos con respecto a su único argumento, la calidad, y que la derivada de la función de probabilidad de descenso es negativa con respecto a este mismo argumento. La desventaja de este modelo es que al tener funciones no explícitas de ingresos ni de probabilidades, no permite avanzar mucho en los resultados que impliquen interacción estratégica entre los participantes, ya que la resolución analítica del modelo se ve limitada rápidamente. Por otra parte, el modelo de Szymanski y Valletti quiere examinar el efecto del descenso en un torneo deportivo sobre la dispersión en las elecciones de calidad de los equipos participantes. Para ello, al igual que Noll, asume que los clubes maximizan utilidades, pero ocupa funciones explícitas tanto para los ingresos como para la probabilidad de descender. La desventaja de este modelo es que ocupa una función de ingresos que asume que los clubes sólo generan ingresos si es que ganan el torneo.

El presente modelo toma elementos de los dos modelos mencionados anteriormente. En primer lugar, sigue la estrategia de Noll (2002) de tomar funciones de ingresos cóncavas con respecto a la calidad propia para poder modelar el hecho de que los clubes no sólo generan ingresos si son campeones, sino que generan ingresos por la calidad exhibida. Además, toma las funciones explícitas de probabilidad de descenso de Szymanski y Valletti (2003) (*logits standard*) para poder avanzar un poco más en los equilibrios que impliquen interacción estratégica. Por último, cabe destacar que este modelo incorpora la heterogeneidad entre los clubes, factor que no fue modelado en los trabajos anteriores.

Con el marco recién descrito, este modelo intenta replicar el modelo de Noll (2002) acerca de los efectos del descenso sobre las elecciones de calidad de los participantes. Sin embargo, lo más relevante es que plantea dos nuevas preguntas: ¿Qué sucede con la calidad del torneo si es que aumenta el número de equipos manteniendo el número de partidos constante? y ¿Qué sucede con la calidad del torneo si es que aumenta el número de partidos manteniendo el número de equipos constante? En resumen, el presente modelo examina cómo se afecta la elección de calidad de los clubes ante cambios en las variables de diseño descenso, equipos y partidos.



## A.2 Descripción del modelo

La interacción estratégica está modelada como un juego estático<sup>12</sup>. Existen dos divisiones: primera división  $A$  y primera división  $B$  ( $A$  y  $B$  en adelante). Cada una tiene  $N$  clubes que denotamos  $\{1, 2, \dots, N\}$ . El análisis se centra en  $A$ . Esta se modela completamente, mientras que  $B$  sólo se incluye para examinar los efectos de las variables de diseño sobre los incentivos generados en los equipos pertenecientes a  $A$ . La elección de calidad determina el pago del juego, compuesto de dos partes: los beneficios contemporáneos y un “premio de división” que representa el premio esperado<sup>13</sup>, medido como beneficios monetarios de comenzar el torneo siguiente en  $A$  o, alternativamente, el castigo de comenzar en  $B$ . Se adopta la noción de “premio de  $A$ ”, tal que para cada equipo  $i$  el premio de  $A$  es  $G_i^A = \bar{G}_i^A + \epsilon_i$ , con  $\epsilon_i \sim iid(0, \sigma^2)$ ,  $\bar{G}_i^A > 0$  y volatilidad suficientemente pequeña para esperar  $G_i^A > 0$ <sup>14</sup>.

La apertura o posibilidad de descenso se denota como  $r \in \{0, 1, \dots, N-1\}$  y es el número de clubes que, una vez observado el desempeño al final de cada torneo, desciende de  $A$  a  $B$  y sube de  $B$  a  $A$ . Si  $r = 0$ , decimos que el torneo es cerrado. Si  $r = 1$ , decimos que es abierto. Analizamos el formato de torneo “todos contra todos”, tal que cada club  $i \in \{1, 2, \dots, N\}$  juega una sola vez contra todo  $j \in \{1, 2, \dots, i-1, i+1, \dots, N\}$  de tal manera que el número de partidos que cada equipo juega,  $P$ , es  $N-1$ .

Cada club  $i \in \{1, 2, \dots, N\}$  escoge su calidad  $q_i \in [0, \bar{q}]$  justo antes de comenzar cada torneo. Pensaremos en  $q$  como el vector de calidades, tal que  $q = [q_1, q_2, \dots, q_N]$ . Se denomina  $Q$  a la calidad total del campeonato, tal que  $Q = f(q)$ , con  $\frac{df(q)}{dq_i} \geq 0 \forall i \in \{1, 2, \dots, N\}$ .

Cada equipo  $i$  genera ingresos de acuerdo a la función  $P^\lambda R^i(q_i)$ , con  $\lambda \in (0, 1)$ , tal que para todo club el ingreso marginal es decreciente en la cantidad de partidos que juega<sup>15</sup>. Aparte del número de partidos, se asume que sólo la calidad individual influye en la generación de ingresos contemporáneos, siguiendo a Noll (2002) y a Szymanski & Valletti (2003). Se supone que  $R^i(0) = 0$ ,  $R_q^i > 0$  y  $R_{qq}^i < 0$ <sup>16</sup> para dar cuenta de: i) no se paga la “no calidad” y ii) el ingreso es creciente a tasas decrecientes en la calidad.

<sup>12</sup> El juego estático es una simplificación del juego repetido (infinitas veces) que sería la representación más fiel de la interacción en torneos deportivos. A pesar de estar omitiendo en su conceptualización el “futuro estratégico”, permite concentrarse en las dimensiones contemporáneas de las elecciones conjuntas de calidad desde una perspectiva más didáctica.

<sup>13</sup> La naturaleza estocástica está dada por la probabilidad de descender, que se modela más adelante.

<sup>14</sup> En un juego repetido, este diferencial sería dependiente de las características de los torneos futuros y de las estrategias de elección de calidad de cada equipo en dichos torneos. En este juego estático el error captura dichos efectos y se asume un diferencial fijo y particular a cada club.

<sup>15</sup> A más de partidos, la disposición a pagar por ellos disminuye.

<sup>16</sup> Como usualmente,  $R_q^i$  y  $R_{qq}^i$  denotan, respectivamente, la primera y segunda derivada en el único argumento  $q_i$ .

El no incluir en la función generadora de ingresos de los equipos la calidad total del torneo,  $Q$  tiene cuatro justificaciones. La primera es que, al incluirla, no se pueden encontrar soluciones analíticas en dos de las tres demostraciones de este modelo si no se explicita algún tipo de función. La segunda, es que si se explicita algún tipo de función como por ejemplo  $R_i(Aq_1^{\alpha_1} q_2^{\alpha_2} \dots q_i^{\alpha_i} \dots q_N^{\alpha_N})$ . Donde  $A > 0$ ,  $0 < \alpha_i < 1 \forall i$ , y además,  $\alpha_i > \alpha_{-i}$ , tal que se suponga funciones de ingreso cóncavas que crezcan más si es que aumenta la calidad propia que la de un equipo rival, hace perder generalidad al modelo y debe ser resuelto por métodos numéricos en vez de analíticos. En tercer lugar, suponer una función como la recién expuesta implica solamente que aumenta el grado de complementariedad estratégica entre las elecciones de calidad de los equipos, lo que refuerza los resultados que se encontrarán más adelante, haciendo que el supuesto sea inocuo. En cuarto lugar, trabajar con una función generadora de ingresos por partido que sólo depende de la calidad propia permite expandir levemente lo realizado por los trabajos de economía deportiva, haciéndolo fácilmente comparable.

Ahora, nos disponemos a caracterizar la heterogeneidad de la función  $R^i(q_i)$ . Suponemos que es distinta para cada  $i$ , ya que los ingresos que se desprenden de una determinada calidad  $\hat{q} \in [q, \bar{q}]$ <sup>17</sup> pueden ser muy diferentes para equipos con diversas hinchadas. Por ejemplo, para un equipo muy popular o con una hinchada muy adinerada, la función de ingresos marginales sería  $R_q^p(\hat{q})$ , mientras que para un equipo poco popular o con hinchada poco adinerada sería  $R_q^u(\hat{q})$ , tales que  $R_q^p(\hat{q}) > R_q^u(\hat{q})$ , es decir, para el mismo monto de calidad escogida por ambos clubes, el club más popular y/o más adinerado obtiene mayores ingresos marginales. Esto sumado a que  $R^i(0) = 0$  y a que las elecciones de calidad están restringidas a ser positivas, implica que  $R^p(\hat{q}) > R^u(\hat{q})$ , ya que en cada punto, la función generadora de ingresos tiene mayor pendiente para el club más popular. Luego, los diversos clubes pueden ordenarse de acuerdo a su función de ingresos totales y marginales de la siguiente forma:  $R^1(\hat{q}) > R^2(\hat{q}) > \dots > R^N(\hat{q})$ ;  $R_q^1(\hat{q}) > R_q^2(\hat{q}) > \dots > R_q^N(\hat{q})$ , siendo el club 1 el más popular y el club N el menos popular.

El costo unitario de la calidad es constante e igual a  $w$ , de modo que todos los clubes tienen la misma tecnología de producción. Así, la función de costos totales de cada club es

$$(A1) \quad CT(q_i) = F + wq_i$$

donde  $F$  es un costo fijo y  $wq_i$  son los costos variables. El hecho de que el costo unitario de la calidad ( $w$ ) sea constante implica que la elasticidad de la oferta de calidad deportiva es infinita. El modelo es robusto si se plantea con costos convexos que implican elasticidad de oferta de calidad positiva.

Justo antes de encontrarse en el torneo, todo club  $i$  maximiza el valor de su pago considerando la estructura del torneo, es decir, maximiza sus utilidades. Alternativamente podemos considerar un escenario en donde los equipos están

<sup>17</sup>  $\underline{q} > 0$ .

maximizando el rendimiento deportivo o satisfacción de los hinchas menos los costos de ese determinado rendimiento o satisfacción, sin variar la estructura de la función objetivo.

Las utilidades de los clubes  $\pi^i$  son

$$(A2) \quad \pi^i(q_i) = P^\lambda R^i(q_i) - CT(q_i)$$

Por simplicidad se asume que el premio otorgado al equipo campeón es cero<sup>18</sup> y que  $\delta$  es el factor de descuento (igual para todos los clubes). En una liga abierta ( $r = 1$ ), los clubes que en  $t = 0$  juegan en  $A$ , en  $t = 1$  pueden comenzar jugando en  $A$  o en  $B$ , dependiendo de si mantuvieron o perdieron la categoría.

La probabilidad de descender  $D^i(q)$  para un equipo  $i$  se modela como una logit Standard siguiendo a Szymanski y Valletti (2003) y a Nti (1997)<sup>19</sup> y es equivalente a la probabilidad de salir último en un torneo donde se enfrentan todos los equipos contra todos:

$$(A3) \quad D^i(q) = \frac{q_i^{-\gamma}}{\sum_{k \in \{1, \dots, N\}} q_k^{-\gamma}}$$

donde  $\gamma$  representa qué tan fuerte afectan las elecciones de calidad a la probabilidad de descender y es mayor o igual a cero. Si  $\gamma = 0$ , la probabilidad de descender es independiente de la elección de calidad, siendo siempre igual a  $1/N$ . Si  $\gamma$  tiende a infinito se puede demostrar que en el límite  $D^i(q) = 1$ .

Además,

$$(A4) \quad \frac{\partial D^i(q)}{\partial q_i} = - \frac{\gamma q_i^{-\gamma-1}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2} \left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\} < 0$$

Lo que implica que si aumenta la calidad escogida por el club  $i$  disminuye su probabilidad de descenso, *ceteris paribus*.

<sup>18</sup> Se asume premio 0 para simplificar, ya que los cambios en los incentivos vienen dados por los cambios en las probabilidades de descender a B. Si hubiese un premio  $X$  por ganar el campeonato y la probabilidad de obtenerlo fuera  $S^i(q)$ , la maximización se plantearía de la siguiente manera:  $\max_{q_i} P^\lambda R_i(q_i) + S^i(q) * X - F - wq_i$ .

Esto no cambiaría los resultados encontrados, sólo lograría que existiera interacción estratégica aun cuando el torneo fuese cerrado.

<sup>19</sup> Esta función es de fácil maniobrabilidad algebraica.

A.3 Desarrollo del modelo

A.3.1 Torneo cerrado o sin descenso

Cada equipo debe maximizar sus utilidades de acuerdo al marco de incentivos presentado por el diseño del campeonato, sin embargo las elecciones de calidad serán distintas para los diversos clubes, ya que ellos difieren en sus funciones generadoras de ingresos. La elección de calidad que realiza un equipo  $i$  en  $t = 0$  obedece a la maximización de las utilidades y al diferencial de premio esperado por la división donde comienza jugando en  $t = 1$ .

Para un torneo cerrado se puede asumir que el premio esperado es cero, ya que en  $t = 1$  se mantiene la división A cualquiera que sea el lugar que ocupe el equipo  $i$  en  $t = 0$  o bien que la probabilidad de ganar el premio es uno, lo que no cambiaría la CPO del problema. En este caso, la maximización de utilidades para un equipo  $i$  asume lo primero, y queda conformada de la siguiente manera:

$$(A5) \quad \max_{q_i} P^\lambda R^i(q_i) - F - wq_i$$

$$(A6) \quad \text{CPO} \quad R_q^i(q_i) = \frac{w}{p^\lambda}$$

Esto implica que mientras menor sea el costo de la calidad ( $w$ ) y mayor la cantidad de partidos se aumenta la elección de  $q_1$  por parte del club  $i$ .

Se denominará  $q_N^{C*}$  al vector de calidades óptimas escogidas por los  $N$  clubes en el diseño de campeonato cerrado de manera de satisfacer la ecuación (A6). Los clubes más populares escogerán necesariamente un nivel de calidad mayor que los menos populares.

A.3.2 Torneo abierto o con descenso ( $r = 1$ )

La maximización queda conformada de la siguiente manera para un equipo  $i$  cualquiera:

$$(A7) \quad \max_{q_i} P^\lambda R_i(q_i) - F - wq_i + \delta \left[ (1 - D^i(q)) E(G_i^A) \right]$$

$$(A8) \quad \text{CPO} \quad R_q^i(q_i) = \frac{w}{p^\lambda} - \frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2 P^\lambda}$$

Sea  $q_N^{O*}$  el vector de calidades óptimas escogidas por los  $N$  clubes en el diseño de campeonato abierto de manera de satisfacer la ecuación (A8). En equilibrio, también se da que los clubes más populares escogerán necesariamente un nivel de calidad mayor que los menos populares para el mismo formato de torneo.

### A.3.3 Comparación de elección de calidad en torneos abiertos versus cerrados

En esta sección se comparan las calidades escogidas por  $N$  equipos en un torneo cerrado versus uno abierto, de manera de llegar a conclusiones acerca del cambio en la calidad total del torneo provocada por la existencia del descenso.

**Proposición 1.** En una liga cerrada la calidad escogida por todos los equipos de  $A$  es menor que en una liga abierta ( $q_N^{C*} < q_N^{O*}$ ), por lo tanto la apertura aumenta la calidad del campeonato.

**Prueba.** Como se observa en la ecuación (A6), la calidad escogida por cualquier equipo  $A$  es tal que  $R_q^i(q_i) = \frac{w}{p^\lambda}$ , sin embargo, se ve que en la ecuación (A8)

$$(A9) \quad R_q^i(q_i) = \frac{w}{p^\lambda} - \frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2 p^\lambda} \quad \text{y dado que } \delta, \bar{G}_i^A, \gamma, q \geq 0 \Rightarrow$$

$$\frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2 p^\lambda} \geq 0. \text{ Por lo tanto } R_q^i(q_i) \text{ debe ser menor en una liga}$$

abierta que en una cerrada. Como la función  $R_q^i(q_i)$  es cóncava, la calidad requerida para satisfacer la ecuación (A7) debe ser menor que la requerida para satisfacer la ecuación (A8) para el equipo  $i$  y como esto se cumple  $\forall i \in \{1, 2, \dots, N\}$ , aumenta la calidad escogida por todos los equipos participantes del torneo, independientemente de las acciones tomadas por el resto de los clubes, es decir, si se abre la liga la estrategia de aumentar la calidad escogida por cualquier equipo es dominante. Entonces, si se permite descender en la división  $A$ ,  $\Delta^+ q_i \forall i \in \{1, 2, \dots, N\} \Rightarrow \Delta^+ q$  y por ende  $\Delta^+ Q$  ceteris paribus.

### A.3.4 Comparación de elección de calidad variando el número de equipos y partidos

En esta sección se asume el escenario de una liga abierta.

Al aumentar la cantidad de clubes participantes de  $N$  a  $N + 1$  se esperan dos efectos sobre la calidad total del torneo. El primero es el “efecto equipos” y el segundo es el “efecto partidos”, ya que se definió el formato del torneo como uno en donde juegan todos los equipos contra todos, lo que implica que al haber más equipos también hay más partidos. A su vez, el “efecto equipos” tiene dos aristas: Las implicancias sobre las elecciones de calidad para los  $N$  equipos pertenecientes al campeonato antes de la expansión y el efecto sobre la calidad propiciado por el equipo entrante.

La intuición detrás del “efecto equipos” es la siguiente: Al incorporar un equipo extra al torneo, los  $N$  clubes iniciales ven más lejana su probabilidad de descender y sus incentivos a escoger calidad disminuyen, lo que conlleva a una merma en la calidad del campeonato. Por otra parte, el efecto que produce la entrada de un nuevo participante es que aporta algún monto positivo de calidad al torneo. Sin embargo, se sabe que el aumento en el número de participantes en una liga abierta se produce porque se acrecenta la cantidad de clubes que provienen de la división inmediatamente inferior por sobre el número normal de equipos que ascienden, o porque se permite permanecer en la categoría a equipos que con la cantidad de descenso habitual la hubiesen perdido. Esto implica que el aumento en la cantidad de participantes viene propiciado por equipos que con una alta probabilidad tienen una función generadora de ingresos menor a la de los equipos pertenecientes a la división  $A$  y por lo tanto escogen montos de calidad inferiores a los que eligen los  $N$  equipos que participaban con anterioridad. Entonces, la calidad total del torneo aumenta sólo de manera modesta por los aportes del equipo entrante. El hecho de que los entrantes sean más débiles<sup>20</sup> que los incumbentes se puede modelar como  $q_{N+1} \leq q_i \forall i \in \{1, 2, \dots, N\}$ .

Por otra parte, la intuición detrás del “efecto partidos” es que conllevaría a escoger mayor calidad a los  $N$  equipos iniciales, debido a que al aumentar el número de partidos se incrementa la rentabilidad de la calidad escogida por los clubes, ya que la función generadora de ingresos depende positivamente de  $P$ .

De los efectos mencionados recientemente, el aumento en la calidad del torneo propiciado por el aporte del equipo entrante se cumple de manera tautológica, pero los efectos sobre las elecciones de calidad de los  $N$  equipos iniciales derivados del aumento en el número de partidos y de equipos hay que analizarlos en detalle. Esto se realiza a continuación.

**Proposición 2.** *Si se mantiene la cantidad de partidos jugados constante, el aumento en el número de equipos participantes disminuye las elecciones de calidad de los  $N$  clubes iniciales y si se mantiene la cantidad de equipos constante, el aumento en la cantidad de partidos jugados incrementa las elecciones de calidad de esos mismos clubes<sup>21</sup>.*

**Prueba.** La condición de primer orden de la maximización de utilidades para un equipo  $i$  en un torneo abierto con  $N + 1$  equipos participantes es:

$$(A10) \quad (P + 1)^\lambda R_q^i(q_i) = w - \frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right]^2}$$

<sup>20</sup> Con funciones generadoras de ingresos menores.

<sup>21</sup> Las implicancias de esta proposición cobran sentido práctico cuando se relaja el supuesto de que los equipos juegan todos contra todos una sola vez y se permiten más rondas para ajustar el número de equipos y el número de partidos.

El vector de calidades que satisface esta ecuación para cualquier equipo  $i$  es  $q_{N+1}^{O*}$ , es decir, ese es el vector resultante de la intersección de las funciones de reacción de los equipos. Análogamente, la condición de primer orden de la maximización de utilidades para un equipo  $i$  en un torneo abierto con  $N$  equipos participantes es:

$$(A11) \quad P^\lambda R_q^i(q_i) = w - \frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2}$$

Para el caso de  $N$  equipos, el vector de calidades que satisface esta ecuación para cualquier equipo  $i$  es  $q_N^{O*}$ . De las ecuaciones (A10) y (A11) se observa que el  $q_i$  resultante puede no ser el mismo, ya que este valor depende del equilibrio completo del sistema de ecuaciones, los cuales pueden ser distintos cuando hay  $N + 1$  y  $N$  participantes. Para recalcar este hecho se ocupa un supraíndice en las funciones de ingresos marginales que denote el proceso maximizador del cual se obtiene el  $q_i$  correspondiente. A la función  $R_q^i(q_i)$  de la ecuación (A10) se le denomina  $R_q^{i,N+1}(q_i)$ , mientras que si proviene de la ecuación (A11) queda como  $R_q^{i,N}(q_i)$ , de ahora en adelante. Si se restan las ecuaciones (A10) y (A11) se obtiene:

$$(A12) \quad \begin{aligned} & \left( R_q^{i,N+1}(q_i) - R_q^{i,N}(q_i) \right) = \\ & \frac{w}{(P+1)^\lambda} - \frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right]^2 (P+1)^\lambda} \\ & - \frac{w}{(P)^\lambda} + \frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2 (P)^\lambda} \end{aligned}$$

Por simplicidad, al lado derecho de la ecuación (A12) se le denomina  $\Psi$ , es decir, la ecuación (A12) puede ser rescrita como:

$$(A13) \quad \left( R_q^{i,N+1}(q_i) - R_q^{i,N}(q_i) \right) = \Psi$$

Entonces, si al reemplazar el vector de calidades que satisface a los  $N$  equipos iniciales cuando sólo existen  $N$  participantes ( $q_N^{O*}$ )<sup>22</sup> en todos los términos de la ecuación (A12) se cumple que  $\Psi > 0$ , implica que la calidad escogida por un equipo  $i \in \{1, 2, \dots, N\}$  disminuye al aumentar el número de clubes de  $N$  a  $N + 1$  conjuntamente con aumentar el número de partidos jugados de  $P$  a  $P + 1$ . La lógica es que como la función generadora de ingresos sólo depende de la calidad propia, el lado izquierdo de la ecuación (A13) sería cero al reemplazar en ambas funciones el mismo  $q_i \in q_N^{O*}$ , mientras que el lado derecho sería mayor que cero, si se cumple que  $\Psi > 0$ . Esto implicaría que la calidad de equilibrio para el caso de  $N + 1$  equipos y  $P + 1$  partidos debe ser menor que la de equilibrio con  $N$  equipos y  $P$  partidos para equilibrar ambos lados de la ecuación. En particular, un  $q_i$  menor en el caso de haber  $N + 1$  equipos y  $P + 1$  hará aumentar el lado izquierdo, ya que  $R_q^{i,N+1}(q_i)$  aumentará, y disminuir el lado derecho, ya

$$\text{que } \frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right]^2 (P+1)^\lambda}$$

también aumentará<sup>23</sup>. Lo dicho recientemente-

se sostiene bajo el supuesto de que las calidades de los otros  $N - 1$  equipos permanecen constantes en las correspondientes al vector  $q_N^{O*}$ . Pero se sabe que los otros clubes no mantienen su calidad constante, sino que también la disminuyen, ya que enfrentan el mismo problema que  $i$ . Considerando dicha interacción, el efecto de la disminución de calidad se refuerza, puesto que éstas son complementos estratégicos<sup>24</sup> y por lo tanto se mengua la calidad escogida por los  $N$  participantes iniciales. Análogamente, si  $\Psi = 0$  las calidades no cambian y si  $\Psi < 0$  éstas aumentan. Por ende, hay que examinar el signo de  $\Psi$  una vez que se haya reemplazado  $q_N^{O*}$  en la ecuación (A12), para saber qué ocurre con la calidad de los clubes si aumenta el número de equipos de  $N$  a  $N + 1$ . Para facilitar la exposición, en el numerador se ocupa la descomposición  $(P + 1)^\lambda = P^\lambda + \Xi$ , donde  $0 < \Xi < 1$ .

<sup>22</sup> Se asume un valor  $q_{N+1}$  cualquiera, tal que  $q_{N+1} > q_i$  para  $i \in \{1, 2, \dots, N\}$

<sup>23</sup> Es sencillo notar esto ocupando límites.

<sup>24</sup> Al asumir que se produce una disminución en la calidad escogida por club  $i$  cuando hay  $N + 1$  equipos participantes en vez de  $N$ , se está pensando en que la función de reacción del equipo  $i$  (con pendiente positiva, ya que las calidades serían complementos estratégicos en el rango relevante) se desplaza y necesita una calidad menor para llegar al equilibrio si es que las otras  $N - 1$  funciones de reacción permanecen constantes. Pero a las otras funciones de reacción les sucede lo mismo y por lo tanto el equilibrio final de las  $N$  funciones es escoger menor calidad para los  $N$  equipos. Es decir, se refuerzan las predicciones una vez que se considera la interacción estratégica.

De todas formas, como la formulación algebraica es amplia, puede haber la posibilidad de que en algún rango las calidades sean sustitutos estratégicos. Sin embargo, sólo en un caso muy extremo de substitutividad estratégica la calidad promedio por equipo (dado un número de partidos fijos) del campeonato disminuye si se considera dicha interacción, por lo que el supuesto de complementariedad estratégica es prácticamente inocuo. Además, la complementariedad estratégica se cumple en una serie de funciones explícitas de acuerdo al marco descrito.



$$(A14) \quad \Psi = \left( \frac{wP^\lambda - w(P+1)^\lambda}{(P+1)^\lambda P^\lambda} \right) -$$

$$\left( \frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1}}{(P+1)^\lambda P^\lambda} \right) * \left( \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\} P^\lambda}{\left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right]^2} - \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\} P^\lambda}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2} \right)$$

La idea es reagrupar términos para poder separar el “efecto equipos” del “efecto partidos”. Después de un poco de álgebra  $\Psi$  se puede expresar como:

$$(A15) \quad \left( \frac{-\Xi}{(P+1)^\lambda P^\lambda} \right) * \left( w - \frac{\delta \bar{G}_i^A \gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2} \right) \text{E. partidos}$$

$$+ \left( \frac{-\delta \bar{G}_i^A \gamma q_i^{-\gamma-1}}{(P+1)^\lambda} \right)$$

$$* \left( \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right]^2} - \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2} \right) \text{E. equipos}$$

En la ecuación (A15) se observan dos efectos contrapuestos. El primer término de la ecuación (A15) representa el “efecto partidos” y es exactamente el diferencial de las funciones de ingresos si es que hubiera aumentado el número de partidos de  $P$  a  $P + 1$ , sin que cambiara el número de equipos participantes. El segundo término es el “efecto equipos” y representa un aumento en el número de equipos participantes desde  $N$  a  $N + 1$  considerando que se juegan  $P + 1$  partidos en el campeonato. El “efecto partidos” tiene signo negativo, ya que  $\Xi$ ,  $P$  y

$$\left( w - \frac{\delta \bar{G}_i^\gamma q_i^{-\gamma-1} \left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2} \right) \text{son positivos}^{25}. \text{ Esto implica que jugar más}$$

partidos afecta negativamente a  $\Psi$  y por lo tanto las calidades de los  $N$  equipos iniciales aumentan por este factor. El efecto equipos tiene signo positivo, ya que

$$\text{el término } \left( \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right]^2} - \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2} \right) \text{ es negativo (como se}$$

verá en la demostración de la proposición 3) y está multiplicando a otro término negativo. Por lo tanto, el efecto neto de aumentar el número de equipos es incrementar  $\Psi$  y en consecuencia las calidades de los  $N$  equipos que inicialmente componían la liga disminuyen. ■

Finalmente, el modelo se cierra con la proposición y demostración 3 que se había dejado pendiente.

**Proposición 3.** Si la probabilidad de descender del equipo  $N$  es menor a 0,618, en un torneo con  $N$  clubes participantes, es condición suficiente para que se cumpla:

$$(A16) \left( \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right]^2} - \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2} \right) < 0 \forall i \in \{1, 2, \dots, N\}$$

**Prueba.** Manipulando algebraicamente la inecuación (A16) se llega a:

$$(A17) \left( \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] \left\{ 2q_i^{-\gamma} - \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right\} - q_{N+1}^{-\gamma} q_i^{-\gamma} \right) < 0$$

<sup>25</sup> El término entre paréntesis es positivo porque es equivalente al lado derecho de la ecuación (A11), el cual debe ser mayor que cero para que se escojan niveles de calidad positivos.

Si

$$\left\{ 2q_i^{-\gamma} - \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right\} < 0$$

es condición suficiente para que la inecuación (A17) se verifique.

Si

$$\left\{ 2q_i^{-\gamma} - \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right\} > 0$$

hay que comparar la multiplicación de los términos entre paréntesis que son positivos con el negativo  $(-q_{N+1}^{-\gamma} q_i^{-\gamma})$  para ver si se cumple la inecuación (A17).

Es sencillo notar que

$$\left\{ 2q_i^{-\gamma} - \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right\} < 0 \forall i \neq N$$

dado que  $q_N^{-\gamma} > q_{N-1}^{-\gamma} \dots > q_1^{-\gamma 27}$ . Entonces, la desigualdad expuesta en la inecuación (A17) se cumple para los  $N - 1$  equipos con mayores funciones generadoras de ingresos.

Sin embargo, aún queda por analizar qué sucede cuando el equipo  $i$  es el con menor función generadora de ingresos y por ende el club con menor elección de calidad, es decir, cuando  $i = N$ . Para examinar este caso particular se desglosa la sumatoria anterior y posteriormente se reemplaza  $q_i$  por  $q_N$  lo que hace que el término pueda ser expresado como:

$$(A18) \left\{ (q_N^{-\gamma} - q_{N-1}^{-\gamma}) - \sum_{k \in \{1, \dots, N-2\}} q_k^{-\gamma} \right\}$$

En este caso,  $(q_N^{-\gamma} - q_{N-1}^{-\gamma}) > 0$  y de hecho si  $q_N$  es muy pequeño,  $q_N^{-\gamma}$  es muy grande, lo que puede llevar incluso a que

$$(A19) \left\{ (2q_i^{-\gamma} - q_{N-1}^{-\gamma} - q_N^{-\gamma}) - \sum_{k \in \{1, \dots, N-2\}} q_k^{-\gamma} \right\} > 0$$

<sup>26</sup> Esto se debe a que los equipos con mayores funciones de ingresos escogen montos de calidad mayores, porque tienen funciones de ingreso marginal ordenables.

Por lo tanto, hay que analizar la inecuación (A17) completa reemplazando  $q_i$  por  $q_N$  para ver bajo qué circunstancias se cumple. Es decir, hay que examinar bajo qué condiciones se verifica que

$$(A20) \left( \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] \left\{ 2q_N^{-\gamma} - \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right\} - q_{N+1}^{-\gamma} q_N^{-\gamma} \right) < 0$$

Para poder encontrar una condición coherente con el modelo, se trata de expresar la inecuación (A20) en términos de la probabilidad de descenso del equipo  $N$  cuando participan  $N$  clubes en el torneo, es decir, en términos de:

$$D^N(q) = \frac{q_N^{-\gamma}}{\sum_{k \in \{1, \dots, N\}} q_k^{-\gamma}}$$

Para eso se divide la ecuación (A20) por

$$\left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]$$

, debiéndose cumplir que

$$(A21) \left\{ \frac{2q_N^{-\gamma}}{\sum_{k \in \{1, \dots, N\}} q_k^{-\gamma}} - 1 \right\} - \left( \frac{q_{N+1}^{-\gamma}}{\sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma}} \right) \left( \frac{q_N^{-\gamma}}{\sum_{k \in \{1, \dots, N\}} q_k^{-\gamma}} \right) < 0$$

Factorizando por

$$\frac{q_N^{-\gamma}}{\sum_{k \in \{1, \dots, N\}} q_k^{-\gamma}}$$

se llega a

$$(A22) \left[ \frac{q_N^{-\gamma}}{\sum_{k \in \{1, \dots, N\}} q_k^{-\gamma}} \right] \left[ 2 - \left( \frac{q_{N+1}^{-\gamma}}{\sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma}} \right) \right] - 1 < 0$$

Como se observa en la inecuación (A22), el caso en donde es más probable que ésta no se cumpla, es cuando el equipo entrante  $N + 1$  escoge la mayor

calidad posible y dado que  $q_{N+1} \leq q_N$ , ese monto es  $q_N$ . Por lo tanto, la condición que satisfaga la inecuación (A22) cuando se reemplace  $q_{N+1}$  por  $q_N$  es la más restrictiva de todas las posibles, es decir, es condición suficiente para que se cumpla la inecuación (A22). Reemplazando  $q_{N+1}$  por  $q_N$  se obtiene

$$(A23) \left[ \frac{q_N^{-\gamma}}{\sum_{k \in \{1, \dots, N\}} q_k^{-\gamma}} \right] \left[ 2 - \left( \frac{q_N^{-\gamma}}{\sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} + q_N^{-\gamma}} \right) \right] - 1 < 0$$

Manipulando términos, se llega a

$$(A24) \left[ \frac{q_N^{-\gamma}}{\left( \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right)} \right] \left[ \frac{(q_N^{-\gamma})^2}{\left( \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right)^2} \right] < 1$$

Finalmente, la inecuación (A24) está expresada en términos de

$$(A25) D^N(q) = \frac{q_N^{-\gamma}}{\sum_{k \in \{1, \dots, N\}} q_k^{-\gamma}}$$

Entonces, se reemplaza por dicha probabilidad.

$$(A26) [D^N(q)] [D^N(q)]^2 < 1$$

Por lo tanto, debe haber un valor crítico de  $D^N(q)$  que determine en el rango en el cual la inecuación (A26) se cumpla. Para encontrar dicho rango, se resuelve lo siguiente:

$$(A27) [D^N(q)] [D^N(q)]^2 = 1$$

El valor de  $D^N(q)$  que satisface la ecuación (A27) es 0,618, lo que implica que si  $D^N(q) < 0,618$  es condición suficiente para que se verifiquen las inecuaciones (A20) a (A27). Por ende queda demostrado que:

$$(A28) \left( \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N, N+1\}} q_k^{-\gamma} \right]^2} - \frac{\left\{ \left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right] - q_i^{-\gamma} \right\}}{\left[ \sum_{k \in \{1, \dots, N\}} q_k^{-\gamma} \right]^2} \right) < 0$$

$\forall i \in \{1, 2, \dots, N - 1\}$  y para  $i = N$  si  $D^N(q) < 0,618$ . ■

Suponer que la probabilidad de descender sea menor a 0,618 es razonable, ya que por lo general existen varios equipos débiles en los torneos de fútbol, los cuales no difieren mucho en sus funciones de ingresos y por lo tanto pensar en que la probabilidad de descenso de un equipo sea tan alta como para superar ese valor crítico es poco real, porque implicaría una disparidad tremenda entre la función generadora de ingresos del club más débil con el resto de los equipos pertenecientes al campeonato.

#### A.4 CONCLUSIONES DEL MODELO

En primer resultado del modelo recién expuesto concluye que la posibilidad de descenso aumenta la elección de calidad de los equipos participantes, ya que todos invierten extra en calidad para evitar la posibilidad de descender. Si bien hubiese sido deseable comparar dos escenarios en los cuales se diera interacción estratégica (torneo cerrado versus abierto) no se consideró así, puesto que hubiese complicado demasiado la comparación entre los escenarios con  $N$  y  $N + 1$  equipos. Esto puede ser tema para futuras investigaciones.

El segundo, y principal resultado del presente modelo es que si se mantiene constante la cantidad de partidos jugados, se espera que al aumentar el número de equipos participantes disminuya la elección de calidad para los  $N$  equipos iniciales, ya que sólo se altera su probabilidad de descender y no su función generadora de ingresos. A su vez, si se mantiene el número de equipos constante y se aumenta la cantidad de partidos jugados, se espera que aumente la calidad del torneo, ya que es posible rentabilizar más cada unidad de calidad adquirida. También se espera que al aumentar el número de equipos participantes, el entrante escoja niveles de calidad menores a las elegidas por los incumbentes y por lo tanto aporte pequeños montos de calidad al torneo. Si ese equipo extra no significa un aumento en el número de partidos implica que la calidad por partido jugado disminuye.

La gran implicancia de este modelo es que se logra examinar por separado el “efecto partidos” del “efecto equipos” logrando observar hacia dónde apunta cada uno. Esto es útil, ya que normalmente cuando se piensa en cambiar el diseño de un torneo se piensa en variar el número de equipos para que haya distinta cantidad de partidos o variar el número de equipos asumiendo las consecuencias de que al mismo tiempo cambie el número de partidos. Ahora, sin embargo,

queda más claro que los incentivos subyacentes de cada uno de estos factores apuntan en sentido contrario y que al aumentar el número de ruedas disminuyendo el número de equipos, se puede lograr que el número de partidos también aumente, aprovechando ambos efectos positivos sobre la inversión en calidad de los clubes.

## B. APÉNDICE B: DEFINICIÓN DE LAS VARIABLES Y SUS FUENTES

### **Variable Dependiente**

*Asistencia*: Asistencia anual al estadio en la primera división del fútbol chileno.

Fuente: Marín (2002).

### **VARIABLES INDEPENDIENTES.**

#### *i) Variables económicas.*

*Ingreso*: Representa el producto interno bruto promedio anual por habitante en Chile. Está expresando en millones de pesos de 1996.

Fuente: Díaz *et al.* (2004).

*Población*: Población total de Chile.

Fuente: Díaz *et al.* (2004).

*Población urbana*: Población total urbana de Chile.

Fuente: Autor en base a información en los censos (1930, 1940, 1952, 1960, 1970, 1982, 1992 y 2002). Como no hay información para todos los años de la población urbana de Chile se completaron los datos faltantes con medias geométricas.

*D75*: Variable dicotómica que toma valor 0 hasta 1974 y 1 de 1975 en adelante.

*DTV*: Variable dicotómica que toma valor 1 cuando el torneo de fútbol fue televisado y 0 en caso contrario.

#### *ii) Variables de diseño.*

*Equipos*: Número de equipos que participaron en la primera división del fútbol chileno en cada año.

Fuente: Marín (2002).

*Partidos*: Número de partidos que se jugaron en la primera división del fútbol chileno en cada año.

Fuente: Marín (2002).

*Descenso*: Número de equipos que descendieron directamente por bajo desempeño en la primera división del fútbol chileno en cada año. Se excluyen los descensos por liguillas promocionales.

Fuente: Marín (2002).

#### *iii) Variables de control.*

*D7374*: Variable dicotómica que toma valor 1 en los años 1973 y 1974 y 0 en otro caso.

D3P: Variable dicotómica que toma valor 1 en los años que se le asignan 3 puntos al equipo ganador de un partido y toma valor 0 en caso contrario.

Violencia: Desde 1988 en adelante es el número de denuncias por daños a la propiedad privada en Chile. De 1987 hacia atrás se deflacta el indicador de 1988 por el crecimiento de la población<sup>28</sup>.

Fuente: Autor a partir de datos de Fundación Paz Ciudadana y Díaz *et al.* (2004).

<sup>27</sup> Este indicador se construyó de esta manera porque no existen *proxies* del fenómeno de la violencia en los estadios anteriores a 1988. Como el fenómeno de la violencia en los estadios y los desmanes en la ciudad de Santiago floreció en la década de los noventa, se considera que el período anterior a ese, la violencia creció a la par con el crecimiento vegetativo de Chile.



CONVERGENCIA Y ESTABILIDAD DE LOS TIPOS DE  
CAMBIO EUROPEOS: UNA APLICACIÓN  
DE EXPONENTES DE LYAPUNOV\*

ELENA OLMEDO

*Universidad de Sevilla*

RICARDO GIMENO

*Universidad Pontificia Comillas*

LORENZO ESCOT

*Universidad Complutense de Madrid*

RUTH MATEOS

*Universidad San Pablo-CEU*

*In this paper we applied the dynamic system theory to the measurement of the stability of the European process of convergence. In particular, Lyapunov's exponents are used to verify the influence of political and economic decisions made during the creation of the European Union on the stability (or instability) of exchange rate fluctuation in different European countries. We find evidence of such relationship.*

*JEL:* C45, C61, F02

*Keywords:* estabilidad, sistemas dinámicos, convergencia europea, euro, exponente de Lyapunov, redes neuronales.

1. INTRODUCCIÓN: SISTEMAS DINÁMICOS Y ESTABILIDAD DE LOS  
TIPOS DE CAMBIO

Es sobradamente conocido que la volatilidad de los tipos de cambio está caracterizada por una considerable persistencia, de manera que grandes movimientos en los mismos están seguidas de a su vez otros grandes movimientos, existiendo correlación serial positiva en sus cuadrados. Por este motivo, la volatilidad presente y pasada puede utilizarse para predecir la volatilidad futura. En definitiva, se puede considerar que el valor del tipo de cambio en el momento

\* E mail: [olmedo@us.es](mailto:olmedo@us.es), [rgimeno@upco.es](mailto:rgimeno@upco.es), [escot@ccee.ucm.es](mailto:escot@ccee.ucm.es), [matcab@ceu.es](mailto:matcab@ceu.es)

presente es una función del valor que presentaba en el pasado, definiéndose así un sistema dinámico.

Una vez enfocado el problema desde este punto de vista, parece lógico pensar que la estabilidad del sistema explicará, en gran medida, la volatilidad observada en el tipo de cambio. Si el sistema dinámico subyacente es inestable el tipo de cambio será también inestable. De esta manera, en el presente trabajo se utiliza una medida de la estabilidad de los sistemas dinámicos para analizar la estabilidad en los tipos de cambio.

¿Cómo se miden las propiedades de estabilidad del sistema dinámico? Si el sistema dinámico es lineal, se trabaja con sus autovalores, distinguiéndose en este caso el comportamiento a largo plazo estable (si el autovalor es de módulo menor que la unidad) o explosivo (si el autovalor es de módulo mayor que la unidad). El problema al que nos enfrentamos, sin embargo, es más complejo debido a que este tipo de sistemas lineales no puede generar un comportamiento volátil en los tipos de cambio, que sea tan complejo como el observado en la realidad. Para ello es necesario introducir no linealidades en el sistema.

Si el sistema dinámico es no lineal, sus propiedades de estabilidad se describen mediante los exponentes de Lyapunov (Lyapunov, 1892): si el máximo exponente de Lyapunov es positivo, el sistema es inestable y si es negativo, el sistema es estable. El sistema será tanto más estable conforme mayor sea, en valor absoluto, el exponente de Lyapunov. Un sistema dinámico se dice que es caótico si su máximo exponente de Lyapunov es positivo. Los sistemas caóticos se caracterizan por las diferentes dinámicas que presentan condiciones iniciales infinitesimalmente próximas, de manera que un sistema dinámico determinista pero caótico aparentemente funciona igual que uno aleatorio, sin serlo. Por este motivo el caos ha supuesto una nueva vía de estudio en la dinámica económica de sistemas, ya que no es necesaria la introducción de perturbaciones aleatorias para conseguir un comportamiento tan irregular y complejo como el observado en la realidad.

Por los motivos enunciados actualmente existen numerosos ejemplos empíricos de cálculo de exponentes de Lyapunov en series temporales procedentes del mundo económico o financiero. Como ejemplos pueden citarse trabajos para datos de tipos de cambio: (Dechert y Gencay, 1992; Bask y de Luna, 2002; Bask, 2002; Weston y Premachandran, 2004), para agregados monetarios: (Serletis, 1995; Barnett *et al.*, 1995), para rendimientos financieros: (Abhyankar *et al.*, 1997; y Shintani y Linton, 2004), para datos bursátiles: (Wang, Chen y Lü, 2004) y para datos reales de producción: (Shintani y Linton, 2003).

En el presente trabajo se plantea analizar la incidencia de decisiones políticas y/o económicas en el ámbito del proceso de Integración Monetaria Europea sobre la estabilidad de los tipos de cambio en el período 1997-2002. De esta manera, en la primera sección se introduce la definición de exponente de Lyapunov y cómo se calcula para series temporales, presentándose los resultados empíricos en la segunda sección.

## 2. CARACTERIZACIÓN DE LA ESTABILIDAD DEL SISTEMA MEDIANTE EL MÁXIMO EXPONENTE DE LYAPUNOV

### 2.1. Introducción

Consideramos primero un sistema dinámico unidimensional, en el que trabajamos con dos puntos  $x$  e  $y$  del espacio de estados  $X \in \mathbb{R}$ , tan próximos como deseemos, siendo  $d_0$  la distancia entre ellos. Aplicamos a estos puntos la acción de un flujo  $\phi_t$  (o una función  $f(x)$  en el caso discreto) y, después de un tiempo  $t$  sea  $d_t$  la distancia entre ellos. La idea base de los exponentes de Lyapunov parte de aproximar mediante una exponencial  $e^{\lambda t}$  el cociente de estas distancias, para medir lo que se llama el crecimiento exponencial

$$(1) \quad \frac{d_t}{d_0} \approx e^{\lambda t}$$

Si para  $t \rightarrow \infty$  se puede ver que  $\lambda_t$  converge a un límite  $\lambda$ , diremos que  $\lambda$  es el exponente de Lyapunov. Además, si este límite es mayor (estrictamente) que cero, podemos decir que bajo la acción del flujo (o de la función) órbitas inicialmente próximas divergen de forma exponencial conforme transcurre el tiempo o, lo que es lo mismo, que el sistema es inestable.

Consideramos el caso discreto. Partimos de un sistema dinámico

$$(2) \quad x_{k+1} = f(x_k)$$

siendo  $x \in X \subset \mathbb{R}$ ,  $f: X \rightarrow \mathbb{R}$ . Sea una condición inicial  $x_0$  y su correspondiente órbita  $\{x_0, x_1, x_2, \dots\}$ . Considerando un desplazamiento infinitesimal con respecto a  $x_0$  dado por el vector  $y_0$ . Entonces se verifica que

$$(3) \quad f(x_0 + y_0) = f(x_0) + f'(x_0)y_0$$

por lo que, siendo  $y_1 = f(x_0 + y_0) - f(x_0)$ , se cumple que

$$(4) \quad y_1 = f'(x_0)y_0$$

Siguiendo el mismo planteamiento, la evolución dinámica del vector  $y_n$  (ver Figura 1) vendrá dada por

$$(5) \quad y_{n+1} = f'(x_n) \cdot y_n$$

con lo cual tenemos la evolución de la distancia entre la órbita original  $\{x_0, f(x_0), f^2(x_0), \dots\}$  y la nueva órbita generada a partir de una condición arbitrariamente próxima  $\{x_0 + y_0, f(x_0 + y_0), f^2(x_0 + y_0), \dots\}$ .

El factor de crecimiento (o decrecimiento) de este desplazamiento infinitesimal viene dado por el cociente

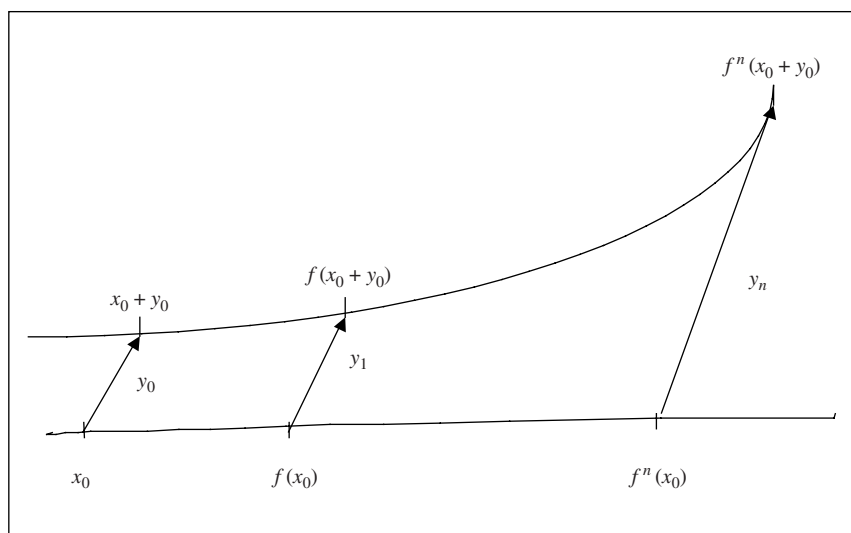
$$(6) \quad \frac{\|y_n\|}{\|y_0\|}$$

y hay crecimiento si  $\|y_n\| > \|y_0\|$  o decrecimiento si  $\|y_n\| < \|y_0\|$ . Aplicando la regla de la cadena tenemos que

$$(7) \quad y_n = f^{(n)'}(x_0) \cdot y_0$$

donde  $f^{(n)'}(x_0) = f'(x_{n-1}) \cdot f'(x_{n-2}) \cdot \dots \cdot f'(x_0)$ , por lo que  $\frac{\|y_n\|}{\|y_0\|} = |f^{(n)'}(x_n)|$ .

FIGURA 1  
EVOLUCIÓN DE LA DISTANCIA ENTRE DOS ÓRBITAS GENERADAS A PARTIR DE CONDICIONES INICIALES ARBITRARIAMENTE CERCANAS



El exponente de Lyapunov será, por tanto, el límite del logaritmo de ese cociente cuando el número de iteraciones sea suficientemente grande, es decir,

$$(8) \quad \lambda(x_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \left[ |f^{(n)'}(x_n)| \right]$$

Cuando el sistema objeto de estudio es  $k$ -dimensional, el análisis de la estabilidad en el sentido de Lyapunov se generaliza del valor de la derivada de la función al valor de la matriz jacobiana asociada a dicho sistema dinámico. Recordemos que la matriz jacobiana es la formada por las derivadas parciales con respecto a todas las variables de estado, es decir,

$$(9) \quad Df = \begin{pmatrix} \partial f_1 / \partial x_1 & \partial f_1 / \partial x_2 & \cdots & \partial f_1 / \partial x_k \\ \partial f_2 / \partial x_1 & \partial f_2 / \partial x_2 & \cdots & \partial f_2 / \partial x_k \\ \vdots & \vdots & \ddots & \vdots \\ \partial f_k / \partial x_1 & \partial f_k / \partial x_2 & \cdots & \partial f_k / \partial x_k \end{pmatrix}$$

Cuando el sistema es  $k$ -dimensional se debe generalizar la definición de exponente de Lyapunov, sustituyendo la derivada de la función por la matriz jacobiana o, más exactamente, por los autovalores de la misma, de manera que en lugar de hablar de exponente de Lyapunov hablamos de espectro de exponentes de Lyapunov, ya que habrá tantos exponentes como dimensión tenga el sistema. El sistema será inestable si el máximo exponente es positivo.

El  $j$ -ésimo exponente de Lyapunov para una condición inicial  $x_0$  se define como

$$(10) \quad \lambda_j(x_0) = \lim_{n \rightarrow \infty} \frac{1}{2n} \ln H_{nj}$$

para  $j = 1, \dots, k$  siendo  $H_{jn}$  los autovalores de la matriz  $H_n(x_0) = (Df^n(x_0))' Df^n(x_0)$ , donde  $Df^n(x_0) = Df(x_{n-1}) \cdot Df(x_{n-2}) \cdot \dots \cdot Df(x_0)$ .

## 2.2. Cálculo del Exponente de Lyapunov en Series Temporales

En economía no son conocidos los sistemas dinámicos que rigen el comportamiento de las variables económicas, pero sí se dispone de series temporales, que se supone que son observaciones o realizaciones de estos sistemas desconocidos, denominados sistemas generadores de datos. Por consiguiente, dado que generalmente no se conoce el modelo generador de datos, los exponentes de Lyapunov habrán de ser calculados no a través del sistema sino de la serie generada.

En los últimos tiempos, la estimación de los exponentes de Lyapunov a partir de una serie temporal ha experimentado un impulso de la mano de la aplicación de la metodología de redes neuronales. Tradicionalmente han existido dos métodos para la estimación de los exponentes de Lyapunov a partir de datos: el método directo y el método indirecto o del jacobiano. El primero fue propuesto por Hahn (1963) siendo Wolf (1985) el que generalizó su uso y el segundo por Eckmann y Ruelle (1985).

Los métodos directos tratan de estimar los exponentes de Lyapunov a partir de su propia definición, midiendo empíricamente la separación de las órbitas generadas a partir de puntos inicialmente próximos. En concreto, se calculan analizando la evolución en el espacio de fases de una elipsoide de condiciones iniciales. Cada exponente mide la tasa media de crecimiento exponencial a largo plazo de cada uno de los ejes del elipsoide y el máximo exponente de Lyapunov marcará la máxima tasa de divergencia de manera que será el más sencillo de calcular. En condiciones ideales (Brock, 1986; Eckmann y Ruelle, 1992), con infinitos datos y sin ningún ruido, el valor proporcionado converge al verdadero valor del exponente máximo de Lyapunov.

Sin embargo, en la realidad las series no son, obviamente, infinitas (es más, en economía las series disponibles suelen ser excesivamente cortas) y además suelen estar perturbadas por factores aleatorios exógenos (ruido). La presencia de ruido afecta muy negativamente a la estimación utilizando estos métodos ya que, en presencia de ruido, pueden proporcionarse exponentes espurios positivos, al incrementarse la divergencia entre las condiciones iniciales consideradas.

En estos casos surge la necesidad de la aplicación de otros métodos robustos a la presencia de ruido. Estos métodos no analizan directamente la evolución seguida por dos condiciones iniciales arbitrariamente próximas, buscando una estimación directa de la tasa de divergencia (o convergencia), sino que estiman esta tasa de divergencia indirectamente, basándose en una estimación previa de la matriz jacobiana (la matriz de derivadas parciales) del sistema subyacente desconocido, generador de la serie temporal. Por este motivo también se los conoce como métodos de estimación del jacobiano. Difieren unos de otros precisamente en la manera de estimar esta matriz.

En un primer momento, los métodos se basaban en estimación lineal local<sup>1</sup>. Estos métodos producen buenas estimaciones del exponente de Lyapunov cuando las series son suficientemente largas y limpias de ruido, circunstancias difíciles de conseguir en economía. Además, como muestran McCaffrey *et al.* (1992) el ajuste lineal genera exponentes espurios cuando la dimensión es elevada, de manera que se hace imprescindible el desarrollo de nuevas técnicas. Este desarrollo pasa, lógicamente, por disminuir las restricciones de la metodología propuesta anteriormente, en especial la referida a la linealidad. Es decir, en lugar de estimar una aproximación lineal al jacobiano, tratan de estimarlo mediante una función no lineal. Estos métodos presentan la ventaja fundamental de ofrecer la posibilidad de estimar los exponentes de Lyapunov a partir de series temporales (aplicando métodos de regresión no paramétrica no lineal), aunque las mismas procedan de sistemas que pueden tener perturbaciones estocásticas, y sin necesidad de abundantes datos.

Uno de los métodos más utilizados en estimación de funciones no lineales es el basado en redes neuronales. Estos métodos, inicialmente propuestos por Nychka *et al.* (1992) y Gencay y Dechert (1992), han funcionado muy bien en la

<sup>1</sup> Ver Eckman y Ruelle, 1985, Eckman *et al.*, 1985, Sano y Sawada, 1985, Sato *et al.*, 1987 y Abarbanel *et al.*, 1991, entre otros.

práctica con series cortas y con ruido, y asimismo han mostrado resultados favorables en simulación como se puede ver en Gallant y White (1992), Kuan y White (1994) y Barnett *et al.* (1995).

Por último, también muestran buenas propiedades teóricas (ver Nychka *et al.*, 1992; McCaffrey, 1991, y McCaffrey *et al.*, 1992), fundamentalmente su buen nivel de aproximación y la escasa sensibilidad al aumento de la dimensión del sistema. La propiedad de aproximación universal también se aplica a sus derivadas (Gallart y White, 1992), propiedad interesante en el contexto de los exponentes de Lyapunov. Los más utilizados han sido los programados por Ellner *et al.* (1992) llamado Lyapunov Exponents Neural NetS (LENNS) y Kuan *et al.* (1997) NETLE 3.01. En ambos casos se utiliza para la estimación del jacobiano el método de mínimos cuadrados no lineales, y se trabaja con una red neuronal SLFFN de una sola capa, trabajando con un número amplio de número de nodos y de *inputs*, eligiendo el mejor ajuste conseguido en función de la minimización del error cometido.

Adicionalmente, en los últimos tiempos se ha estado trabajando en el desarrollo de la distribución asintótica de probabilidad de los exponentes de Lyapunov calculados mediante redes neuronales. En concreto, Shintani y Linton (2003 y 2004) han probado que convergen asintóticamente a la distribución normal. La obtención de esta distribución es fundamental para extraer conclusiones fiables acerca de la estabilidad o inestabilidad del sistema generador de la serie temporal.

Se parte de una sucesión de escalares aleatorios  $\{x_t\}_{t=1}^T$  generada por el siguiente modelo no lineal autorregresivo  $x_t = f(x_{t-1}, \dots, x_{t-d}) + u_t$  donde  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  es una función no lineal y  $\{u_t\}_{t=1}^T$  es una sucesión de perturbaciones aleatorias con las hipótesis habituales. Este modelo puede ser expresado como  $Z_t = F(Z_{t-1}) + U_t$  donde  $Z_t = (x_t, \dots, x_{t-d+1})' \in \mathbb{R}^d$ ,  $f: \mathbb{R}^d \rightarrow \mathbb{R}^d$  y  $U_t = (u_t, 0, \dots, 0)'$ . El máximo exponente de Lyapunov del sistema dinámico está definido por

$$(11) \quad \lambda \equiv \lim_{M \rightarrow \infty} \frac{1}{2M} \ln v_1(T'_M T_M) \quad T_M = \prod_{t=1}^M J_{M-t} = J_{M-1} \cdot \dots \cdot J_0$$

donde  $v_1(A)$  es el máximo autovalor de una matriz  $A$  y  $J_t$  es la matriz jacobiana de la función  $F$  evaluada en  $Z_t$ . Específicamente,

$$(12) \quad J_t = \begin{bmatrix} \Delta f_{1t} & \Delta f_{2t} & \dots & \Delta f_{d-1,t} & \Delta f_{dt} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

para  $t = 0, 1, \dots, T-1$ , donde  $\Delta f_{jt} = D^{e_j} f(Z_t)$  para  $j = 1, \dots, d$  siendo  $e_j$  el  $j$ -ésimo vector elemental  $e_j = (0, \dots, 1, \dots, 0) \in \mathbb{R}^d$ .

Si  $\hat{f}(\cdot)$  es el estimador de  $f(\cdot)$  basado en redes neuronales, sustituyendo este estimador en la fórmula de la matriz jacobiana, se obtiene el estimador de la misma que se denota por  $\hat{J}_t$ . El estimador por redes neuronales del exponente de Lyapunov se define entonces como

$$(13) \quad \hat{\lambda}_M = \frac{1}{2M} \ln v_1(\hat{T}_M \hat{T}_M) \quad \hat{T}_M = \prod_{t=1}^M \hat{J}_{M-t} = \hat{J}_{M-1} \cdot \dots \cdot \hat{J}_0$$

donde  $M$  es el número de puntos utilizados para la estimación del exponente.

### 2.3. Redes Neuronales y Estimación del Máximo Exponente de Lyapunov

Las redes neuronales (ver, por ejemplo, Nychka *et al.*, 1992; Jungeilges, 1996 y White, 1989) son una clase de modelos no lineales inspirados por la arquitectura neuronal básica del cerebro humano. Reflejan la idea de un proceso de aprendizaje entendido como un procedimiento estadístico de tipo recursivo. Son modelos de tipo *input-output* en los que la función que relaciona la entrada con la salida está determinada por una estructura neuronal. Estos métodos están siendo aplicados a numerosas áreas tan diferentes como la psicología, la computación, la genética, la lingüística y la ingeniería debido a su capacidad para resolver una gran diversidad de problemas como la decodificación de lenguajes, reconocimiento de caracteres escritos a mano, decodificación de sucesiones genéticas proteicas... en definitiva, aplicando este tipo de técnicas se resuelven problemas relacionados con el reconocimiento de patrones, clasificación, detección de no linealidad y predicción no lineal.

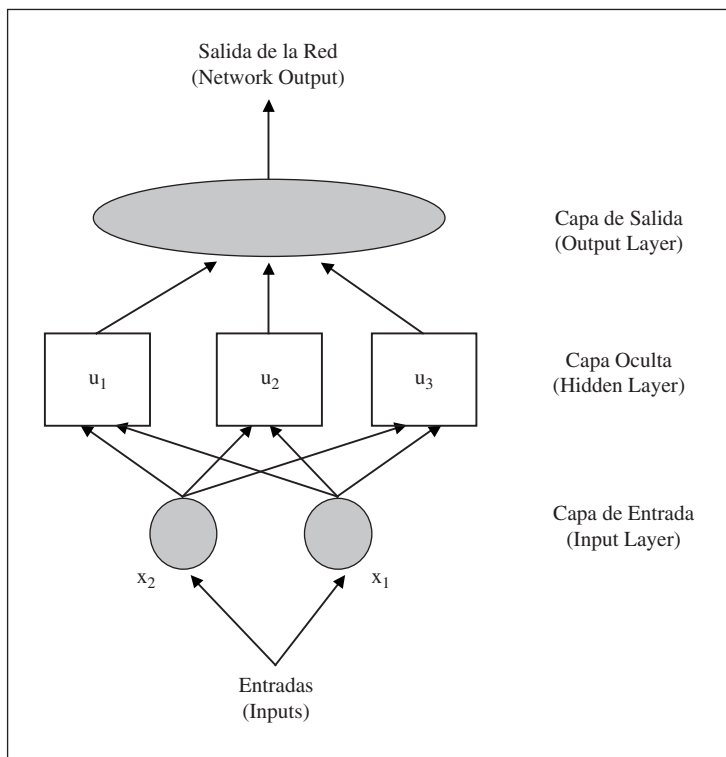
En el caso que nos ocupa, la estimación de los exponentes de Lyapunov, podemos decir que gracias a este tipo de modelos se puede aproximar funciones continuas en espacios finito dimensionales. Por consiguiente, es un método que proporciona una aproximación global a la función. Concretamente los autores proponen la utilización de una red simple (la información sólo fluye en un sentido y no existe retroalimentación) con una única capa oculta y una única salida. Este tipo de red neuronal se suele denotar como *SLFFN* (*single-layer feedforward network*) y se muestra en la Figura 2, para dos valores de entrada:  $x_1$  y  $x_2$ .

Su funcionamiento es el siguiente: las unidades de entrada transforman linealmente los *inputs* determinada por los denominados factores de intensidad asociado a la conexión entre el nodo *input*  $j$  y el nodo oculto  $i$ . Así, el *input* total que entra en el  $i$ -ésimo nodo oculto suponiendo  $d$  *inputs*  $x$  es  $\sum_{j=1}^d \chi_{ij} x_j$ . Cada uno de estos nodos ocultos lleva a cabo una transformación no lineal sobre el *input* total que le entra, aplicándole una función de activación y produciendo el *output*

$$(14) \quad o_i = \psi \left( \sum_{j=1}^m \chi_{ij} x_j + \chi_{i0} \right) \quad i = 1, \dots, k$$



FIGURA 2  
RED NEURONAL SINGLE-LAYER FEED FORWARD NETWORK (SLFFN)



de manera que habrá un output por unidad o nodo en la capa oculta. Esta función de activación es la misma para todas las unidades, pero cada una de ellas tiene su propio sesgo  $\chi_{i0}$  que representa un input externo o el nivel de actividad neuronal intrínseco. Esta función de activación  $\psi(u)$  tiene en general forma de ese<sup>2</sup>, con dos asíntotas, en cero cuando  $u \rightarrow -\infty$  y en uno cuando  $u \rightarrow \infty$ . Los autores

utilizan la función  $\psi(u) = \frac{e^u}{1 + e^u}$ .

Los outputs de la capa oculta pasan a través de las unidades de salida que llevan a cabo una transformación sobre su input total. Suponiendo  $q$  unidades outputs, la salida de la red será, por tanto

$$(15) \quad O_j = \Phi \left( \beta_j + \sum_{s=1}^k \beta_{js} o_s \right) \quad j = 1, \dots, q$$

<sup>2</sup> Se toman funciones de este tipo para modelizar a grandes rasgos las propiedades “umbral” de las neuronas reales.

En el caso particular de una unidad de salida, el output global se suele considerar como una transformación lineal del input global, de manera que

$$(16) \quad O = \sum_{i=1}^k \beta_i \psi(\chi_i' X + \chi_{i0})$$

para  $k$  unidades en la capa oculta y  $d$  inputs, donde  $\chi_i' = (\chi_{i1}, \chi_{i2}, \dots, \chi_{id})$  y  $X \in \mathbb{R}^d$  es el vector de inputs.

En definitiva, un modelo de redes neuronales general será un modelo que intenta aproximar, mediante la estructura especificada de capas y conexiones, las  $q$  unidades de salida  $y_i$   $i = 1, \dots, q$  mediante funciones no lineales de los  $d$  inputs,  $f_i(x_1, \dots, x_d)$   $i = 1, \dots, q$ . Se tratará, por tanto, de estimar el conjunto de ponderaciones y el número de nodos de la capa oculta necesarios para que los valores de entrada, después de pasar por la red, produzcan valores de salida lo más parecidos posible a los que realmente toman las variables. Esta aproximación se realiza en la fase denominada entrenamiento mediante el que, aplicando un proceso recursivo de ajuste mediante minimización del error cometido, la red intenta estimar las relaciones ocultas en los datos.

Según este enfoque y para el caso concreto de la estimación del exponente máximo de Lyapunov, por tanto, la estimación de la función que utilizaremos para estimar la matriz jacobiana y a partir de ella el exponente de Lyapunov máximo será

$$(17) \quad \hat{f}(X) = \sum_{i=1}^K \beta_i \psi(\chi_i' X + \chi_{i0})$$

donde  $\psi(u) = \frac{e^u}{1+e^u}$  es la función de distribución logística  $X = x_t^m \in \mathbb{R}^m$  y  $\chi_i \in \mathbb{R}^m$ . Es decir,

$$(18) \quad x_t = \hat{f}(x_t^m) = \sum_{i=1}^K \beta_i \psi \left( \chi_{i0} + \sum_{j=1}^m \chi_{ij} x_{t-j\tau} \right)$$

donde  $m$  y  $\tau$  son, respectivamente, la dimensión de absorción y el retardo necesarios para una correcta reconstrucción<sup>3</sup>. Los parámetros se estiman aplicando mínimos cuadrados no lineales dentro de un algoritmo recursivo de minimiza-

<sup>3</sup> Esta idea se basa en las nociones de la reconstrucción del atractor que garantiza, gracias al Teorema de Takens (Takens, 1985), la invariabilidad de la dinámica subyacente en el sistema y en su reconstrucción utilizando la serie temporal generada y los parámetros adecuados; en concreto, los exponentes de Lyapunov se mantienen invariantes. La elección de estos parámetros, en concreto, la dimensión de absorción y el retardo, es crítica para la obtención de resultados fiables, desarrollándose gran cantidad de métodos, siendo uno de los más populares la maximización de la bondad de la predicción (Casdagli 1989 y 1992, Sugihara y May, 1990).

ción del error cuadrático medio cometido. Este método de aproximación de la función presenta la ventaja de no ser sensible ante un incremento en la dimensión de absorción. Los autores comprueban que esta estimación es factible para sistemas con ruido con un número escaso de observaciones.

En cuanto al número de nodos de la capa oculta, se trata de la elección más arbitraria en esta metodología. En general, se suele considerar que el número correcto será aquel que minimice el error cometido, evaluando este error por alguno de los métodos habituales.

Por tanto, la principal utilidad de los modelos de redes neuronales referidos al cálculo de los exponentes de Lyapunov se deriva del hecho de su capacidad de aproximar cualquier función no lineal. De hecho, proporcionan la clase de aproximaciones más rica y flexible. Concretamente, como señala White (1989), se ha demostrado que una red *SLFNN* con una función en forma de ese puede aproximar cualquier función medible con cualquier exactitud (en la métrica apropiada), dado un número suficiente de unidades escondidas.

En el presente trabajo se ha utilizado el algoritmo *LENNS* (Lyapunov Exponents Neural NetS, 1992). Trabaja con el ajuste de redes neuronales recurrentes de una capa, y el algoritmo de optimización es el estándar BFGS, con el algoritmo BHHH de longitud de paso (Berndt *et al.*, 1974) y el criterio de convergencia sugerido por Gill *et al.* (1981). Su principal ventaja es la de no restringir la reconstrucción, es decir, que calcula el exponente para un amplio rango de posibles valores de la dimensión de absorción y del retardo, además de los nodos utilizados en la capa de la red (en nuestro caso, las tres variables oscilan entre uno y 10). El exponente de Lyapunov elegido será aquel para el que se maximice la bondad de predicción. Basándose en técnicas de cros-validación, y teniendo en cuenta que se ajustan modelos de redes neuronales a datos generados a partir de sistemas no lineales y ruidosos, se utiliza la minimización del Criterio de Información de Bayes (Nychka *et al.*, 1992; Granger y Teräsvirta, 1992). Básicamente este criterio consiste en penalizar el error cuadrático medio por los parámetros utilizados en el ajuste:

$$BIC = \frac{1}{2} \cdot \left( 1 + \ln(2 \cdot \pi) + 2 \cdot \ln \left( \frac{1}{N_m} \sum_t e_t^2 \right) + k \frac{\ln N_m}{N_m} \right)$$

donde  $k$  es el número de nodos o unidades de la capa oculta,  $N_m$  es la longitud de la serie de  $m$ -historias y  $e_t$  es el error cometido por la red en el instante  $t$ .

### 3. MEDICIÓN DE LA ESTABILIDAD DE LOS TIPOS DE CAMBIO EN EL PROCESO DE CONVERGENCIA EUROPEA (1997-2002)

Se han considerado datos de tipos de cambio con respecto al dólar de las distintas monedas de la Comunidad Económica Europea, esto es, Alemania

(DEM), Austria (ATS), Bélgica (BEF), Dinamarca (DKK), España (ESP), Finlandia (FIM), Francia (FRF), Gran Bretaña (GBP), Grecia (GRD), Holanda (NLG), Irlanda (IEP), Italia (ITL), Luxemburgo (LUF), Noruega (NOK), Portugal (PTE), Suecia (SEK) y Suiza (CHF). El período de tiempo considerado ha sido desde el 2 de enero de 1997 hasta el 31 de diciembre de 2002 (1504 datos). De estas monedas se han considerado un primer grupo que sólo llega hasta el 31 de diciembre de 1998, constituido por las de los Estados que forman parte de la Unión, con el tipo de cambio fijo y por tanto una misma evolución a partir de este momento. Estos Estados, que cumplían las condiciones de convergencia y que conforman la primera Unión Monetaria son: Alemania, Austria, Bélgica, Finlandia, Francia, Irlanda, Italia, Luxemburgo, Holanda, Portugal y España. Otros dos países que cumplieron los criterios eligieron no participar: Dinamarca y Reino Unido. Y otros dos países no cumplieron los criterios: Grecia (Grecia entró a formar parte de la Unión el 2 de enero de 2001) y Suecia. Por este motivo, y para conseguir un elemento común en toda la serie, se ha considerado adicionalmente la serie del tipo de cambio del euro frente al dólar en todo el período considerado.

Para cada variable se ha subdividido el período 1997-2002, de acuerdo a eventos importantes que podrían haber influido en la evolución de los tipos de cambio mencionados. El Cuadro 1 resume los eventos considerados.

CUADRO 1  
EVENTOS CONSIDERADOS PARA EL CORTE DE LAS SERIES

Fecha de corte de la serie	Evento	Observaciones
17.06.1997	Cumbre de Amsterdam	La cumbre de Amsterdam (16 y 17 de junio de 1997), con la aprobación del denominado Pacto de Estabilidad y Crecimiento para dar credibilidad al euro, limitando el déficit permitido y llegando a acuerdos sobre seguridad, empleo, política exterior, defensa y legitimidad de las instituciones europeas, que se ratificaron en el Tratado de Amsterdam.
4.05.1998	Consejo Europeo de Bruselas	El Consejo Europeo de Bruselas (3 de mayo de 1998), en el que se acuerdan los Estados que conformarían la primera Unión Monetaria.
4.01.1999	Consejo de la Unión Europea	El Consejo de la Unión Europea (31 de diciembre de 1998) establece de manera definitiva los tipos de conversión entre el euro y las monedas nacionales de los miembros participantes de la Unión.

27.03.2000	Consejo Europeo de Lisboa	El Consejo Europeo de Lisboa (23 y 24 de marzo de 2000), en el que, para potenciar el dinamismo económico en los países de la Unión, se adoptan diferentes estrategias en materia de empleo, reforma económica y cohesión social.
25.09.2000	Intervenciones de Bancos Centrales; no danés	Intervenciones del Banco Central Europeo, Reserva Federal de Estados Unidos y Banco de Japón a favor del euro (septiembre de 2000). En este mismo mes (28 de septiembre) Dinamarca rechaza en referéndum su incorporación a la zona Euro.
7.06.2001	No ratificación por Irlanda	Irlanda, mediante referéndum, no ratifica el Tratado de Niza (8 de junio de 2000).
1.01.2002	Entrada en vigor del euro	Entrada en vigor del euro (1 de enero de 2002) como moneda única en Bélgica, Alemania, Finlandia, Francia, Grecia, Irlanda, Italia, Luxemburgo, Países Bajos, Austria, Portugal y España.

Fuente: Elaboración propia.

De esta manera, las series consideradas se han dividido en ocho subperíodos distintos, para así medir la influencia de estos distintos eventos en la estabilidad de los tipos de cambio. Para analizar esta influencia, así como la diferencia entre volatilidad e inestabilidad de los tipos de cambio, se ha calculado, para cada una de estas subseries, el máximo exponente de Lyapunov. A continuación se reflejan los resultados, primero para el período que transcurre desde el 2 de enero de 1997 hasta el 31 de diciembre de 1998 para todas las monedas, y luego para el período completo para las monedas que no entraron en la Unión, el euro y el yen. En los dos casos se proporciona una primera tabla con los exponentes de Lyapunov, y sus respectivas tasas de variación relativa global, y otra segunda tabla con sus variaciones relativas período a período.

TABLA 1  
MAXIMO EXPONENTE DE LYAPUNOV ANTES DEL EURO (BLOQUE 1)

	Exponente de Lyapunov											
	ATS	BEF	FRF	DEM	GRD	ITL	ESP	PTE	IEP	FIM	NLG	LUF
<b>Per. 1</b>	0,096	0,155	0,129	0,085	0,041	0,077	0,183	0,022	0,123	0,075	0,080	0,155
<b>Per. 2</b>	0,053	0,071	0,074	0,063	0,031	0,075	0,056	0,084	0,102	0,079	0,063	0,055
<b>Per. 3</b>	0,057	0,020	0,072	0,027	0,512	0,035	0,150	0,062	0,041	0,048	0,019	0,046
<b>T.V.</b>	-0,406	-0,871	-0,442	-0,682	11,49	-0,545	-0,180	1,818	-0,667	-0,360	-0,763	-0,703

Nota: Per. corresponde a cada período considerado y T.V. es la tasa de variación relativa global.

Fuente: Elaboración propia.

TABLA 2  
MAXIMO EXPONENTE DE LYAPUNOV ANTES DEL EURO (BLOQUE 2)

	Exponente de Lyapunov				
	GBP	DKK	NOK	SEK	CHF
<b>Per. 1</b>	0,076	0,089	-4,49	-0,430	0,016
<b>Per. 2</b>	0,029	0,044	0,077	0,044	0,049
<b>Per. 3</b>	0,122	0,018	0,384	0,038	0,019
<b>T.V.</b>	0,605	-0,798	1,085	1,088	0,187

Nota: Per. corresponde a cada período considerado y T.V. es la tasa de variación relativa global.  
Fuente: Elaboración propia.

TABLA 3  
MAXIMO EXPONENTE DE LYAPUNOV (BLOQUE 1)

	Exponente de Lyapunov					
	GBP	DKK	NOK	SEK	CHF	EURO
<b>Per. 1</b>	0,076	0,089	-4,49	-0,430	0,016	0,146
<b>Per. 2</b>	0,029	0,044	0,077	0,044	0,049	0,050
<b>Per. 3</b>	0,122	0,018	0,384	0,038	0,019	0,067
<b>Per. 4</b>	-0,040	-0,029	-0,047	-0,057	-0,030	-0,027
<b>Per. 5</b>	-0,086	0,026	0,039	0,087	0,194	0,026
<b>Per. 6</b>	-0,048	-0,025	0,434	-0,010	-0,033	-0,025
<b>Per. 7</b>	0,126	0,061	0,332	0,154	0,082	0,572
<b>Per. 8</b>	-0,062	-0,028	-0,020	-0,022	-0,020	-0,028
<b>T.V.</b>	-1,816	-1,315	0,995	0,948	-2,250	-1,192

Nota: Per. corresponde a cada período considerado y T.V. es la tasa de variación relativa global.  
Fuente: Elaboración propia.

TABLA 4  
VARIACION RELATIVA DEL MAXIMO EXPONENTE  
DE LYAPUNOV (BLOQUE 2)

	Variación Relativa del Exponente de Lyapunov					
	GBP	DKK	NOK	SEK	CHF	EURO
<b>Per. 1</b>	-0,618	-0,506	-1,017	-1,102	2,063	-0,658
<b>Per. 2</b>	3,207	-0,591	3,987	-0,136	-0,612	0,340
<b>Per. 3</b>	-1,328	-2,611	-1,122	-2,500	-2,579	-1,403
<b>Per. 4</b>	1,150	-1,897	-1,830	-2,526	-7,467	-1,963
<b>Per. 5</b>	-0,442	-1,962	10,128	-1,115	-1,170	-1,962
<b>Per. 6</b>	-3,625	-3,440	-0,235	-16,400	-3,485	-23,880
<b>Per. 7</b>	-1,492	-1,459	-1,060	-1,143	-1,244	-1,049

Nota: Per. corresponde a cada período considerado y T.V. es la tasa de variación relativa global.  
Fuente: Elaboración propia.

Lo primero que llama la atención en estas tablas es que ninguno de los países de la Unión que no se incorporaron al euro aumentó su estabilidad a lo largo del período considerado anterior a la entrada del mismo. En concreto, hablamos de Gran Bretaña, Noruega, Dinamarca, Suecia y Suiza, que presentan variaciones relativas positivas del máximo exponente de Lyapunov, salvo en el caso de Dinamarca. Y recordemos que Dinamarca cumplió los denominados ‘criterios de convergencia’, pero escogió no incorporarse a la zona euro. Es curioso, sin embargo, el caso de Gran Bretaña, que cumplió los criterios de convergencia pero muestra un crecimiento de la inestabilidad en el tipo de cambio, pero concentrado en el tercer período anterior al euro.

En cuanto a los países que se incorporaron a la zona euro, en general mostraron un crecimiento de la estabilidad en el comportamiento de los tipos de cambio en el período anterior al euro, salvo en el caso de Portugal y Grecia. Recordemos, en este sentido, que Grecia cumplió los criterios de convergencia en un momento posterior, mientras que la economía portuguesa, salvo en los últimos tiempos, ha mostrado síntomas de mayor inestabilidad.

Si se considera el período completo 1997-2002 existe una clara diferenciación en el período anterior y posterior al euro. Mientras que antes del euro los máximos exponentes de Lyapunov son positivos, después del euro dichos exponentes pasan a ser, en su mayor parte, negativos. Es decir, que se pasa de un sistema dinámico subyacente inestable a un sistema estable, excluyendo además la posibilidad de existencia de dinámica caótica. Incluso, esto ocurre independientemente de si el país considerado pertenece o no a la zona euro, lo cual indica que este proceso de convergencia y la Unión Monetaria subsiguiente contribuyó a la estabilización económica.

Es curioso además que en el período quinto se observa que los exponentes de Lyapunov para todos los países considerados exceptuando a Gran Bretaña pasan a ser positivos, indicando un claro aumento de la inestabilidad en el sistema dinámico subyacente, que se contrarresta en el período siguiente gracias a la intervención de los Bancos Centrales a favor del euro. Por otro lado, el efecto estabilizador de esta intervención compensa el posible efecto desestabilizador de la negativa danesa a la Unión. El siguiente punto de inflexión se produce con la negativa irlandesa, que nuevamente se traduce en un aumento de inestabilidad del tipo de cambio, siendo de nuevo los máximos exponentes de Lyapunov positivos. Esta inestabilidad cesa con la entrada en vigor del euro, que refuerza la estabilidad, mostrando otra vez exponentes positivos.

Por último, pero no menos importante, parece claro que a lo largo de todo el período considerado se ha conseguido una clara armonización en las propiedades de estabilidad de los sistemas dinámicos de los países considerados, ya que, comenzando con valores muy diversos del máximo exponente de Lyapunov, los últimos valores observados son todos muy similares, en torno a  $-0,02$ , excepto en el caso de Gran Bretaña, con  $-0,058$  que, de todos modos, sigue siendo un valor bastante próximo al resto.

#### 4. CONCLUSIONES

En el presente trabajo queda demostrada la validez del cálculo de los exponentes de Lyapunov para medir la estabilidad de la evolución de los tipos de cambio en el proceso de convergencia europeo, ya que los resultados observados empíricamente son coherentes con los hechos políticos y económicos reales.

Además, los exponentes de Lyapunov ponen de manifiesto el diferente comportamiento, antes de la creación de la zona euro, de los países que la conformaron y aquellos que no cumplieron los criterios de convergencia, o que decidieron no incorporarse a la zona euro. También se pone de manifiesto el cambio que supuso la creación de la Unión Monetaria para la estabilidad de los tipos de cambio en Europa, ya que pasaron de mostrar valores positivos del máximo exponente de Lyapunov a valores negativos. A lo largo del período considerado se observa un aumento de la estabilidad, además de una clara convergencia en las propiedades dinámicas del sistema subyacente, ya que los valores obtenidos para los exponentes de Lyapunov son similares para los distintos países. Por último, es interesante resaltar el ligeramente distinto comportamiento observado en Gran Bretaña.

#### REFERENCIAS

- Abarbanel, H.D.I.; R. Brown y M. B. Kennel (1991), *Local Lyapunov exponents computed from observed data*. Journal of Nonlinear Science 1: 175-199.
- Abhayandar, A.; L.S. Copeland y W. Wong (1997), *Uncovering nonlinear structure in real-time stock-market indexes: the S&P 500, the DAX, the Nikkei 225 and the FTSE-100*. Journal of Business and Economic Statistics 15: 1-14.
- Barnett, W.A.; A.R. Gallant; M.J. Hinich; J.A. Jungeilges; D.T. Kaplan y M.J. Jensen (1995), *Robustness of nonlinearity and chaos tests to measurement error, inference method, and sample size*. Journal of Economic Behavior and Organization 27: 301-320.
- Bask, M. (2002), *A positive Lyapunov exponent in Swedish exchange rates?* Chaos, Solitons and Fractals 14: 1295-1304.
- Bask, M. y X. de Luna (2002), *Characterizing the degree of stability of nonlinear dynamic models*. Studies in Nonlinear Dynamics & Econometrics 6, 1: art. 3.
- Berndt, E.K.; G.H. Hall; R.E. Hall y J.A. Hausman (1974), *Estimation and inference in nonlinear structural models*. Annals of Economics and Social Measurement 3/4: 653-665.
- Besné, R.; J.R. Canedo; B. Pérez (1998), *La Unión Europea: historia, instituciones y sistema jurídico*. Bilbao, Universidad de Deusto.
- Brock, W.A. (1986), *Distinguishing random and deterministic systems. abridged version*. Journal of Economic Theory 40: 168-195.
- Casdagli, M. (1989), *Nonlinear prediction of chaotic time series*. Physica D 35: 335-356.
- Casdagli, M. (1992), *Chaos and deterministic versus stochastic modelling*. Journal of the Royal Statistical Society B 54, 2: 303-328.
- Dechert, W.D. y R. Gencay (1992), *Lyapunov exponents as a nonparametric diagnostic for stability analysis*. Journal of Applied Econometrics 7: S41-S60.



- Eckman, J.-P y D. Ruelle (1985), *Ergodic theory of chaos and strange attractors*. Reviews of Modern Physics 57, 3: 617-656.
- Eckman, J.-P. y D. Ruelle (1992), Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems. Physica D 56: 185-187.
- Ellner, S.; D.W. Nychka y A.R. Gallant (1991), *LENNS, a program to estimate the dominant Lyapunov exponent of noisy nonlinear system from time series data*. Institute of Statistics Mimeo Series #2235, Statistics Department, North Carolina State University, Raleigh, NC 27695-8203.
- Gallant, A.R. y H. White (1992), On learning the derivatives of an unknown mapping with multilayer feedforward networks. Neural Networks 5: 12-138.
- Gencay, R. y W.D. Dechert (1992), *An algorithm for the n Lyapunov Exponents of an n dimensional unknown dynamical system*. Physica D 59: 142-157.
- Gill, P.E.; W. Murray y M.H. Wright (1981), *Practical Optimization*. London y New York. Academic Press.
- Granger, C.W.J. y T. Teräsvirta (1992), *Experiments in modeling nonlinear relationships between time series*. En *Nonlinear modelling and forecasting* (Eds. M. Casdagli y S. Eubank). Santa Fe Institute Studies in the Sciences of Complexity, Proceedings v. XII. New York. Addison-Wesley: 199-226.
- Jungeilges, J.A. (1996), *Operational characteristics of White's test*. En *Nonlinear dynamics and economics* (Eds. W.A. Barnett, A.P. Kirman y M. Salmon). Cambridge. Cambridge University Press: 219-266.
- Kuan, C.M.; T. Liu y R. Gencay (1997), *Nettle 3.01: feedforward neural networks and Lyapunov exponents estimation*. Ball State University.
- Kuan, C.M. y H. White (1994), *Artificial neural networks. an econometric perspective*. Econometric Reviews 13: 1-91.
- Lyapunov, A.M. (1892), *The general problem of the stability of motion*. Mathematical Society of Kharkov. PhD. Traducido en International Journal of Control (1992) 55: 131-773.
- McCaffrey, D.F.; S. Ellner; A.R. Gallant y D.W. Nychka (1992), *Estimating the Lyapunov exponent of a chaotic system with nonparametric regression*. Journal of the American Statistical Association 87: 682-695.
- Nychka, D.; S. Ellner; A.R. Gallant y D. McCaffrey (1992), *Finding chaos in noisy systems*; Journal of the Royal Statistical Society B. 54, 2: 399-426.
- Sano, M. y Y. Sawada (1985), *Measurement of the Lyapunov Spectrum from a chaotic time series*. Physical Review Letters 55, 10: 1082-1085.
- Sato, S.; M. Sano e Y. Sawada (1987), *Practical methods of measuring the generalized dimension and the largest Lyapunov exponent in high dimensional chaotic systems*. Progress of Theoretical Physics 77, 1: 1-5.
- Serletis, A. (1995), *Random walks, breaking trend functions, and the chaotic structure of the velocity of money*. Journal of Business and Economic Statistics 13: 453-458.
- Shintani, M. y O. Linton (2003), *Is there chaos in the world economy? A nonparametric test using consistent standard errors*. International Economic Review 44, 1: 331-358.
- Shintani, M. y O. Linton (2004), *Nonparametric neural networkd estimation of Lyapunov exponents and a direct test of chaos*. Journal of Econometrics 120: 1-33.
- Sugihara, G. y R.M. May (1990), *Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series*. Nature 334: 734-741.

- Takens, F. (1986), *Detecting strange attractors in turbulence* en D.A. Rand y L.-S. Young (eds.): *Lecture Notes in Mathematics: Dynamical Systems and Turbulence 89*; New York, Springer-Verlag: 366-381.
- Tamames, R. (1998), *Unión Monetaria y el Euro: la recta final*. Madrid, Espasa Calpe.
- Wang, H.; G. Chen y J. Lü (2004), Complex dynamical behaviors of daily data series in stock exchange. *Physics Letters A* 333: 246-255.
- White, A. (1989), *Some asymptotic results for learning in single hidden-layer feedforward network models*. *Journal of the American Statistical Association* 84, 408: 1003-1013.
- Wolf, A.; J.B. Swift; H.L. Swinney y J.A. Vastano (1985), *Determining Lyapunov exponents from a time series*. *Physica D* 16: 285-317.
- Weston, R. y P. Premachandra (2004), *A chaotic analysis of the New Zealand exchange rate 1985-2004*. NZAE 2004.

## EVIDENCE OF A BANK LENDING CHANNEL FOR ARGENTINA AND COLOMBIA\*

JOSÉ GÓMEZ-GONZÁLEZ

*Cornell University and Banco de la República*

FERNANDO GROSZ

*Cornell University*

*In this paper we find empirical evidence of bank lending channel for Colombia and Argentina. For Argentina, we do not find evidence that changes in the interbank interest rate affect the growth rate of total loans directly. However, the interbank interest rate affects the loan supply through its interactions with capitalization and liquidity. For Colombia, there is direct bank lending channel, which is reinforced through interactions with capitalization and liquidity. Also, using a panel data of more than 3,300 firms, we provide additional support to the existence of a bank lending channel for Colombia.*

*JEL classification: E5; E52; E59; G21*

*Keywords: Monetary Transmission; Bank Lending Channel; Argentina; Colombia.*

### 1. INTRODUCTION

While economists agree that monetary policy can affect real output in the short run, there is still a lot of controversy on how does monetary policy exactly operate.

The most commonly discussed channel, the traditional interest rate channel, suggests that when the Fed decreases money supply (exchanging bonds for bank reserves), nominal and real interest rates will increase –the effect of monetary policy on the real interest rates works through an assumption of sticky prices in the short run. Consequently, investment and present consumption decrease, and therefore aggregate demand (AD) decreases. However, as Bernanke and Gertler (1995) point out, empirical studies of the interest rate channel have

\* We want to thank comments received from the participants at a seminar in Cornell University. We are also indebted to Nick Kiefer, Karl Shell, Levon Barseghyan, Tao Zhu and an anonymous referee for valuable comments and suggestions. All remaining errors are our own, and opinions here do not necessarily reflect those of the Banco de la República nor its Board of Directors.  
Email: jeg44@cornell.edu, fag3@cornell.edu

not been entirely successful in explaining large changes in GDP and AD components due to moderate changes in the interest rates induced by monetary policy. These empirical findings have led to a vast literature that tries to identify and quantify other monetary transmission mechanisms, which complement and amplify the interest rate channel.

A long-standing question is whether financial institutions in general, and banks in particular, play an important role in the transmission of monetary policy to the real economy. The literature identifies two channels in which financial institutions might play a significant role, namely, the balance sheet channel (or broad credit channel) and the bank lending channel (or narrow credit channel).

The balance sheet channel was first introduced by Bernanke and Gertler (1989). The main idea of this mechanism is that in the presence of imperfect capital markets, informational asymmetries between borrowers and lenders cause a gap in the cost of internal and external sources of funding to borrowers. In general, this gap has a negative relation with the collateral of the borrower. A contractionary monetary policy has the effect of increasing real interest rates, therefore reducing the value of assets that act as collateral, which has the effect of deteriorating credit worthiness of borrowers. Therefore, consumption and investment plans that would be profitable if financed entirely with internal sources of funding are no longer profitable when financed partially with external sources. This effect leads to a lower level of AD in the economy, magnifying the effects of the interest rate channel. Note that there is no specific role played by banks in the broad credit channel.

On the other hand, the bank lending channel gives a specific role to banks. The basic idea is that a contractionary monetary policy that reduces bank deposits creates a need for alternative funds in order to maintain the level of loans. If such alternative funds are scarce or not available, then banks will necessarily reduce their loan supply, affecting negatively consumption and investment plans. Therefore, the bank lending channel amplifies the effect on AD of a contractionary monetary policy.

In order to have a bank lending channel two conditions are required: first, some firms must be dependent on bank loans; second, the central bank must be able to shift bank loan supply schedules. Regarding the first condition, there is evidence that suggests that small firms are bank dependent<sup>1</sup>. This occurs because banks have a comparative advantage in the sense of having lower costs of obtaining information about (and monitoring) their customers than other investors. Also, small firms generally lack access to securities markets. This effect will be more important for countries with less developed capital markets.

With respect to the second condition, a contractionary monetary policy has the effect of reducing the aggregate level of deposits<sup>2</sup>. Since these are one of

<sup>1</sup> For US, see for example Fazzari *et al.* (1988).

<sup>2</sup> Bernanke and Blinder (1992) show that aggregate deposit fall immediately when the Fed tightens monetary policy. Using bank disaggregated data, Kashyap and Stein (1995) also provide evidence that supports this.

the least expensive sources of financing for banks, it will be costly (for some banks) and even impossible (for others)<sup>3</sup> to offset the shortage in deposits with other sources of funding. In particular, if the Modigliani-Miller financial irrelevance theorem (Modigliani and Miller, 1958) does not hold for the banking firm, some banks will not be able to obtain loanable funds required to maintain their level of lending, and therefore their loan supply will drop<sup>4</sup>. Then financial variables that measure banks' financial health can play an important role, in the sense that banks with weak balance sheets are more affected by informational asymmetries than banks with stronger balance sheets.

In this paper, we look for evidence consistent with the bank lending channel literature for Argentina and Colombia. We find evidence that supports the existence of a lending channel of monetary policy for Colombia, but there does not seem to exist a strong evidence of a bank lending channel for Argentina. We do not find evidence that changes in the interbank interest rate affect the growth rate of total loans directly in Argentina. However, it does indirectly through interactions: the interbank interest rate affects the loan supply through its interactions with capitalization and liquidity. As for Colombia, there is direct bank lending channel, which is reinforced through interactions with capitalization and liquidity. Also, using a panel data of more than 3300 Colombian firms, we provide additional support to the existence of a bank lending channel for Colombia.

The rest of the paper is organized as follows. Section 2 presents a theoretical model developed by Kishan and Opiela (2000), which has testable implications for the empirical analysis. Section 3 discusses the data and summary statistics and evidence of a bank lending channel for other countries that has been established in other's work. Section 4 presents the empirical analysis and results, and Section 5 concludes.

## 2. A MOTIVATING MODEL

Kishan and Opiela (2000) provide evidence of the existence of a bank lending channel of monetary policy in the US from 1980 to 1995. Following

<sup>3</sup> Because demand deposits are insured, they are less subject to informational asymmetries relative to other sources of funding (i.e. large time deposits). Meanwhile, other sources of funding for banks are uninsured, which makes them more likely to suffer information problems. Therefore, bank characteristics that are not so important for the obtention of deposits become very relevant for the acquisition of other sources of funds, such as large CDs.

<sup>4</sup> Another way in which a contractionary monetary policy can affect bank lending is through its impact on the capitalization ratio. Banks face interest rate risk given their role in maturity transformation: they hold long term assets (many of them with a fixed interest rate), which they finance issuing short term liabilities. Therefore, a contractionary monetary policy, which increases short-term interest rates, increases the debt of the banks and decreases the net present value of its assets, thus reducing bank profits. If banks cannot reduce dividend payments substantially, then equity is reduced. Given the minimum capitalization requirements, some banks that initially had a low capitalization ratio, will have to cut lending to meet the capitalization requirement as a short term response.

Kashyap and Stein (1995), they try to identify the bank lending channel by studying cross sectional differences on the response of bank lending to monetary policy. To do so, they develop a model of a representative bank that has three assets, namely required reserves ( $RR$ ), loans ( $LN$ ) and securities ( $SEC$ ); and three liabilities, namely, demand deposits ( $DD$ ), large time deposits ( $TD$ ) and capital ( $K$ ). For simplicity, the bank does not hold excess reserves, so  $RR = \alpha DD$ , where  $\alpha \in [0,1]$  is determined by the central bank.

$DD$  are assumed to be inversely related to the Fed funds rate:  $DD = a_0 - a_1 r_{ff}$ . A bank is assumed to have market power in the  $TD$  market as well as in the  $LN$  market. Thus, it can raise  $TD$  by increasing its rate ( $r_{TD}$ ) over the mean rate in the market ( $\overline{r_{TD}}$ ), and can change loans by moving its loan rate ( $r_{LN}$ ) with respect to the mean rate in the loan market ( $\overline{r_{LN}}$ ):

$$(1) \quad TD = b_0 + b_1(r_{TD} - \overline{r_{TD}})$$

$$(2) \quad LN = d_0 - d_1(r_{LN} - \overline{r_{LN}})$$

Capital markets are assumed to be imperfect. This is introduced by assuming that the interest rate sensitivities of  $TD$  and  $LN$  respectively ( $b_1, d_1$ )  $> 0$ , depend on bank size and capitalization. Specifically, it is assumed that  $b_1$  depends positively on both bank size and capitalization, following the idea that larger and better capitalized banks will find it easier to raise funds by issuing time deposits. Meanwhile,  $d_1$  depends positively on bank size only, reflecting the idea that larger banks tend to give credit to larger firms which have better access to alternative sources of funding. Thus, larger banks have a demand for  $LN$  which is more sensible to changes in the interest rate than that of small banks<sup>5</sup>.

$$(3) \quad b_1 = b_1^+(A, K)$$

$$(4) \quad d_1 = d_1^+(A)$$

Securities are held as a buffer stock against liquidity shocks, and the mean market rates of  $TD$ ,  $SEC$  and  $LN$  are assumed to be directly related to the Fed funds rate with fixed spreads:

$$(5) \quad SEC = c_0 + c_1 DD - RR$$

$$(6) \quad \overline{r_{TD}} = e_0 + \phi r_{ff}$$

$$(7) \quad \overline{r_{SEC}} = f_0 + \phi r_{ff}$$

$$(8) \quad \overline{r_{LN}} = g_0 + \phi r_{ff}$$

<sup>5</sup> Kishan and Opiela (2000) assume that while large banks can also attend small firms, their major clients are large firms with access to capital markets.

Banks are assumed to choose  $LN$ ,  $TD$  and  $SEC$  to maximize profits,

$$(9) \quad Profit = (\overline{r_{LN}} - \Phi)LN + \overline{r_{SEC}}SEC - \overline{r_{DD}}DD - \overline{r_{TD}}TD$$

subject to the balance sheet constraint ( $LN + SEC + RR = DD + TD + K$ ) and the equations given above.  $\Phi LN$  represents loan losses.

The first order conditions of this maximization problem yield the optimal portfolio for the bank (i.e.,  $LN$ ,  $SEC$ ,  $TD$ ). Taking derivatives of  $LN$ ,  $SEC$  and  $TD$  with respect to the Fed funds rate generates some testable implications. In particular, assuming  $c_1 < 1$ , the model predicts that an increase in this rate should increase  $TD$ , decrease  $LN$ , and have an ambiguous effect on  $SEC$  (the sign of the partial derivative will depend on the values of the parameters, thus becoming an empirical issue).

$$(10) \quad \frac{\partial LN}{\partial r_{ff}} = -\frac{a_1 d_1 (1 - c_1)}{b_1 + d_1} < 0$$

$$(11) \quad \frac{\partial TD}{\partial r_{ff}} = \frac{a_1 b_1 (1 - c_1)}{b_1 + d_1} > 0$$

$$(12) \quad \frac{\partial SEC}{\partial r_{ff}} = -a_1 (c_1 - \alpha) \begin{matrix} \leq \\ \geq \end{matrix} 0$$

More interesting testable implications, however, derive from the introduction of the dependence of interest rate sensitivities of  $LC$  and  $TD$  on bank size and capitalization:

$$(13) \quad \frac{\partial(\frac{\partial LN}{\partial r_{ff}})}{\partial A} = \frac{\partial(\frac{\partial TD}{\partial r_{ff}})}{\partial A} = \frac{a_1 (1 - c_1) (b_1 \frac{\partial d_1}{\partial A} - d_1 \frac{\partial b_1}{\partial A})}{(b_1 + d_1)^2} \begin{matrix} \leq \\ \geq \end{matrix} 0$$

$$(14) \quad \frac{\partial(\frac{\partial LN}{\partial r_{ff}})}{\partial K} = \frac{\partial(\frac{\partial TD}{\partial r_{ff}})}{\partial K} = \frac{a_1 d_1 (1 - c_1) \frac{\partial b_1}{\partial K}}{(b_1 + d_1)^2} > 0$$

First, the net effect of asset size on the sensitivity of  $LN$  to  $r_{ff}$  is undetermined and depends on parameter values. This reflects the idea that two factors play an important role for big banks: maybe they may face less asymmetries of information than smaller banks, and therefore they can obtain easier alternative sources of funds ( $TD$ ) when a contractionary monetary policy reduces  $DD$ ; but, their clients are also bigger firms which tend to be more sensitive to interest rates, so if big banks increase  $r_{LN}$  due to higher costs of funding, they will loose more demand for loans than smaller banks. Similarly, it is unclear the effect of bank size on the sensitivity of  $TD$  to the Fed funds rate.

Furthermore, the sensitivity of  $LN$  to  $r_{ff}$  is lower for better capitalized banks; that is, better capitalized banks experience a lower reduction in loans than

less well capitalized banks do. Similarly, better capitalized banks will increase more TD in times of monetary policy tightening.

Thus, capitalization and bank size appear to matter for lending. In Section 4 we test the predictions for bank lending that derive from this model for Argentina and Colombia.

### 3. DATA

#### 3.1 Data sources

The data for banks of Argentina are taken from the *Información de Entidades Financieras* releases of the Central Bank of Argentina, which consists of the monthly balance sheets that each bank is required to report to the *Superintendencia de Entidades Financieras y Cambiarias* (the financial institutions' regulator). The period used is 2003:08 to 2005:11, that is after the 2001-2002 crisis. Unfortunately, such releases only include those banks that existed as of November 2005; therefore, banks that disappeared throughout the period are not included in the dataset.

The data for banks of Colombia comes from financial statements that banks report monthly to the *Superintendencia Financiera de Colombia*, which regulates the financial system. The sample used covers the period 1995:1 to 2005:9, and consists of all the commercial banks that were operating at every moment of time. The panel of banks is unbalanced, with a maximum number of 40 banks in 1997:12, and a minimum of 21 banks at the final period. The number of banks at the beginning of the sample period was 38. The reduction in the number of banks during the last ten years reflects the consequences of the period of stress experienced by financial institutions in Colombia between 1997 and 1999, which led to bank failures, acquisitions and merges among financial institutions<sup>6</sup>.

The macroeconomic variables used were taken from several sources: International Financial Statistics data was used for CPI and bilateral exchange rate peso-US dollar for both countries. As a proxy of GDP, which is not available on a monthly basis, the *Estimador Mensual de Actividad Económica Index* (EMAE) without seasonality was used for Argentina, and the *Índice de Producción Real Manufacturera* (IPM) for Colombia. The first one is available online in a monthly frequency from *Instituto Nacional de Estadísticas y Censos of Argentina* (INDEC), while the second one is available online at the Central Bank of Colombia's (BANREP) web page.

<sup>6</sup> The way in which the *Superintendencia Financiera* handles acquisitions and merges is the following: when a big bank acquires a small bank (definition), the small bank disappears while the big bank's figures become those of the group; when a merge takes place, both banks disappear at the time of the merge, and a new one starts operating the period after.



The interbank interest rate is used as the monetary policy instrument. This variable was chosen, because it is the most commonly used in monetary transmission mechanism and inflation targeting studies (see, for example, Gómez and Julio, 2001). The interbank interest rate of Argentina is available online at the webpage of the Central Bank of Argentina (BCRA), and that of Colombia is available online at the webpage of BANREP.

### 3.2 Construction of variables

As a proxy of bank-specific characteristics, capitalization ratio and liquidity were used. All these variables were constructed from balance sheet data. The ratio of equity over assets was used as our measure of capitalization. Liquidity was measured as the ratio between liquid assets minus short-term liabilities *and* deposits (checking, savings and time deposits).

These variables are included to test whether bank specific variables affect lending, and whether banks with different characteristics respond differently to monetary policy shocks. According to the bank lending channel, capital market imperfections affect the relationship between banks and their investors. Therefore, it is expected that variables that proxy for the financial health of a bank, such as capitalization and liquidity, matter in these relationships.

### 3.3 Characteristics of the banking sectors in Argentina and Colombia

Tables 1 to 3 show summary statistics of characteristics of the banking systems in Argentina and Colombia at different points in time. These characteristics are presented in two ways: the first four columns in each table group banks according to asset size, while the last four group them according to the capitalization ratio. For instance, the fourth column “>75”, stands for the group composed by the 25 percent of the largest banks in terms of assets. Similarly, the eighth column, also labeled “>75”, stands for the group composed by the 25 percent banks with largest capitalization ratio.

According to Table 1, the best capitalized banks have significantly lower liquidity ratios. This can be explained by the fact that holding liquid assets has an opportunity cost which well capitalized banks need not incur.

It is also important to note that there are differences in the ratios of capitalization among groups of banks in Argentina. For instance, for November 2005, while the 25 percent least capitalized banks had a capital ratio of 5.6%, the top 25 percent had a capital ratio of almost 55%<sup>7</sup>. These differences are much

<sup>7</sup> There is a caveat regarding the way capitalization is measured. Since it is calculated as equity over assets, and after the crisis of 2001-2002 the Central Bank of Argentina allowed the banks to choose the way in which they would account for the losses, some banks may have lower capitalization ratios even though they decided to account for all the losses early in 2002 or 2003, while others still have not accounted for all the losses due to the crises and default. This problem does not have a solution since the Central Bank of Argentina does not make this information public.

TABLE 1  
CHARACTERISTICS OF THE BANKING SYSTEM IN ARGENTINA  
(NOVEMBER 2005)

Characteristics	By Asset Size (percentile)				By Capitalization (percentile)			
	< 25	25-50	50-75	> 75	< 25	25-50	50-75	> 75
<i>Market Share (percent)</i>								
Total Assets	0.65	3.44	9.60	86.31	28.15	54.33	15.74	1.78
Total Loans	0.74	3.14	8.44	87.68	24.34	58.91	14.60	2.15
Securities	0.51	2.80	10.15	86.53	33.69	51.71	12.98	1.62
Deposits	0.21	3.16	9.20	87.44	24.08	64.03	11.36	0.52
<i>Ratios as group aggregate (percent)</i>								
Loans to Asset	41.16	33.15	31.88	36.86	31.36	39.34	33.66	43.89
Securities to Assets	23.29	24.06	31.18	29.58	35.30	28.08	24.33	26.91
Deposits to Liabilities	38.38	70.44	72.37	69.34	55.52	80.61	57.33	42.89
Capitalization	49.25	16.17	14.36	10.79	5.59	11.26	18.45	54.88
Liquidity	11.80	17.70	26.60	28.47	32.99	24.72	30.57	11.53
<i>In millions of pesos of November 2005</i>				<i>In millions of US dollars of November 2005</i>				
Total Assets of the system	208159			Total Assets of the system	70658			
Total Loans of the system	75522			Total Loans of the system	25636			
Total Securities of the system	61406			Total Securities of the system	20844			
Total Deposits of the system	133317			Total Deposits of the system	45254			
Number of Banks	72							

Source: Author's elaboration based on "Información de Entidades Financieras", Banco Central, Argentina.

stronger than those of banks in the US and other countries such as Colombia. It can also be observed that the group of more capitalized banks has a significantly lower ratio of deposits to liabilities. This can be explained by the fact that more capitalized banks have better access to alternative sources of funding, and therefore they depend less on deposits to finance their positions in assets.

As for Colombia, some important facts can be observed from Table 2. When looking at the differences according to size, it can be noticed that smaller banks tend to have a bigger participation of loans in their assets, vis-a-vis larger banks. Similarly, they have a smaller participation of securities. That trend explains the fact that smaller banks appear to have lower liquidity ratios than larger banks do. Similarly, the biggest banks have a high capitalization ratio, relative to the mean of the system. Equally as important, the banks in the smallest percentile are the less capitalized among commercial banks. Note also that banks of any size obtain their financing basically from deposits<sup>8</sup>.

<sup>8</sup> Deposits include demand deposits as well as time deposits.

TABLE 2  
CHARACTERISTICS OF THE BANKING SYSTEM IN COLOMBIA  
(SEPTEMBER 2005)

Characteristic	By Size (percentile)				By Capitalization (percentile)			
	< 25	25-50	50-75	> 75	< 25	25-50	50-75	> 75
<i>Market Share (percent)</i>								
Total Assets	7.9	15.4	25.8	50.9	23.1	19.8	10.7	46.4
Total Loans	9.5	17.6	24.3	48.7	22.4	19.9	12.3	45.4
Securities	5.9	11.7	29.5	52.9	24.2	19.9	8.0	47.9
Deposits	7.5	15.9	26.7	49.9	23.1	20.4	11.0	45.5
<i>Ratios (percent)</i>								
Loans to Assets	62.6	59.2	49.0	49.8	50.4	52.3	59.9	50.9
Securities to Assets	25.3	25.6	38.7	35.2	35.5	34.0	25.3	34.9
Deposits to Liabilities	75.8	85.9	84.6	81.4	79.0	84.2	84.5	83.1
Liquidity	15.9	24.8	41.4	37.8	36.2	36.1	24.1	36.8
Capitalization	8.8	12.7	10.8	12.1	7.9	10.7	11.9	13.8
<i>Participation of Total Loans (percent)</i>								
Commercial	52.4	51.1	60.5	70.7	56.5	57.8	52.4	71.5
Consumer	46.8	25.6	30.6	17.5	27.6	33.8	26.6	19.1
Mortgage	0.4	24.3	8.7	10.0	9.6	9.4	22.1	10.0
<i>In millions of Colombian pesos of November 2005</i>				<i>In millions of US dollars of November 2005</i>				
Total Assets of the system	110,231,651			Total Assets of the system				47,690
Total Loans of the system	57,357,481			Total Loans of the system				24,815
Total Securities of the system	37,273,492			Total Securities of the system				16,126
Total Deposits of the system	80,323,813			Total Deposits of the system				34,721
Total Number of Commercial Banks	21							

Source: Author's elaboration based on monthly report of Superintendencia Financiera de Colombia.

Meanwhile, when looking at the characteristics of banks according to their capitalization ratio, trends seem to be less clear. The most capitalized banks (those above the 75th percentile) are also the biggest ones, according to market shares. But, for lower percentiles, there is no clear relationship between size, as proxied by market shares, and the degree of capitalization. More importantly, when dividing banks according to the capitalization ratio, there does not appear to be a pattern in asset composition; the percentage of loans to total assets are very similar among groups, as well as the percentage of securities. But, what can be observed is that the most capitalized banks have a larger proportion of commercial loans and a smaller participation of consumption loans.

Nevertheless, banks' characteristics in Colombia appear to have changed over time. Table 3 replicates the information shown in Table 3 for December

1996. By the end of 1996, there was a negative relation between capitalization and the ratio of loans to assets; while banks in the first quartile according to capitalization had a ratio of loans to assets above 75%, the quartile of the most capitalized banks had a ratio below 60%. Correspondingly, the ratio of securities to assets had a positive relationship with the degree of capitalization. When comparing asset composition between 1996 and 2005, it is noticeable that, in general, banks turned to have more securities and fewer loans in their portfolios after the crisis. Note that the ratio of loans to assets reduced importantly for all groups of banks, while the ratio of securities to assets rose significantly. This can probably be explained by the systematic default of banks that took higher risks in the period of credit boom before the crisis, and by more conservative lending

TABLE 3  
CHARACTERISTICS OF THE BANKING SYSTEM IN COLOMBIA  
(DECEMBER 1996)

Characteristic	By Size (percentile)				By Capitalization (percentile)				
	< 25	25-50	50-75	> 75	< 25	25-50	50-75	> 75	
<i>Market Share (percent)</i>									
Total Assets	4.6	11.8	28.8	54.8	31.8	15.5	25.2	27.4	
Total Loans	4.7	11.2	29.6	54.5	36.8	16.5	22.7	24.0	
Securities	4.2	12.1	25.7	58.0	15.3	11.1	29.6	44.0	
Deposits	4.0	11.6	31.1	53.3	38.7	16.8	22.8	21.8	
<i>Ratios (percent)</i>									
Loans to Assets	66.4	62.2	67.4	65.2	75.7	69.8	59.1	57.2	
Securities to Assets	10.4	11.9	10.3	12.1	5.5	8.3	13.5	18.5	
Deposits to Liabilities	69.4	78.9	81.9	82.8	93.1	86.3	73.6	70.5	
Liquidity	8.8	9.4	9.7	10.7	4.0	7.1	11.7	19.5	
Capitalization	14.8	13.6	12.2	13.6	5.9	10.5	15.0	21.9	
<i>Participation of Total Loans (percent)</i>									
Commercial	82.5	70.5	40.9	49.7	26.5	44.8	71.6	73.0	
Consumer	17.5	28.9	19.7	14.6	5.7	22.4	28.1	23.7	
Mortgage	0.0	0.6	39.4	35.7	67.8	32.8	0.3	3.3	
<i>In millions of Colombian pesos of November 2005</i>				<i>In millions of US dollars of November 2005</i>					
Total Assets of the system	94,962,893			Total Assets of the system					53,207
Total Loans of the system	62,216,689			Total Loans of the system					34,859
Total Securities of the system	10,956,440			Total Securities of the system					6,139
Total Deposits of the system	65,782,971			Total Deposits of the system					36,858
Total Number of Commercial Banks	40								

Source: Author's elaboration based on monthly report of Superintendencia Financiera de Colombia.

policies taken by surviving banks that tend to account better for risks coming from the lending business.

Another important feature, now regarding liability composition, is that more capitalized banks in 1997 tended to have a lower ratio of deposits to total liabilities than less capitalized banks. This seems to provide some indirect evidence of the presence of capital market imperfections for banks, as predicted by the bank lending channel theory, and also of the importance of the capitalization ratio as an indicator of the degree of informational asymmetries faced by banks. Basically, the fact that there is a negative relation between the degree of capitalization and the ratio of deposits to liabilities supports the idea that better capitalized banks are less deposit-dependent, in the sense that they can find easier substitutes for these than less capitalized banks do. Nevertheless, this relationship tended to disappear in time, becoming less clear in recent years. This could have happened because the differences in capitalization between different groups of banks has reduced, due to failure of poorly capitalized banks and also to financial decisions taken by surviving banks.

As a final point concerning the evolution of the banking sector characteristics in time for Colombia, it is important to mention that more capitalized banks appear to have a larger proportion of commercial loans in their portfolios. This could suggest that banks that lend to large firms are healthier and better capitalized banks.

One can observe several differences for Argentina and Colombia from Tables 1 to 3. One is that the banking system of Argentina is much more concentrated than that of Colombia. For example, while the largest banks in Argentina hold 86% of the assets of the system, that figure is 51% for Colombia, even though the number of commercial banks in Argentina is more than three times that of Colombia. Next, the capitalization ratios for Argentina are much higher than those for Colombia, especially in recent years. Also, banks in Argentina with the highest capitalization ratios are small banks in terms of asset size, whereas in Colombia, those with the highest capitalization ratios are big banks in terms of asset size.

### 3.4 Evidence of a Bank Lending Channel in other Countries

Kashyap and Stein (1995) use quarterly data on banks operating in the United States from 1976.1 to 1992.2. They classify banks by their asset size and use the Fed Funds rate as the monetary policy instrument. They find that a tightening in monetary policy declines deposits for all size categories and in a similar fashion. They find evidence of a bank lending channel for all but the biggest asset size category –that is, increases in the Fed funds rate has a negative and statistically significant effect on the growth rate of total loans. If instead of total loans, C&I loans are used, they find similar results. They also find that small bank securities holdings are more sensitive to changes in monetary policy, although they recognize that the results they get are not as clear as with growth rate of loans.

Kishan and Opiela (2000) also provide evidence that support the existence of a bank lending channel in the United States. They use quarterly data from 1980.1 to 1995.4 for 13042 commercial banks and divide banks according to their size and capital adequacy. They find that loans of small undercapitalized banks are more responsive to monetary policy. They also find that large time deposits of small undercapitalized banks are unresponsive to monetary policy. They conclude that the asset size and the capital ratio have important implications for the strength and distributional effects of monetary policy and for the relation between stabilization and regulatory policy.

Altunbas *et al.* (2002) use yearly balance sheet data to estimate the response of monetary policy stance between 1991 and 1999 for countries in the European Monetary Union (EMU). They classify banks according to asset size and capital and find that undercapitalized banks of any size tend to respond more to changes in monetary policy when they use the pooled data on the eleven EMU countries. When they look at individual country estimates for France and Germany they do not find evidence supporting a bank lending channel, but they do find evidence for Italy and Spain. In Spain, they find some evidence that a bank lending channel exists for the largest undercapitalized and adequately capitalized banks. In Italy, they find some evidence of a bank lending channel for adequately capitalized and overcapitalized banks. They conclude that more research is needed to further investigate bank lending channels in the smaller EMU countries and that balance sheet data availability should be improved and different model specifications should be used –i.e., more indicators of monetary policy stance and different lag structures.

Alfaro *et al.* (2005) use quarterly data from 1990.1 to 2002.2 and find evidence supporting a bank lending channel in Chile with a significant impact on macroeconomic activity. The evidence of a bank lending channel is more important for banks with low liquidity ratios and to lesser extent smaller and less capitalized banks.

#### 4. EMPIRICAL RESULTS

The empirical specification in this panel approach is the following:

$$(15) \quad y_{it} = \sum_{j=1}^6 X_{it-j} \beta_j + \sum_{j=1}^6 x_{3it-j} \lambda_j + Z_{it-1} \phi_j + \alpha \text{Dummycrisis}_t + u_{it}$$

where  $y_{it}$  represents the growth rate of total loans for bank  $i$  at time  $t$ ;  $X$  is a matrix of macroeconomic variables including a proxy for the growth rate of GDP, growth rate of real exchange rate and the policy instrument, which is the real interbank interest rate (denoted by  $x_{3it}$  and calculated as the nominal interbank interest rate minus observed monthly inflation rate).  $Z$  is a matrix of bank specific variables, namely, capitalization and liquidity. Monthly dummies

to control for seasonality in the data were also included; *Dummyscrisis* is a dummy variable included for Colombia to control for the financial crisis period between July 1998 and December 2000.  $1_2'$  is a  $2 \times 1$  row vector of ones and “ $\odot$ ” is the Hadamard product. The error term was assumed *i.i.d.* as well as to account for bank specific AR(1) structure. For Colombia, another two regressions are considered, namely, one for commercial loans and one for consumer loans.

The estimated empirical specification for both countries is

$$(16) \quad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} X_1^* \\ X_2^* \\ \vdots \\ X_m^* \end{pmatrix} \delta + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix}$$

where 1, ...,m are the banks and the  $X^*$  matrix contains all the variables in the Right Hand Side of (15)

$$E(\mathbf{u} \mathbf{u}') = \begin{pmatrix} \sigma_{1,1} \Omega_{1,1} & \sigma_{1,2} \Omega_{1,2} & \cdots & \sigma_{1,m} \Omega_{1,m} \\ \sigma_{2,1} \Omega_{2,1} & \sigma_{2,2} \Omega_{2,2} & \cdots & \sigma_{2,m} \Omega_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m,1} \Omega_{m,1} & \sigma_{m,2} \Omega_{m,2} & \cdots & \sigma_{m,m} \Omega_{m,m} \end{pmatrix}$$

As a robustness test, the variance structure of the errors was also specified to account for autocorrelation of order 1 specific to each panel<sup>9</sup>.

The results for Argentina suggest that there is no direct bank lending channel triggered by changes in the interbank interest rate. But there is indirect effect on bank lending through the interactions of the interbank interest rate and capitalization and, to a less extent, liquidity.

These results suggest that when the interbank interest rate changes, the growth rate of loans will be affected less the better capitalized (and the higher its liquidity ratio) the bank is.

<sup>9</sup> Some studies have used dynamic panel data models to look for evidence of a bank lending channel. However, since the endogenous variable is the growth rate of loans, it is not clear why the growth rate of loans of today depending on its previous realizations. The current period growth rate of loans might depend on past periods realizations through demand side influences, but these should be captured by the macroeconomic and bank specific variables.

TABLE 4  
TOTAL EFFECT COEFFICIENTS FOR ARGENTINA  
DEPENDENT VARIABLE: GROWTH RATE OF TOTAL LOANS

<i>Variable</i>	<i>Coefficient</i>	<i>Std.Err</i>
Interbank rate	-0.032	0.022
Interactions:		
Interbank rate x Liquidity	0.048**	0.015
Interbank rate x Capitalization	0.143*	0.025

\*, \*\* Significant at the 1 and 10 percent level.

The results for Colombia agree with the basic ideas of the bank lending channel. See Tables 5 and 6, which show the results of the regressions corresponding to total loans and commercial loans. Note that the impact of monetary policy on the growth rate of loans is negative; increments in the interbank interest rate lead to reductions in the growth rate of loans. However, the impact is not the same across banks; those institutions that have lower capitalization and liquidity ratios are affected more.

This suggests that these bank specific variables, capitalization and liquidity, affect lending decisions done by banks, and also the ability that they have to obtain alternative sources of funding when a monetary policy shock affects the amount of core deposits in the economy. The interpretation of the coefficient on the interaction terms is the following: when the interbank rate goes up by one percentage point (*i.e.* from 6 to 7 percent), there is a direct effect on total loans, that according to Table 5 will fall by 0.68 percent. But the more capitalized a bank is, the smaller the fall in deposits will be. For example, total loans a bank with a capitalization of 0.10 will fall by  $0.6824 - 4.542 * 0.10 = 0.228$ , while a bank with capitalization ratio of 0.15, the fall will be of smaller magnitude, in fact negligible:  $0.6824 - 4.542 * 0.15 = 0$ .

The results are quite different when the growth rate of consumption loans is used in the regressions. The interest rate is not significant in explaining changes in the growth rate of these loans, and bank specific variables do not appear to matter either. This result suggests that consumer loans have a different dynamics than commercial loans. However, as it was before seen, commercial loans represent the great majority of total loans, and that can explain that even when there is no evidence of a bank lending channel for consumer loans, there is evidence for total loans.



TABLE 5  
TOTAL EFFECT COEFFICIENTS FOR COLOMBIA  
DEPENDENT VARIABLE: GROWTH RATE OF TOTAL LOANS

<i>Variable</i>	<i>Coefficient</i>	<i>Std.Err</i>
Interbank rate	-.6824*	0.0551
Interactions:		
Interbank rate x Liquidity	1.6949*	0.1379
Interbank rate x Capitalization	4.5424*	0.3608

\* Significant at the 1 percent level.

TABLE 6  
TOTAL EFFECT COEFFICIENTS FOR COLOMBIA  
DEPENDENT VARIABLE: GROWTH RATE OF COMMERCIAL LOANS

<i>Variable</i>	<i>Coefficient</i>	<i>Std.Err</i>
Interbank rate	-0.6710*	0.0757
Interactions:		
Interbank rate x Liquidity	1.5629*	0.1892
Interbank rate x Capitalization	4.7544*	0.4947

\* Significant at the 1 percent level.

## 5. BANK LENDING CHANNEL FOR COLOMBIA USING FIRMS' DATA

As a robustness test for the evidence of bank lending channel for Colombia, we analyze balance sheet data for Colombian firms. In order to identify the bank lending channel with these data, we separate firms into two groups according to the degree of leverage<sup>10</sup>. Group 1 is composed by firms in the lowest quartile of leverage and Group 2 is composed by firms in the highest quartile. The intuition behind this way of grouping firms is that in moments of tight monetary policy, firms with higher indebtedness will find it harder to obtain liquidity. Therefore, the impact of a change in interest rates on these firms should be stronger than on those with lower leverage ratios.

Kashyap *et al.* (1993) developed a simple theoretical framework to model changes in optimal debt structure of the firm over time. In times of contractionary monetary policy, bank loans experience a supply shock due to the fall in deposits.

<sup>10</sup> Leverage is defined as the ratio of total liabilities to total assets of the firm, following the conventional definition.

As the spread between the interest rates on bank loans and non-bank debt rises, the lending channel would anticipate a decrease in bank lending in the firm's optimal debt structure. Furthermore, as risk increases in periods of contractionary monetary policy, banks will be more reluctant to lend to firms with poor information availability.

### 5.1 Data Source

The data for firms was collected by the *Superintendencia de Sociedades*, the organism that regulates non-financial firms in Colombia. The sample consists of a panel of annual observations on firms from 1995 to 2004. For each year, more than six thousand firms submit information about their balance sheets. The panel is unbalanced, due to the entry and exit of firms, and also to the fact that some firms report their balances for some years but not for others (not necessarily for consecutive years).

After eliminating the data base for firms that have an early exit or a late entry, or that do not report for the whole span of time, we are left with a balanced panel of around four thousand non-financial firms belonging to various economic sectors.

### 5.2 Empirical Specification and Results

The empirical specification in this panel approach is the following:

$$(18) \quad BD_{it} = \sum_{j=1}^2 BD_{it-j} \gamma_j + \sum_{j=1}^2 \Delta \log(R_{t-j}) \theta_j + \sum_{j=1}^2 inv_{it-1} \eta_j + cap_{it-1} \mu + \alpha Dummycrisis_t + u_{it}$$

where  $BD_{it}$  represents the ratio of bank debt to total debt for firm  $i$  at time  $t$ ;  $R_t$  is the real interbank interest rate, which instruments for monetary policy;  $inv$  controls for inventories (the decision of increasing or decreasing inventories might affect the debt structure in the margin) and  $cap$  for capitalization. The parameters of interest are  $\gamma$ ,  $\theta$ ,  $\eta$  and  $\mu$ ; in particular, the vector  $\theta$ , which multiplies the instrument for monetary policy.  $Dummycrisis$  is a dummy variable included for Colombia to control for the financial crisis period of 1999 and 2000. The error term is assumed *i.i.d.*

Because lag values of the dependent variable are included as regressors, we use Arellano and Bond (1991) methodology, which corrects the inconsistency of the within estimator and provides a consistent and efficient estimator. The intuition of including lags of the ratio of bank debt as covariates is that the ratio presents some inertia over time, in the sense that the debt structure of firms does not change significantly in short periods of time.

The main results are summarized in Table 7, which presents the long run coefficients of the real interest rate for both groups of firms.

As shown in Table 7, the ratio of bank debt to total debt of both groups of firms goes down when the real interest rate increases. This is due to both supply

TABLE 7  
EVIDENCE OF BANK LENDING CHANNEL FOR FINNS  
DEPENDENT VARIABLE: RATIO OF BANK DEBT TO TOTAL DEBT

Group 1: Low-Leveraged Firms

<i>Variable</i>	<i>Coefficient</i>	<i>Std.Err.</i>
Interbank rate	-0.0068*	0.0037

\* Significant at the 10 percent level.

Group 2: High-Leveraged Finns

<i>Variable</i>	<i>Coefficient</i>	<i>Std.Err.</i>
Interbank rate	-0.0191*	0.0068

\* Significant at the 1 percent level.

and demand factors and debt substitution. However, the impact is clearly bigger and more significant for high leveraged firms than for low leveraged firms. This provides some additional support for the existence of a bank lending channel of monetary policy in Colombia.

## 6. CONCLUSION

In this paper we find empirical evidence of bank lending channel for Colombia. As for Argentina, we do not find evidence that changes in the interbank interest rate affect the growth rate of total loans directly. It does, however, indirectly through interactions: the interbank interest rate affects the loan supply through its interactions with capitalization and liquidity. As for Colombia, there is direct bank lending channel, which is reinforced through interactions with capitalization and liquidity.

Bank specific variables are key in analyzing how changes in the interbank rate affect the growth rate of loans for Argentina and Colombia. It is particularly important the way in which capitalization affects the lending ability of banks; better capitalized banks should be able to lend more in moments of constrained liquidity. This might be the result of a combination of minimum capital regulations and informational frictions in markets for banks' funds, which affect both the decisions taken by banks and their ability to raise funds in markets alternative to the deposits one.

Finally, using a panel data of more than 3300 firms, we provide additional support to the existence of a bank lending channel for Colombia.

## REFERENCES

- Alfaro, R.; H. Franjen; C. García y A. Jara (2005), "The Bank Lending Channel in Chile". BIS Papers No 22, part 8, April 2005.
- Altunba, Y.; O. Fazylov y M. Philip (2002), "Evidence on the bank lending channel in Europe". *Journal of Banking & Finance* 26: 2093-2110.
- Arellano, M. and S. Bond (1991), "Some Tests of Specification for Panel Data: Monte Carlo evidence and an application to employment equations". *Review of Economic Studies* 58: 277-297.
- Bernanke, B.S. and A. S. Blinder (1988), "Credit, Money and Aggregate Demand". *American Economic Review*, 78(2): 435-9.
- Bernanke, B.S. and A. S. Blinder, (1992), "The Federal Funds Rate and the Channels of Monetary Transmission". *American Economic Review*, 82: 901-21.
- Bernanke, B.S. y M. Gertler (1995), "Agency Costs, Net Worth and Business Fluctuations". *American Economic Review*, 79: 14-31.
- Bernanke, B.S. and M. Gertler (1995), "Inside the Black Box: The Credit Channel of Monetary Policy Transmission". *Journal of Economic Perspectives*, 9: 27-48.
- Fazzari, S.; G. Hubbard y B. Petersen (1988), "Financing Constraints and Corporate Investment". *Brookings Papers on Economic Activity*, 141-95.
- Gómez, J. y J.M. Julio (2001), "Transmission mechanisms and inflation targeting: the case of Colombia's disinflation". *Borradores de Economía* 168. Banco de la República.
- Huang, Z. (2003), "Evidence of a Bank Lending channel in the UK". *Journal of Banking and Finance* 27: 491-510.
- Kashyap, A.K. and Stein, J.C. (1995), "The impact of Monetary Policy on Bank Balance Sheets". *Carnegie-Rochester Conference Series on Public Policy* 42, 151-95.
- Kashyap, A.K.; Stein, J.C. and W. Wilcox (1993), "Monetary Policy and Credit Conditions: Evidence from the Composition of External Finance". *American Economic Review* 83: 78-98.
- Kishan, R.P. and T. P. Opiela (2000), "Bank Size, Bank Capital and the Bank Lending Channel". *Journal of Money, Credit and Banking* 32: 121-41.
- Modigliani, F. and M. Miller (1958), "The Cost of Capital, Corporate Finance and the Theory of Investment". *American Economic Review* 48: 261-97.