



BANCO DE LA REPUBLICA
Gerencia Técnica
Unidad de Estudios Económicos

Determinantes de la elección de administradora de pensiones: primeras estimaciones a partir de agregados

Luis Eduardo Arango y Luis Fernando Melo *

larangth@banrep.gov.co, lmelovel@banrep.gov.co

Resumen

Utilizando información del sistema de ahorro individual entre 1998 y 2005, se encuentra evidencia de que la tasa de retorno real de los fondos y la población ocupada son los determinantes principales del número de cotizantes a las distintas AFP. El valor promedio del fondo, utilizado como proxy de otras variables como la capacidad que tiene cada administradora de difundir su producto y de contactar potenciales afiliados, no resultó significativo. Se utiliza la técnica de cointegración panel de Groen y Kleinbergen (2003).

Clasificación JEL: G23, C23.

Palabras clave: Administradora de Fondos de Pensiones (AFP), rentabilidad, cointegración panel.

* Las opiniones contenidas en este documento son responsabilidad exclusiva de sus autores y no coinciden necesariamente con las del Banco de la República ni su Junta Directiva. Se agradece la asistencia de Oscar Becerra, John Jairo León y Carlos Sandoval así como los comentarios y sugerencias de Yaneth Rocío Betancourt, Luz Adriana Flórez, Diego Jara, Gabriel Piraquive y Carlos Esteban Posada.

1. Introducción

La reforma laboral promulgada mediante la Ley 100 de 1993 dio lugar a la existencia de un sistema dual de administración del ahorro obligatorio, conocido como el Sistema General de Pensiones, que constituye la base para la pensión a la edad del retiro. Desde entonces, además del Régimen Solidario de Prima Media con Prestación Definida (*pay-as-you-go*) cuyo eje es el Instituto de Seguro Social, está el Régimen de Ahorro Individual con Solidaridad (*fully-funded*) del cual se encargan las seis administradoras de los fondos de pensiones (*AFP*) que hoy operan, después de las fusiones y adquisiciones que se han registrado. Los recursos de las cuentas individuales conforman un portafolio de inversiones administrado por las *AFP* siguiendo el modelo chileno¹.

Según las normas actuales, la afiliación al Sistema General de Pensiones es obligatoria para todos los trabajadores dependientes e independientes. Antes de la vigencia de la Ley 797 de 2003, la afiliación era obligatoria solo para los asalariados, vinculados mediante un contrato de trabajo, y los servidores públicos mientras que para los trabajadores independientes o cuenta-propia dicha afiliación era voluntaria². Al momento de la creación del Sistema General de Pensiones, los trabajadores debieron escoger, de manera no irreversible para muchos de ellos, entre permanecer en el Régimen de prima media o cambiarse al nuevo sistema.

Actualmente, las *AFP* que conforman el Régimen de Ahorro Individual con Solidaridad³ son Colfondos, Horizonte, Porvenir, Protección, Santander y Skandia, las cuales cuentan con cerca de seis millones de afiliados (aportantes y no aportantes), de los cuales tres millones se encuentran activos⁴ (aportantes), y administran un ahorro conjunto, incluyendo sus rendimientos, de \$ 28.5 billones, lo que significa un valor aproximado de \$4.8 millones por afiliado a abril de 2005. A comienzos del año 2000 los afiliados constituían el 50.1% y los activos el 25.5% de la población ocupada (13 principales ciudades) mientras que a abril de 2005 los afiliados eran algo

¹ Para una descripción del sistema chileno véase Arenas de Mesa y otros (2004). El sistema colombiano de pensiones también ha sido analizado ampliamente. Véanse, por ejemplo, Clavijo (2002), Echeverry et al. (2001) y Bonilla (2001). Más recientemente, Jara et al. (2006) estudian la eficiencia de los portafolios de las *AFP*. Sobre las estructuras de los sistemas de pensiones en general y sus implicaciones véase Barr (2000).

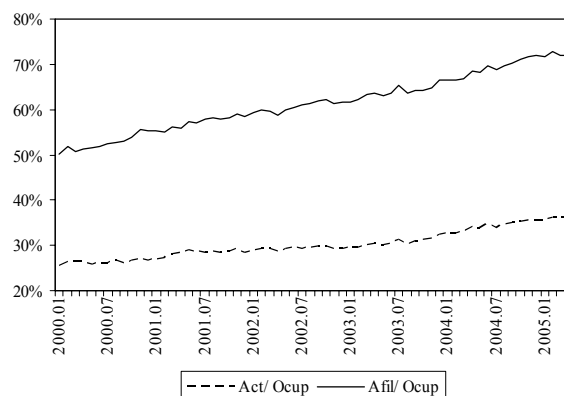
² Llama la atención que, pese al cambio introducido en 2003 por la Ley 797, no se observa ningún cambio importante en el número de afiliados para esa época.

³ Además de las entidades, el régimen lo conforman normas y procedimientos.

⁴ Los afiliados que no realizan ninguna cotización durante seis meses se consideran afiliados inactivos. El desempleo, la informalidad y los cambios en la contratación laboral son las principales razones para la existencia de inactivos (El Tiempo, Septiembre 28 de 2005). A mediados de 2004 el Instituto de Seguros Sociales tenía 2 millones afiliados activos aproximadamente.

menos de 72% y los activos el 36% de la población ocupada. Esto significa que el avance mensual en materia de cobertura de afiliados activos ha sido de 0.17 puntos porcentuales. La Figura 1 presenta la evolución de los afiliados y los afiliados activos al sistema de ahorro individual en relación con la población ocupada (13 ciudades).

Figura 1. Afiliados y activos en relación con ocupados (13 ciudades)



Fuente: Superbancaria – DANE.

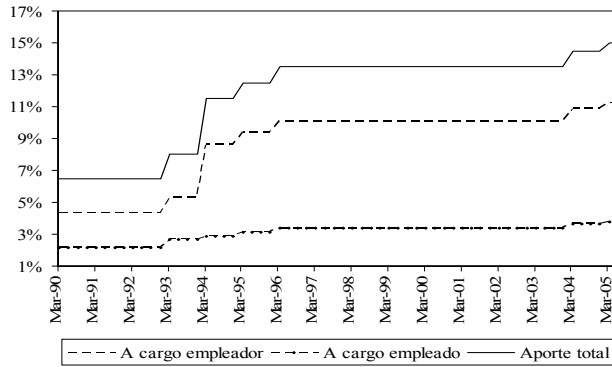
De acuerdo con las disposiciones vigentes, a las *AFP* llega un aporte del 15% del salario mensual al cual contribuye el empleador en un 75% y el empleado en un 25% (Figura 2). El artículo 7° de la Ley 797 de 2003, que modifica el artículo 20 de la Ley 100, establece que, a la fecha (diciembre de 2005), el 15% del ingreso base de cotización se destinará al Sistema y será distribuido de la siguiente manera: 10.5% a las cuentas individuales de ahorro pensional, 1.5% al Fondo de Garantía de Pensión Mínima del Régimen de Ahorro Individual con Solidaridad y el 3% restante a financiar los gastos de administración, la prima de reaseguros de Fogafin y las primas de los seguros de invalidez y sobrevivientes. A partir de enero de 2006 la contribución total será de 15.5%, destinándose los 0.5 puntos porcentuales (*pp*) adicionales a las cuentas de ahorro individuales⁵.

El esquema de ahorro individual asegura una cotización, un rendimiento mínimo, inducido por la regulación que ejerce la Superintendencia Bancaria, y, como veremos más adelante, una pensión mínima para quienes cumplan ciertas condiciones. Sin embargo, por

⁵ Según la legislación actual, el gobierno nacional podrá incrementar el aporte total a 16.5% a partir del año 2008 siempre y cuando el crecimiento del *PIB* sea igual o superior al 4%, en promedio, durante 2006 y 2007.

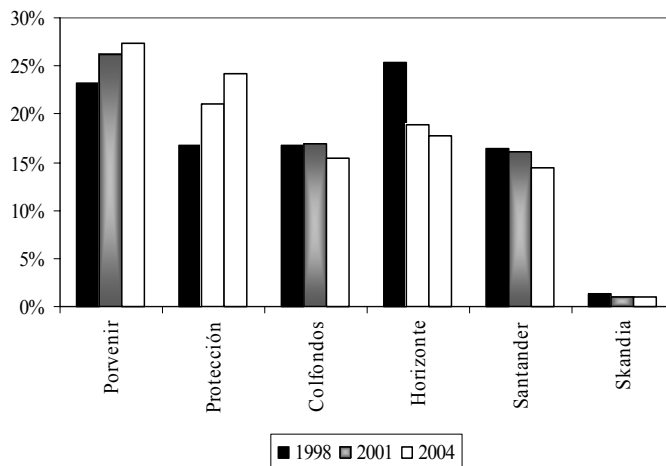
encima de esta pensión mínima existe incertidumbre sobre los beneficios que obtendrá cada cuenta de ahorro en exceso sobre ese retorno mínimo.

Figura 2. Aporte mensual por concepto de pensión



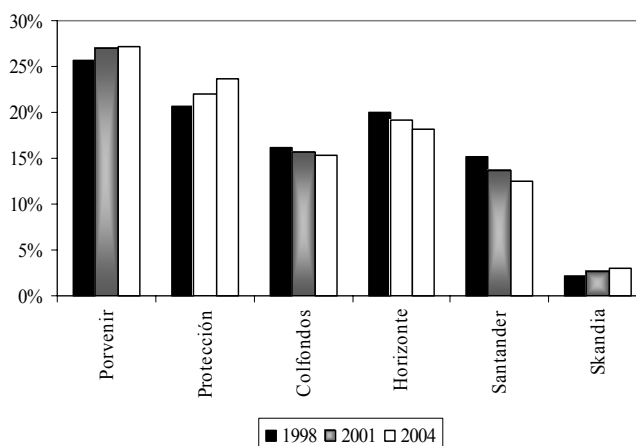
En la Figura 3 se observa la participación de cada *AFP* en los afiliados activos entre 1998 y 2004. Se presentan allí variaciones importantes. Por ejemplo, se destaca la manera cómo Porvenir y Protección han ganado presencia en este aspecto, mostrando la segunda de ellas una mejor dinámica, mientras que Colfondos, Santander y Horizonte han perdido espacio, en especial esta última. De igual manera, los fondos administrados por Porvenir y Protección han ganado participación en el valor total del ahorro pensional mientras que los demás apenas si se han mantenido cuando no han disminuido su participación (Figura 4).

Figura 3. Composición porcentual de los cotizantes entre 1998 y 2004



Fuente: Superintendencia Bancaria.

Figura 4. Participación promedio en valor total del sistema AFP



Fuente: Superintendencia Bancaria

De las Figuras 3 y 4 surgen algunas preguntas como por ejemplo: ¿cuál es el determinante del cambio en la participación que tienen las AFP en el total de cotizantes?. ¿Es la rentabilidad real que ofrecen las administradoras la variable fundamental o hay otras explicaciones para los cambios en la participación?. La teoría económica señala que la tasa de retorno que ofrecen los activos es fundamental al momento de decidir cuánto invertir. También señala la teoría que cuando los agentes tienen una alta aversión no están dispuestos a efectuar cambios muy importantes en su trayectoria de consumo debido a los movimientos tan pronunciados que se presentan en su utilidad marginal. Sin embargo, este punto es de menor importancia en nuestro análisis ya que, dada la contribución forzosa que significa cotizar para la pensión, los individuos no están sacrificando más que el costo de obtener información sobre el retorno que ofrecen los fondos para obtener a cambio un mayor consumo futuro. Es decir, dada esa contribución forzosa, no dejan de consumir nada hoy para consumir más mañana. Por eso, el análisis estático es suficiente para entender el problema del individuo como se verá en la sección 3.

Al verificar la hipótesis se debe tener en cuenta que el ahorro pensional tiene ciertas particularidades. En primer lugar, se trata de un ahorro forzoso, que establecen las normas tanto para trabajadores dependientes como independientes, de manera que al final del período productivo tengan la posibilidad de preservar una corriente de consumo que no diste mucho de la que tenían durante su vida laboral.

La magnitud del ahorro no es, pues, una decisión autónoma, endógena ni óptima de los agentes⁶. Se trata de una obligación determinada por las autoridades económicas que afecta el ingreso disponible de los trabajadores. La primera instancia de la única decisión autónoma de estos se produce al momento de escoger el régimen al cual desean vincularse: al de ahorro individual o al de prima media. Si la decisión se tomó en favor del régimen de cuentas individuales, la segunda instancia de la decisión está en determinar la *AFP* que va administrar sus recursos de ahorro. El traslado de una *AFP* a otra se puede hacer después de permanecer en alguna de ellas un mínimo de seis meses.

En segundo lugar, existe la garantía de una pensión mínima que hace que la gente de salarios bajos se preocupe en menor medida por el retorno que reportan las *AFP*. En este sentido, el artículo 65 de la Ley 797 de 2003 señala que los afiliados que a los 62 años de edad, si se trata de hombres, y a los 57, si se trata de mujeres, que no hayan alcanzado a generar la pensión mínima de que trata el artículo 35 de la Ley 100 de 1993, y hubiesen cotizado por lo menos 1.150 semanas tendrán derecho a que el Fondo de Garantía de Pensión Mínima del Régimen de Ahorro Individual con Solidaridad, en desarrollo del principio de solidaridad, les complete la parte que haga falta para obtener dicha pensión. A propósito, el artículo 35 de la Ley 100 señala que el monto mensual de la pensión mínima de vejez o jubilación no podrá ser inferior al valor del salario mínimo legal mensual vigente.

El alcance de nuestro trabajo es modesto en el sentido de que solamente se concentra en la segunda instancia de la decisión. La muestra que utilizamos considera solamente el sector privado del mercado de pensiones; es decir, los afiliados activos a las *AFP*. Una vez en este sistema nos preguntamos ¿qué determina la vinculación a una *AFP* u otra?. Utilizando técnicas de cointegración panel este trabajo verifica la hipótesis de que, entre enero de 1998 y abril de 2005, el retorno ofrecido por las administradoras constituye un determinante de la composición de los cotizantes a una *AFP* específica, una vez se controla por la población ocupada, la cual opera como variable de escala.

El trabajo se desarrolla de la siguiente manera. La sección 1 es esta introducción; la sección 2 muestra algunos hechos distintivos del sector; la sección 3 esquematiza la decisión de

⁶ Lo sería si los agentes decidieran cada período la proporción de su ingreso corriente (transitorio y permanente) que desean ahorrar. Esto no quiere decir que la existencia de un ahorro obligatorio para la vejez sea una política equivocada; simplemente, que tiene ciertas características de unilateralidad por parte del “planificador central” quien no tiene en cuenta la tasa de impaciencia de los agentes, su grado de aversión al riesgo, su pesimismo con respecto al futuro, ni otras características de sus gustos y preferencias.

escogencia de la *AFP* por parte de los agentes económicos; la sección 4 describe el método de estimación; la sección 5 presenta y discute los resultados; y, finalmente, la sección 6 elabora algunas conclusiones.

2. Algunas características del sistema de cuentas individuales

Recientemente Colombia se ha identificado por sus bajas tasas de ahorro. Así, por ejemplo, mientras que en 1991 dicha tasa era superior a 22% del *PIB*, al comienzo de la presente década, de acuerdo con la información de cuentas nacionales, se ubicó en 12%⁷. La reducción de las restricciones de liquidez, la inflación de activos y las mayores tasas impositivas han figurado en la lista de explicaciones a este fenómeno (Urrutia, 1995)⁸.

El ahorro pensional en las *AFP*, por su parte, ha venido incrementándose como proporción del *PIB*; pasó de significar cerca de 6% de este en 1998 a representar algo más de 43% en 2005 en frecuencia trimestral. Sobre esta base puede decirse que las pensiones del Régimen de Ahorro Individual, al cierre de 2004, eran ligeramente superiores al 10% del *PIB* anual. Sin embargo, si se tiene en cuenta que el ahorro pensional es una variable *stock* y el *PIB* una variable flujo, esta relación podría ser engañosa. La Figura 5 muestra cómo la relación entre el ahorro en cuentas individuales y el *PIB* acumulado entre el tercer trimestre de 1994 (cuando se dio inicio al nuevo régimen) y el primer trimestre de 2005 ha venido en aumento ya que pasó de significar 0.41% del *PIB* acumulado hasta 1998:3 a significar 1.08% del acumulado hasta 2005:3. Esto, si bien puede ser el resultado conjunto de un mayor número de afiliados (activos y no activos), mayores contribuciones (por salarios más altos y mayores porcentajes de cotización) y rendimientos más altos, constituye un hecho positivo.

Los cuatro paneles de la Figura 6 contienen información complementaria del régimen. Así, por ejemplo, en el panel A se observa cómo los afiliados de edades entre 60 y 64 años tienen la menor participación. De igual manera se observa una rápida caída en el número de afiliados de 65 años y más y de 15 a 19 años, aunque la caída en este último número es mucho más lenta que

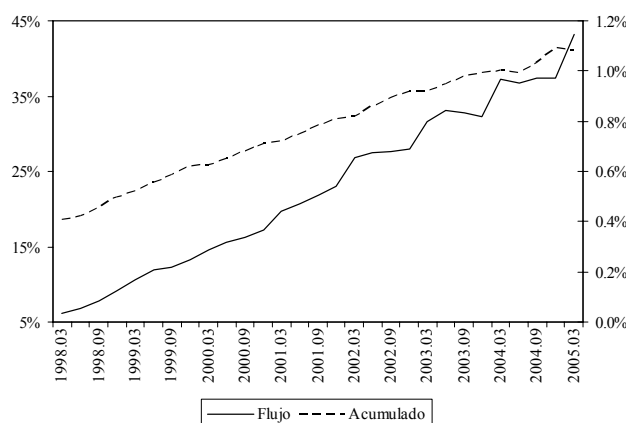
⁷ Esto es un argumento a favor de la existencia de un mecanismo de ahorro forzoso, como las pensiones, para financiar el consumo en la edad adulta. Según Flórez (2006), entre 1996 y 2000 las tasas de ahorro de Japón, los países de la zona Euro y Estados Unidos fueron de 30%, 21% y 18%, respectivamente. Entre 2001 y 2005, las tasas de ahorro fueron de 27%, 21% y 15%, respectivamente.

⁸ Sánchez (1998) y GRECO (2002) contienen estudios recientes sobre la evolución del ahorro y sus determinantes en el mediano y largo plazos.

en el anterior. La mayor participación en este Régimen corresponde a los afiliados de 20 a 39 años de edad.

El panel B muestra una mayor participación de afiliados pertenecientes al género masculino; sin embargo, la brecha por géneros ha tendido a cerrarse ya que en 1997 los hombres representaban 62.9% mientras que en mayo de 2005 su participación era de 59%. Para percibir el efecto de las decisiones (óptimas) de los agentes en cuanto a su participación en el mercado laboral, estos porcentajes deben compararse con otros indicadores. Por ejemplo, para 13 ciudades, la participación de hombres en la población total es de 46.9%; en la población en edad de trabajar de 45.6%; en la población económicamente activa de 52.4%; en la población ocupada de 53.8%; y de 43.8% entre los desocupados. Sugieren los números anteriores que la composición por género que se observa en los afiliados al sub-sistema de ahorro individual parece estarse acercando paulatinamente a la que se observa en la población económicamente activa. Sin embargo, el tema de la afiliación por género amerita una reflexión que toca aspectos más profundos del mercado laboral colombiano.

Figura 5. Ahorro pensional obligatorio sobre PIB trimestral y sobre PIB acumulado

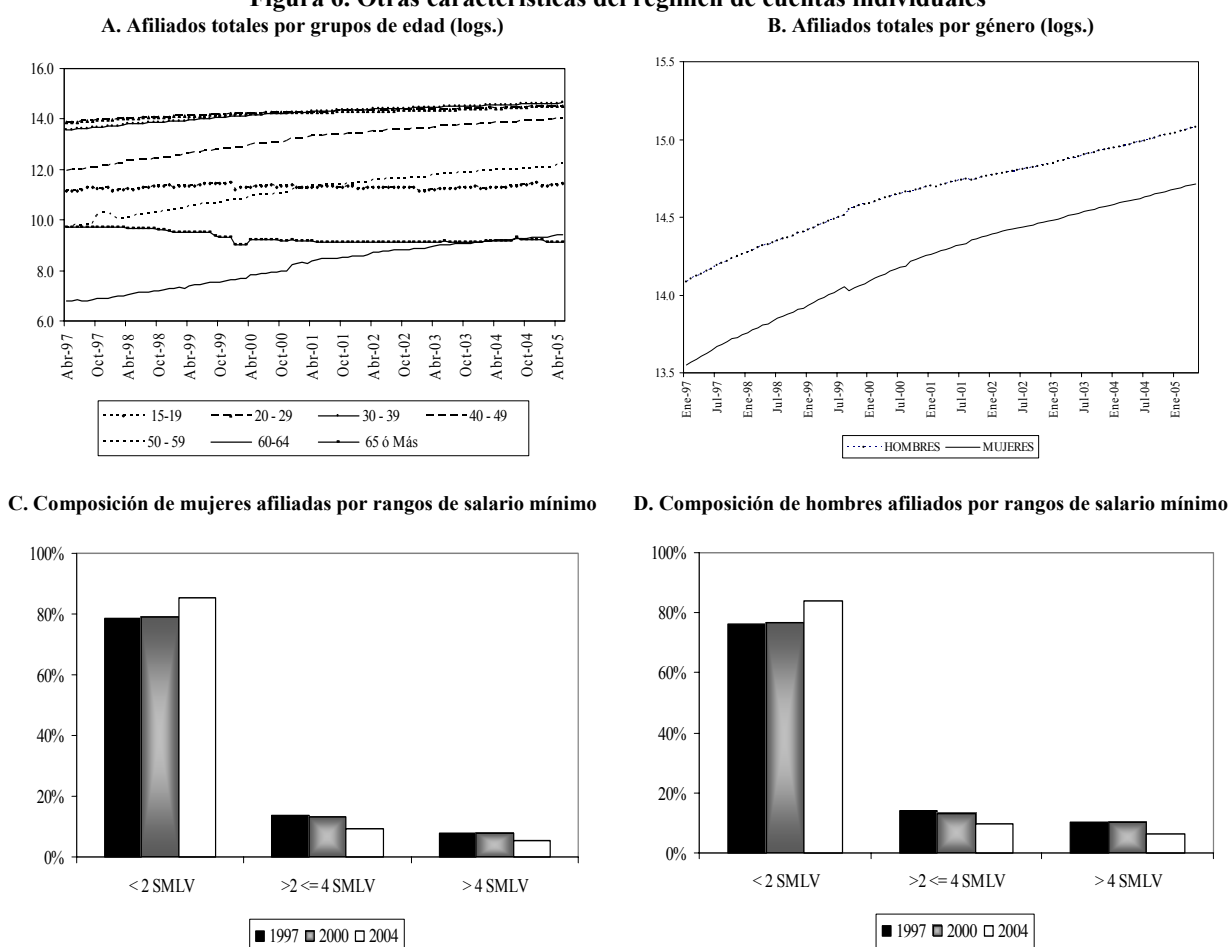


Nota: se refiere al cociente del ahorro pensional sobre el PIB de cada trimestre y al cociente del ahorro pensional sobre el PIB acumulado desde el segundo semestre de 1994, año en que surgieron las administradoras de fondos. Fuente: Superbancaria-DANE; cálculos de los autores.

Los paneles C y D de la Figura 6 muestran la afiliación de mujeres y hombres, respectivamente, por número de salarios mínimos. Las participaciones son relativamente homogéneas; sin embargo, llama la atención la forma en que han venido ganando participación los afiliados que ganan entre 1 y 2 salarios mínimos legales vigentes (*SMLV*) y se han reducido

los que ganan más de 2⁹. El aumento en la participación de mujeres que ganan entre 1 y 2 *SMLV* entre 1997 y 2004 es de 6.9 puntos porcentuales (*pp*) y de 7.8 *pp* para los hombres. Por su parte, la caída en la participación de las mujeres que ganan entre 2 y 4 *SMLV* es de 4.2 *pp* y de 4.1 *pp* en los hombres mientras que entre quienes ganan más de 4 *SMLV* la caída en la participación de mujeres y hombres fue de 2.7 y 3.7 *pp*, respectivamente¹⁰.

Figura 6. Otras características del régimen de cuentas individuales



Nota: SMLV: salario mínimo legal vigente.

⁹ De inmediato surge la pregunta sobre el grado de representatividad que esta estructura salarial tiene del mercado laboral colombiano.

¹⁰ Sin embargo, el nivel de detalle de la información aún no permite hacer análisis más profundos ni verificar otro tipo de hipótesis relacionadas con el grado de aversión al riesgo por género, rango de edad, nivel salarial, y otras características de los afiliados. Tampoco permite modelar algún tipo de heterogeneidad ni la estimación de la reacción ante diferenciales sostenidos de la tasa de rentabilidad real entre los fondos para las categorías mencionadas, etc. La Superintendencia Bancaria y Asofondos, entidad que agremia las *AFP*, deberían buscar un mayor apoyo de estas para generar información a nivel de ahorrador.

2.1. Rentabilidad de las *AFP* y rentabilidades mínima y del portafolio de referencia

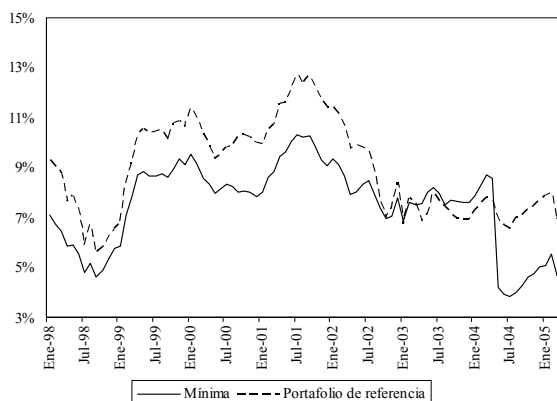
En este momento es conveniente señalar en qué consiste la rentabilidad mínima obligatoria. De acuerdo con el decreto 1592 de 2004 esta debe ser equivalente al promedio simple de: 1) el 70% del promedio ponderado de las rentabilidades acumuladas efectivas anuales para los últimos 36 meses y 2) el promedio ponderado de: *i*) el 70% del incremento porcentual efectivo anual durante los últimos 36 meses del índice de la Bolsa de Valores de Colombia, ponderado por el porcentaje del portafolio de los fondos invertido en acciones y en fondos de inversión en la proporción invertida en acciones; *ii*) el 70% del incremento porcentual efectivo anual durante los últimos 36 meses del índice representativo del mercado accionario del exterior indicado por la Superintendencia Bancaria, ponderado por el porcentaje del portafolio de los fondos invertido en acciones de emisores extranjeros y en fondos de inversión internacionales en la proporción invertida en acciones; *iii*) el 70% de la rentabilidad acumulada efectiva anual arrojada para los últimos 36 meses por un portafolio de referencia para pensiones valorado a precios de mercado, ponderado por el porcentaje invertido en las demás inversiones admisibles. Por su parte, la rentabilidad del portafolio de referencia hace relación al portafolio de base para el cálculo de la rentabilidad mínima que estipula la regulación expedida por la Superbancaria.

La Figura 7 muestra la evolución tanto de la rentabilidad mínima exigida a las *AFP* por la Superintendencia Bancaria como de la rentabilidad del portafolio de referencia. Se observa allí el efecto sobre la rentabilidad mínima del cambio en la fórmula de cálculo ocurrido en abril de 2004. Entre las justificaciones a dicha modificación, estuvo la mayor posibilidad de inversión para las *AFP* dada la menor probabilidad de penalización a la que están sometidas, desde abril de 2004, de no llegar a obtener al menos dicha rentabilidad.

Tanto la legislación reseñada como la Figura 8 que se observa un poco más adelante sugieren relaciones entre la rentabilidad de las *AFP* y las rentabilidades derivadas de la regulación. Por ello, a continuación se verifica la hipótesis de que el cambio en la rentabilidad de los fondos se relaciona, en un contexto panel, con el cambio en la rentabilidad mínima y con el cambio en el rendimiento del portafolio de referencia. Estas verificaciones se hacen utilizando el enfoque de “regresiones aparentemente no relacionadas” (o *Seemingly Unrelated Regressions, SUR*)¹¹.

¹¹ Las pruebas de raíz unitaria indican que estas series son I(1). Adicionalmente, se hicieron pruebas de cointegración tanto panel como individual pero no se encontró evidencia clara de comovimientos entre las rentabilidades de las

Figura 7. Rentabilidades mínima y del portafolio de referencia.

Tabla 1. Modelo *SUR* de la rentabilidad en función de la rentabilidad mínima

Variable	1998:01- 2004:04	1998:01- 2005:04	
Constante	0.00003 (0.0002)	-0.0001 (0.0002)	
Δ rentabilidad mínima	0.5680 (0.0440)	Colfondos	0.7058 (0.0689)
		Horizonte	0.7059 (0.0678)
		Porvenir	0.6233 (0.0712)
		Protección	0.7665 (0.0712)
		Santander	0.6794 (0.0671)
		Skandia	0.6035 (0.0689)
Δ rentabilidad mínima \times Δ dummy 2004:05	---	-0.5261 (0.0679)	
Δ dummy 2004:12	---	Colfondos	0.053 (0.0003)
		Horizonte	0.0080 (0.0003)
		Porvenir	0.0114 (0.0003)
		Protección	0.0052 (0.0003)
		Santander	-0.0026 (0.0003)
		Skandia	0.0082 (0.0003)
R^2	0.367	0.420	
Prueba de autocorrelación de Wooldridge (2002), Valor- <i>p</i>	0.865	0.208	
Prueba de normalidad J-B, valor- <i>p</i>	0.043	0014	

Nota: estimaciones panel *SUR* con corrección de errores estándar panel en el tiempo (Beck y Katz, 1995). Los residuos del modelo siguen un proceso *AR*(1). La variable dummy 2004:05 toma el valor de 1 si $t > 2004:05$ y 0 en otro caso. La variable dummy 2004:12 toma el valor de 1 si $t > 2004:12$ y 0 en otro caso. Δx corresponde a la diferencia de x y los números entre paréntesis a las desviaciones estándar. La frecuencia de la información es mensual.

AFP y la rentabilidad mínima o la rentabilidad del portafolio de referencia. Por lo tanto, el modelo es estimado sobre las diferencias de las series.

Los resultados de la Tabla 1 constituyen evidencia de asociación entre el cambio en la tasa de rentabilidad de las *AFP* y el cambio en la rentabilidad mínima tanto para el período 1998:1 - 2004:4, antes de introducir el cambio en la reglamentación que modifica el cálculo de la rentabilidad mínima, como para el período 1998:1-2005:4. Como se observa en la Tabla 1, en ambos períodos la rentabilidad mínima es determinante de la rentabilidad de las *AFP*. Esto es, ambos períodos el cambio en la rentabilidad mínima se transmite al cambio en la rentabilidad de las *AFP*.

Tabla 2. Modelo *SUR* de la rentabilidad en función de la rentabilidad del portafolio de referencia

Variable	1998:01- 2004:04	1998:01- 2005:04	
Constante	0.0002 (0.0003)	0.0001 (0.0003)	
Δ Rentabilidad Portafolio de Referencia	0.3158 (0.0391)	0.3320 (0.0594)	
Δ dummy 2004:12	----	Colfondos	0.0059 (0.0004)
		Horizonte	0.0086 (0.0003)
		Porvenir	0.0121 (0.0003)
		Protección	0.0059 (0.0005)
		Santander	-0.0016 (0.0003)
		Skandia	0.0088 (0.0003)
Δ dummy 2005:03	----	Colfondos	-0.0117 (0.0010)
		Horizonte	-0.0093 (0.0010)
		Porvenir	-0.0075 (0.0009)
		Protección	-0.0140 (0.0011)
		Santander	-0.0071 (0.0010)
		Skandia	-0.0085 (0.0010)
R^2	0.224	0.351	
Prueba de autocorrelación de Wooldridge (2002), Valor- <i>p</i>	0.806	0.958	
Prueba de normalidad J-B, valor- <i>p</i>	0.051	0.010	

Nota: estimaciones Panel SUR con corrección de errores estándar panel en el tiempo (Beck y Katz, 1995). Los residuos de los dos modelos son ajustados por un proceso AR(1). La variable dummy 2004:12 toma el valor de 1 si $t \geq 2004:12$ y 0 en otro caso. La variable dummy 2005:03 toma el valor de 1 si $t \geq 2005:03$ y 0 en otro caso. $\Delta(x)$ corresponde a la diferencia de x y los números entre paréntesis a las desviaciones estándar. La frecuencia de la información es mensual.

Durante el primer sub-período no se rechaza la hipótesis de que el impacto de la rentabilidad mínima es el mismo para la rentabilidad de todas las *AFP*. Cuando el período muestral se amplía hasta abril de 2005, el efecto de la rentabilidad mínima varía de *AFP* en *AFP*. Durante el período completo, fue necesario introducir dos variables *dummy* para tomar en cuenta variaciones significativas en la rentabilidad de los fondos. Notemos cómo la interacción (Δ rentabilidad mínima \times Δ dummy 2004:05), reduce en forma importante el efecto de la rentabilidad mínima en la rentabilidad de las *AFP* en ese mes particular. Los resultados de la Tabla 1 sugieren que el regulador tiene capacidad de alterar la rentabilidad de las *AFP* a través de la fijación de la rentabilidad mínima.

El cambio en la rentabilidad de las *AFP* está igualmente relacionado con el cambio en la rentabilidad del portafolio de referencia de acuerdo con los resultados de la Tabla 2. Cuando aumenta la rentabilidad del portafolio de referencia aumenta la rentabilidad de las *AFP*. Al igual que en el caso anterior, se puede rechazar la hipótesis de que el efecto marginal de la variación en la rentabilidad del portafolio de referencia en la rentabilidad es igual para todas las *AFP*. Se observa, sin embargo, que el efecto es mucho menor que el reportado en la Tabla 1 de la rentabilidad mínima.

3. Enfoque teórico y hechos subyacentes

La hipótesis fundamental de este estudio es que la tasa de rentabilidad real que ofrecen las *AFP* es un determinante fundamental en la composición del número de afiliados activos. Para racionalizar el asunto, suponemos que, dependiendo de la etapa de la vida en la que se encuentre, la laboral o la de retiro, el individuo tiene un problema en dos dimensiones: una intertemporal y una intratemporal. Suponemos que sus preferencias no se modifican al pasar de una etapa de la vida a la siguiente pero las restricciones que enfrenta sí. De esta manera, la primera dimensión del problema lo lleva a escoger la trayectoria de consumo que maximiza su satisfacción sujeto a una restricción de presupuesto. Para analizar esto, supongamos que la utilidad esperada por el individuo está dada por:

$$U = \sum_{s=t}^T \beta^{s-t} u(c_s) \quad (1)$$

donde U es una función que crece monótonamente, $u(c_s)$ la función de utilidad periódica, c_s el consumo en el momento s y β el factor de descuento. Suponemos que el individuo no deriva

ninguna utilidad del ocio y que, por tanto, ofrece su trabajo inelásticamente a la tasa de salario w'_s , que corresponde al salario de mercado w_s adicionado en la proporción δ_p , $0 < \delta_p < 1$, que es el aporte correspondiente a la firma en la que labora el individuo por concepto de pensión, neto de gastos de comisión de administración y seguros; de esta manera $w'_s = (1 + \delta_p)w_s$. Durante su etapa laboral el individuo desea maximizar (1) sujeto a:

$$A_{s+1} = (1 + r_s)A_s + (1 - \delta_e)w_s - c_s \quad (2)$$

siendo A_s el nivel de riqueza al comienzo del periodo s , r_s la tasa de interés real y δ_e , $0 < \delta_e < 1$, la fracción del ingreso laboral que el individuo debe trasladar periódicamente a la *AFP* que él haya elegido, neto de gastos de administración y seguros. Para ajustar nuestros supuestos a las normas (Ley 797 de 2003), podemos pensar que a la fecha $\delta_p + \delta_e = 10.5\%$ del salario y que a partir de 2006, $\delta_p + \delta_e = 11\%$. La diferencia entre estos valores y el aporte total (15% en 2005 y 15.5% en 2006) se destina a financiar el fondo de pensión mínima, los gastos de administración y la prima de reaseguro que mencionamos al comienzo.

La ecuación de Euler derivada del proceso de optimización con respecto a A_{s+1} y c_s , en esta dimensión es típicamente:

$$1 = \beta(1 + r_s)[u'(c_{s+1})/u'(c_s)] \quad (3)$$

Para analizar la dimensión intratemporal, supongamos que el individuo tiene libertad de elegir la *AFP* en la cual quiere capitalizar su ahorro pensional y que el criterio racional es el de la rentabilidad neta de comisiones de administración: la *AFP* que le pague el mejor retorno será la que administre sus ahorros. En este caso el individuo desea escoger el retorno más alto entre dos¹² posibles alternativas ofrecidas por las AFP_i y AFP_j , respectivamente, $\max(r_s^i, r_s^j)$, sujeto a un costo η_s que el individuo está dispuesto a asumir y que, en esencia, está representado en tiempo de búsqueda y procesamiento de la información relevante para la toma de la decisión que le permita llevar sus ahorros al máximo, al finalizar el período $s+1$, de la siguiente manera:

$$\text{Max } f_{s+1}^k = (1 + r_s^k)f_s^k + \bar{\delta}w_{s+1} \quad k = i, j \quad (4)$$

$$\text{con } \bar{\delta} = (\delta_p + \delta_e) \quad (5)$$

¹² En esta exposición suponemos la existencia de solo dos *AFP* por conveniencia.

siendo f_s^k la suma capitalizada en el sistema de ahorro individual bien sea que la coloque en la AFP_i o en la AFP_j y r_s^k , $k = i, j$, son las tasas de retorno que ofrece cada una de ellas, respectivamente, y $\bar{\delta}$ es el aporte conjunto de empleado y empleador. En este problema las tasas de interés se asocian de la siguiente manera:

$$r_s^k = r_s + \varepsilon_s^k \quad k = i, j$$

donde $\varepsilon_s^k \sim i.i.d.(0, \sigma_{\varepsilon_s^k}^2)$ y ε_s^i y ε_s^j pueden estar contemporáneamente correlacionados. La suma f_{s+1}^k solo se puede utilizar después del retiro, el cual se produce en el período \bar{T} , cuando $w_s = 0$. En ese momento (cuando $s > \bar{T}$) el problema tiene solamente la dimensión intertemporal (suponemos que la persona ya no puede volver cambiar de AFP) y, por tanto, la restricción que enfrentará el individuo es:

$$A_{s+1} = (1 + r_s)A_s + \bar{p}_s - c_s \quad \text{para } s > \bar{T} \quad (6)$$

donde \bar{p}_s es el componente permanente de la corriente pensional, p_s , que recibe el individuo y que acá asociamos a la cuota periódica correspondiente a una tasa de interés variable. Dicho componente se calcula como:

$$\sum_{s=t}^T \frac{\bar{p}_s}{\prod_{v=t+1}^s (1 + r_v^k)} = \sum_{s=t}^T \frac{p_s}{\prod_{v=t+1}^s (1 + r_v^k)} \quad (7)$$

haciendo: $R_{t,s} = 1/\prod_{v=t+1}^s (1 + r_v^k)$, la ecuación (7) se puede escribir como:

$$\sum_{s=t}^T R_{t,s} \bar{p}_s = \sum_{s=t}^T R_{t,s} p_s \quad (8)$$

de manera que:

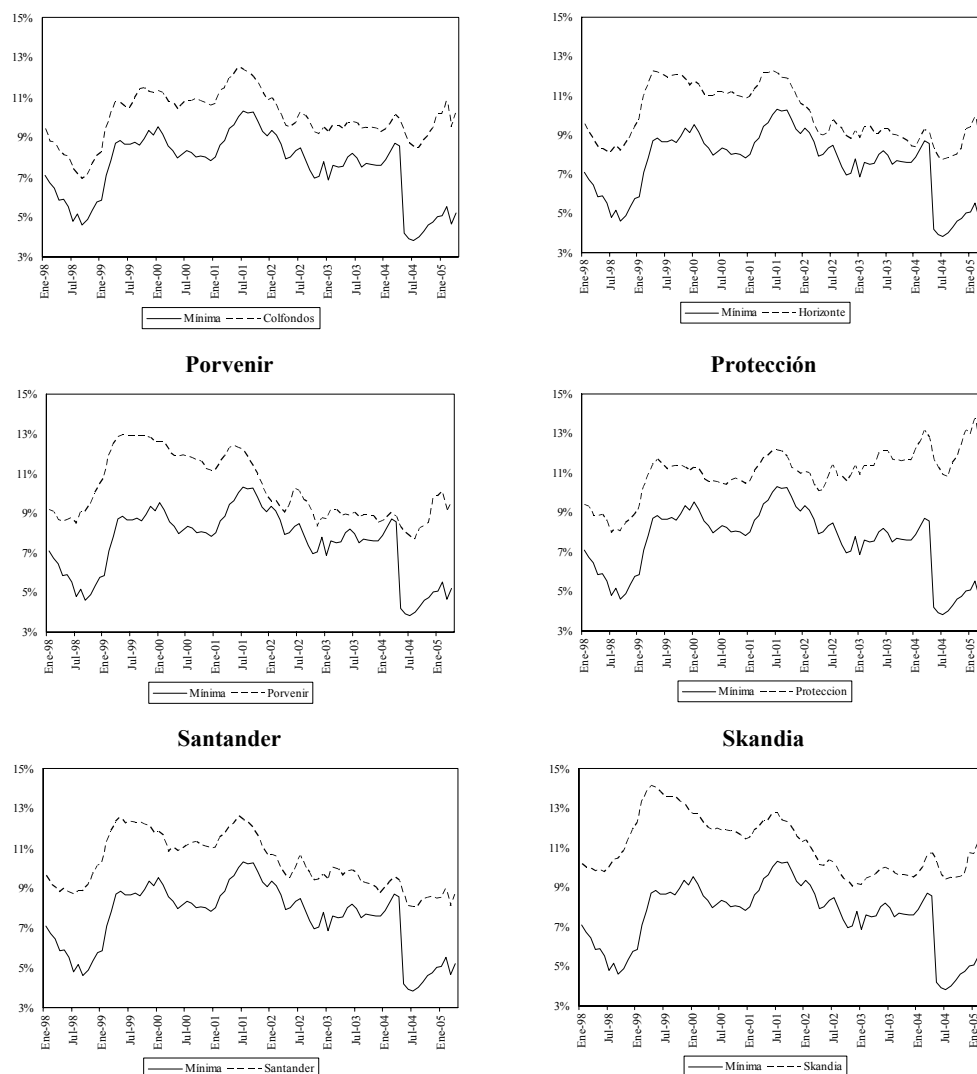
$$\bar{p}_s = \frac{\sum_{s=t}^T R_{s,t} p_s}{\sum_{s=t}^T R_{t,s}} \quad (9)$$

Con estos factores, las condiciones terminales para los problemas del individuo son:

$$\lim_{T \rightarrow \infty} R_{t,t+T}^{-1} A_{t+T+1} = 0 \quad \text{y} \quad \lim_{T \rightarrow \infty} [R_{t,t+T}^k]^{-1} f_{t+T+1}^k = 0 \quad (10)$$

El problema es convencional y la hipótesis que surge, dado que se trata de instituciones vigiladas por un órgano de control estatal como es la Superbancaria, es que el retorno que ofrecen las AFP es una variable determinante de la afiliación a una AFP u otra. Esta hipótesis se verifica a continuación siguiendo el enfoque de cointegración panel de Groen y Kleibergen (2003).

Figura 8. Rentabilidad de los portafolios administrados por las AFP y rentabilidad mínima



Nota: la rentabilidad real se calcula como $[1 + \text{rentabilidad efectiva anual}_t] / (1 + \pi_t) - 1$; la rentabilidad efectiva anual t es igual a rentabilidad efectiva anual del portafolio de 36 meses; π_t es la inflación de 36 meses efectiva anual. Fuente: cálculos de los autores con base en información de la Superintendencia Bancaria y la Dirección de Programación Macroeconómica e Inflación – Banco de la Republica, basada, a su vez, en información del DANE.

La Figura 8 muestra la rentabilidad real de los portafolios administrados por las AFP comparada siempre con la rentabilidad mínima. De allí se destacan varios puntos: *i*) el nivel relativamente alto de la rentabilidad real comparado con las estimaciones habituales de la tasa real de largo plazo las cuales fluctúan alrededor de 5% anual¹³; *ii*) el comovimiento aparente de la

¹³ Para Posada (1998) la tasa de interés real de equilibrio de largo plazo en Colombia se ubicó más o menos entre 3% y 5% hasta comienzos de los ochentas y, adicionada en un margen, ha sido muy similar a la de Estados Unidos. Flórez (2006) calcula que para el periodo 1996-2000 la tasa de interés real para las economías desarrolladas y

rentabilidad de las *AFP* y la rentabilidad mínima por lo menos hasta abril de 2004¹⁴; *iii*) el comportamiento a la baja del retorno real de algunos portafolios desde mediados de 2001 hasta mediados de 2004¹⁵; *iv*) el repunte reciente que ha tenido la rentabilidad en el cual se destaca la de Protección. Este evento se produce durante el período en el cual se reduce la rentabilidad mínima (o de penalización) como claramente se observa en la Figura 8, lo cual podría estar sugiriendo que la restricción que significa una rentabilidad mínima ha perdido vigencia práctica. Esto es, que el riesgo de obtener una rentabilidad por debajo de ella es realmente bajo y esto da más libertades a las *AFP*.

Si la hipótesis que se desprende de la teoría se cumpliera con exactitud, en el largo plazo, Skandia debería ser la *AFP* que más afiliados tenga, de continuar administrando el portafolio que produce el mayor rendimiento real como hasta ahora (11.06% anual durante el período de análisis). Sin embargo, se debe tener en cuenta la creencia de que Skandia se concentra en un segmento particular de asalariados y que, de momento, no tiene interés en apartarse de dicho segmento¹⁶. Un resultado de esta estrategia de Skandia se observa al comparar la participación de esta *AFP* en el patrimonio total del sub-sistema con la participación en los afiliados activos; la primera es casi cuatro veces la segunda (Figura 9), lo cual puede ser síntoma de que su estrategia de concentrarse en un segmento particular de los ocupados parece rendir los frutos esperados.

Dado lo anterior, sería entonces Protección el fondo que más afiliados tenga en el largo plazo ya que su retorno sigue muy de cerca el de Skandia y no parece tener interés particular o exclusivo en algún segmento de la población ocupada.

Por último, en relación con las implicaciones de nuestra hipótesis, en el largo plazo, Porvenir no debería ser la *AFP* que más afiliados activos tenga, como ocurre hoy en día, y Colfondos y Horizonte deberían ser las *AFP* con el menor número de afiliados dada su menor rentabilidad relativa.

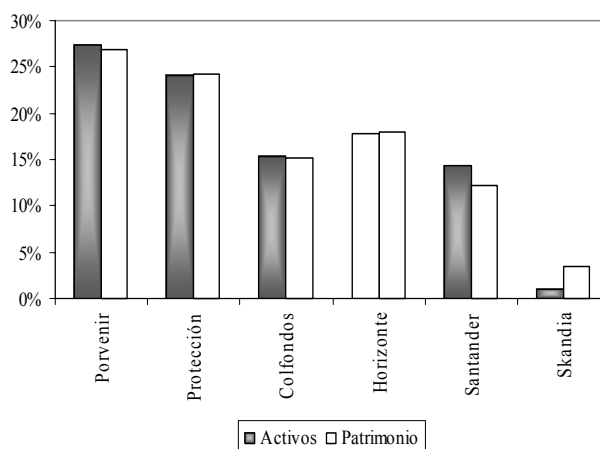
emergentes es de 3.2% y 1.8%, respectivamente. Para el período 2001-2005 las tasas de ubicaron en 2.1% y 4.0%, respectivamente.

¹⁴ Ya en el apartado 2.1 señalamos que no se logró obtener evidencia sobre dicho comovimiento ni en un contexto panel ni individualmente.

¹⁵ A este comportamiento escapan el fondo administrado por Protección, el cual creció en forma persistente desde 1999, y Colfondos que tuvo una destorcida menos fuerte.

¹⁶ La explicación racional a dicha estrategia, de llegar a ser cierta, es que posiblemente Skandia juzga costoso administrar portafolios pertenecientes a personas de bajos salarios y por esta razón solo se concentra en clientes de salarios altos.

Figura 9. Participación en patrimonio total y en afiliados activos al Régimen (2004)



Fuente: Superbancaria – cálculos de los autores.

4. Aproximación empírica

La estrategia empírica que vamos a seguir supone la existencia de otras variables que acompañen a la tasa de interés en la explicación de los afiliados activos a las distintas *AFP*. Como alternativa, vamos a considerar la población ocupada que sirve como variable de escala y es además la fuente natural de los afiliados activos a las *AFP*, y el valor de los fondos en relación con los afiliados como *proxy* de otros factores determinantes de la capacidad de cada entidad para atraer potenciales ahorradores.

Siendo clara la hipótesis que deseamos verificar y dado que disponemos de la serie reciente de los afiliados activos a cada una de las *AFP*, es conveniente plantear un ejercicio panel que nos permita aprovechar toda la información disponible tanto en forma longitudinal como transversal.

Una vez obtenida evidencia que no permite rechazar la hipótesis de que las variables de interés tienen raíz unitaria (ver Anexo 1) se establecen distintos modelos panel en los que se somete a prueba la hipótesis de que el número de afiliados activos y la rentabilidad de los portafolios de las *AFP* están cointegradas.

Las metodologías de cointegración panel se pueden clasificar en dos grupos. En el primero, se encuentran las contribuciones de Pedroni (1995, 1996), Kao (1999) y McCoskey y Kao (1998), las cuales corresponden a una extensión del procedimiento de Engle y Granger (1987). Sin embargo, este enfoque tiene dos desventajas. En primer lugar, no se puede estimar

más de un vector de cointegración y, en segundo, no permite modelar relaciones cruzadas entre individuos.

El segundo grupo, más cercano al procedimiento de Johansen (1991, 1996), incluye las contribuciones de Larsson, Lyhagen y Löthgren (2001) y Groen y Kleibergen (2003), las cuales, en contraste con las metodologías del primer grupo, permiten estimar más de un vector de cointegración.

La metodología de Groen y Kleibergen (2003), *GK* en lo sucesivo, tiene adicionalmente tres ventajas. En primer lugar, utiliza el método de máxima verosimilitud, el cual, además de sus propiedades asintóticas, permite evaluar diferentes tipos de hipótesis utilizando pruebas de razón de verosimilitud. En segundo término, el modelo de *GK* permite una retroalimentación instantánea entre los individuos que componen el panel¹⁷. Finalmente, esta metodología puede ser utilizada para probar la existencia de cointegración panel con parámetros de largo plazo homogéneos combinados con parámetros de corto plazo heterogéneos.

Para ilustrar un poco, supongamos un panel especificado para k variables, N individuos, p rezagos y T observaciones en el tiempo. En estos términos, la metodología *GK* de cointegración panel supone:

$$\Delta Y_t = \Pi^* Y_{t-1} + \Gamma^* W_t + \varepsilon_t \quad (11)$$

siendo:

$$\Delta Y_t = \begin{bmatrix} \Delta Y_{1,t} \\ \vdots \\ \Delta Y_{N,t} \end{bmatrix}; \quad \Pi^* = \begin{bmatrix} \Pi_1 & 0 & \dots & 0 & 0 \\ 0 & \ddots & & & 0 \\ 0 & 0 & \dots & 0 & \Pi_N \end{bmatrix}; \quad \Delta Y_{t-1} = \begin{bmatrix} \Delta Y_{1,t-1} \\ \vdots \\ \Delta Y_{N,t-1} \end{bmatrix}$$

$$\Gamma^* = \begin{bmatrix} \Gamma_1 & 0 & \dots & 0 & 0 \\ 0 & \ddots & & & 0 \\ 0 & 0 & \dots & 0 & \Gamma_N \end{bmatrix}; \quad \Delta W_t = \begin{bmatrix} \Delta Y_{1,t-1} \\ \vdots \\ \Delta Y_{1,t-p} \\ \vdots \\ \Delta Y_{N,t-1} \\ \vdots \\ \Delta Y_{N,t-p} \end{bmatrix}; \quad \Delta \varepsilon_t = \begin{bmatrix} \varepsilon_{1,t} \\ \vdots \\ \varepsilon_{N,t} \end{bmatrix}$$

¹⁷ Este efecto se captura mediante la matriz de covarianzas de los errores del modelo, la cual para cada observación en el tiempo no es una matriz diagonal.

donde $\Delta Y_{i,t}$, $Y_{i,t-1}$, $\Delta Y_{i,t-1}$ y $\varepsilon_{i,t}$ son vectores de órdenes $k \times 1$, Π_i es una matriz de orden $k \times k$ y Γ_i es una matriz de orden $k \times pk$. El primer componente del lado derecho de (11) representa la dinámica de largo plazo mientras que el segundo componente representa la de corto plazo. En forma compacta, este modelo puede ser escrito como¹⁸:

$$\Delta Y = Y_{-1}\Pi + W\Gamma + U \quad U \stackrel{i.i.d.}{\sim} N(0, \Omega) \quad (12)$$

Como en la metodología de Johansen (1991, 1996), la prueba de hipótesis sobre el número de relaciones de cointegración, para las k variables integradas de orden uno del vector $Y_{i,t}$, se realiza con base en una versión de rango reducido de la matriz Π^* especificada en (11). Dependiendo del grado de heterogeneidad de los parámetros de largo plazo (β_i), *GK* proponen dos especificaciones de rango reducido:

$$\Pi_B^* = \begin{bmatrix} \alpha_1 \beta_1' & 0 & \dots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \dots & 0 & \alpha_N \beta_N' \end{bmatrix} \quad \text{y,} \quad \Pi_C^* = \begin{bmatrix} \alpha_1 \beta_1' & 0 & \dots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \dots & 0 & \alpha_N \beta_N' \end{bmatrix}$$

donde α_i , β_i , β son matrices de dimensiones $k \times r$ para $i = 1, \dots, N$ y r representa el número de vectores de cointegración. Bajo los supuestos de que la matriz Π^* es diagonal por bloques y de que el rango de cointegración es común¹⁹, *GK* utilizan pruebas de hipótesis tipo *LR* para su evaluación.

En la Tabla 3 se describen las pruebas donde Π_A^* denota la matriz Π^* de rango completo²⁰. La distribución de las estadísticas *LR*, bajo la hipótesis nula, depende de los componentes determinísticos incluidos en el modelo y presentados en la Tabla 2. Al igual que en las metodologías tradicionales, los componentes determinísticos se refieren a la inclusión o no de intercepto y tendencia en el modelo como variables exógenas o dentro de los vectores de cointegración. La información de la Tabla 4 también considera si los parámetros β_i son homogéneos o heterogéneos entre individuos.

Para la estimación se asume que el proceso generador de datos es el correspondiente a la ecuación (12) y que el vector de errores U se comporta como un proceso ruido blanco

¹⁸ La especificación de las matrices y vectores de la ecuación (12) se encuentran en el Anexo 3.

¹⁹ $\text{Rango}(\Pi_i) = r$, para $i = 1, \dots, N$.

²⁰ El Anexo 2 contiene las distribuciones límite de las estadísticas de la Tabla 3.

multivariado con distribución normal y matriz de covarianza Ω , $U \stackrel{i.i.d.}{\sim} N(0, \Omega)$. El logaritmo de la función de verosimilitud está dado por:

$$l(\Pi, \Gamma, \Omega) = -\frac{NKT}{2} \ln(2\pi) - \frac{T}{2} \ln |\Omega| - \frac{1}{2} \text{vec}(U)' (\Omega^{-1} \otimes I_T) \text{vec}(U) \quad (13)$$

GK muestran que el estimador ML es equivalente en términos asintóticos a un estimador GMM . El Anexo 3 profundiza en algunos detalles de la metodología de estimación de GK .

Tabla 3. Pruebas de cointegración panel de GK

Prueba	Hipótesis nula	Estadística
Π_B^{\bullet} vs. Π_A^{\bullet}	$H_0 : \text{rango}(\Pi_{B,i}^{\bullet}) = \text{rango}(\alpha_i \beta_i') = r \quad \forall_i$	$LR(\Pi_B^{\bullet}(r) \Pi_A^{\bullet}) = T \left[\ln \Omega(\Pi_B^{\bullet}(r)) - \ln \Omega(\Pi_A^{\bullet}) \right]$
Π_C^{\bullet} vs. Π_A^{\bullet}	$H_0 : \text{rango}(\Pi_{C,i}^{\bullet}) = \text{rango}(\alpha_i \beta_i') = r \quad \forall_i$	$LR(\Pi_C^{\bullet}(r) \Pi_A^{\bullet}) = T \left[\ln \Omega(\Pi_C^{\bullet}(r)) - \ln \Omega(\Pi_A^{\bullet}) \right]$

Tabla 4. Componentes determinísticos considerados por GK

$$\Delta y_{i,t} = \delta_{1,i} + \delta_{2,i} t + \alpha_i \beta_i' y_{i,t-1} + \sum_j \gamma_{i,j} \Delta y_{i,t-j} + \varepsilon_{i,t}, \quad i = 1, \dots, N; \quad t = 1, \dots, T$$

Nombre del modelo	β_i	$\delta_{1,i}$	$\delta_{2,i}$
B_1	No restringido	No restringido	No restringido
B_2	No restringido	No restringido	$\delta_{2,i} = \alpha_i \mu_i'$
	No restringido	$\delta_{1,i} = \alpha_i \mu_i'$	0
C_1	$\beta_i = \beta$	No restringido	No restringido
C_2	$\beta_i = \beta$	No restringido	$\delta_{2,i} = \alpha_i \mu_i'$
	$\beta_i = \beta$	$\delta_{1,i} = \alpha_i \mu_i'$	0
C_3	$\beta_i = \beta$	No restringido	$\delta_{2,i} = \alpha_i \mu_i'$
	$\beta_i = \beta$	$\delta_{1,i} = \alpha_i \mu_i'$	0

5. Resultados

El primer modelo estimado, que adopta la forma C_1 de la Tabla 4, incluye además de la rentabilidad, el valor promedio del fondo por afiliado y la población ocupada. La primera variable se deriva directamente del modelo expuesto; la segunda opera como *proxy* de diversos factores (distintos de la rentabilidad) determinantes de la capacidad de la *AFP* para incorporar

nuevos clientes²¹. Finalmente, la población ocupada captura la fuente natural que surte al sistema de cuenta individual como son las personas ocupadas, dependientes e independientes; opera además como variable de escala.

La evidencia presentada en la Tabla 5 sugiere la existencia de, al menos, dos vectores de cointegración, el segundo de los cuales reporta lo esperado en cuanto a signo y magnitud de los coeficientes. Sin embargo, con ninguno de los dos vectores se puede rechazar la hipótesis de que los coeficientes son iguales para todos los fondos. Esto es, si en el largo plazo una *AFP* desea tener un mayor número de afiliados debe aumentar o bien la rentabilidad del portafolio que maneja o su capacidad para atraer nuevos ahorradores lo cual significa contar con una mayor difusión en los medios de información, una mayor fuerza comercial, mejores estrategias de aproximación a la población ocupada, etc. El aumento en los ocupados favorece al sistema en su totalidad y a todas la *AFP* por igual a juzgar por el vector de cointegración 2.

Tabla 5. Modelo C_1 para activos, rentabilidad, valor promedio del fondo y ocupados

$$\Delta y_{i,t} = \delta_{1,i} + \delta_{2,i}t + \alpha_i \beta' y_{i,t-1} + \sum_j \gamma_{i,j} + \varepsilon_i, \quad i = 1, \dots, N; \quad t = 1, \dots, T$$

Relación de largo plazo $\Rightarrow [Activos = \beta_0 + \beta_1 \times Rentabilidad\ media + \beta_2 \times Valor\ promedio + \beta_3 \times Ocupados]$

<i>AFP</i>	Variable	Vector de cointegración 1	Vector de cointegración 2
Efecto general	Rentabilidad media	14.00 (0.003)	1.95 (<0.001)
Efecto general	Valor promedio	2.36 (<0.001)	0.37 (0.170)
Efecto general	Población ocupada	-1.25 (<0.001)	0.43 (0.003)
	Razón de verosimilitud	Valor-p	Rango
	275.319	<0.001	0
	199.973	<0.001	1
	79.009	0.021	2
	33.135	0.056	3

Nota: el período muestral es 2000:1 - 2005:4. La variable correspondiente al logaritmo de los ocupados fue desestacionalizada con X11. Las variables de valor promedio del fondo y ocupados están en logaritmos. Entre paréntesis aparece el valor-p asociado a la hipótesis de que el coeficiente es igual a cero. La frecuencia de la información es mensual.

Sin embargo, la existencia de dos vectores de cointegración solo se produce cuando se utiliza un nivel de significancia de 1%²². Por lo tanto, se llevó a cabo un segundo ejercicio de cointegración panel modificando la medición de la rentabilidad real ya que en lugar de trabajar

²¹ Personas vinculadas al sector han expresado opiniones en el sentido de que, más que la rentabilidad que ofrezca el fondo, la variación en el número de afiliados está determinada por políticas de mercadeo y publicidad, segmentos de interés y estrategias de posicionamiento.

²² Adicionalmente el coeficiente del valor promedio no es significativo en el vector de cointegración 2.

con el promedio de 6 meses se utilizó la rentabilidad contemporánea. De igual manera, se dejó de incluir la variable valor promedio del fondo ya que las pruebas así lo recomendaban; en este caso los componentes determinísticos del modelo seleccionado solo incluyen constantes dentro del vector de cointegración.

Los resultados aparecen en la Tabla 6 en la cual el modelo seleccionado corresponde al C_2 de la Tabla 4. De nuevo las pruebas sugieren la existencia de dos vectores de cointegración; la diferencia con respecto al ejercicio de la Tabla 5 está en que ahora los coeficientes de ambos vectores de cointegración son significativos y exhiben los signos esperados.

De acuerdo con los resultados de la Tabla 6, en el largo plazo, tanto la rentabilidad como los ocupados en la economía aumentan los afiliados activos a las *AFP*. Cada variable, a su turno, afecta por igual a todas las instituciones; esto es, el coeficiente asociado a la tasa de rentabilidad no cambia de *AFP* en *AFP*. Igual sucede con el coeficiente asociado a la población ocupada. Sin embargo, debe tenerse en cuenta que esta última variable es exógena mientras que las *AFP* están en capacidad de hacer gestión de portafolio para modificar la rentabilidad y, con esta, el número de afiliados aportantes sujetándose en todo caso a las restricciones que impone la Superintendencia Bancaria.

Tabla 6. Modelo C_2 para activos, rentabilidad y ocupados

$$\Delta y_{i,t} = \delta_{1,i} + \alpha_i \mu_i' t + \alpha_i \beta_i' y_{i,t-1} + \sum_j \gamma_{i,j} \Delta y_{i,t-j} + \varepsilon_{i,t} \quad i = 1, \dots, N; \quad t = 1, \dots, T$$

$$\text{Relación de largo plazo} \Rightarrow [\text{Activos} = \beta_0 + \beta_1 \times \text{Rentabilidad media} + \beta_2 \times \text{Ocupados}]$$

<i>AFP</i>	Variable	Vector de cointegración 1	Vector de cointegración 2
Efecto general	Rentabilidad	5.63 (<0.001)	5.21 (<0.001)
Efecto general	Población ocupada	4.12 (<0.001)	3.44 (<0.001)
	Razón de verosimilitud	Valor-p	Rango
	470.568	<0.001	0
	406.163	<0.001	1
	44.796	0.100	2

Nota: la muestra va de 2000:1 a 2005:4. La variable correspondiente al logaritmo de los ocupados fue desestacionalizada con X11. La variable "población ocupada" está en logaritmos. Entre paréntesis aparece el valor-p asociado a la hipótesis de que el coeficiente es igual a cero. La frecuencia de la información es mensual.

En síntesis, los resultados sugieren que, dada la estrategia de cada *AFP*, el número de afiliados a los distintos fondos que se encuentran activos están asociados en el largo plazo con la tasa de retorno real que obtenga cada *AFP* y la población ocupada del país. Este resultado se obtiene a pesar de que la gran masa de afiliados (más del 80%) cotizan sobre la base de menos de

dos salarios mínimos y sería de esperar que estas personas estén menos atentos a la evolución de la tasa de retorno de las *AFP* ya que el Régimen de Ahorro Individual con Solidaridad garantiza una pensión de vejez que como mínimo debe ser el salario mínimo.

Dos implicaciones inmediatas de nuestros resultados son las siguientes. En primer lugar, una manera de conseguir que las entidades del Régimen de Ahorro Individual logren atraer más ahorradores es permitiendo “mayores libertades” a las *AFP* para elegir los activos que van a conformar sus portafolios de manera que la rentabilidad real que se pueda obtener sea atractiva para los futuros pensionados ya que para ellos la tasa de retorno real es decisiva. La segunda tiene que ver con modificaciones al esquema de incentivos que hagan, de nuevo, aumentar la rentabilidad que pueden ofrecer las *AFP*. Esto se podría lograr haciendo que, por ejemplo, las comisiones de administración, a las que nos referíamos en la introducción dependan, al menos parcialmente, de los rendimientos que logren las *AFP* para los cuenta-habientes y no se cobren únicamente sobre los aportes mensuales.

6. Conclusiones

Este trabajo utiliza la información de ahorro individual del Sistema General de Pensiones para poner a prueba la hipótesis de que los afiliados activos a las *AFP* responden a la tasa de retorno de los portafolios que ellas administran. Para formular esta hipótesis el documento presenta un modelo simple en el que hay tanto decisiones intertemporales como intratemporales de los agentes.

Siguiendo un enfoque de cointegración panel (Groen y Kleibergen, 2003) en el que se utiliza la información de cada *AFP* entre 1998 y 2005, se estiman dos modelos cada uno con una medida distinta de la tasa de retorno real.

La evidencia que se desprende de nuestra especificación favorita sugiere que el número de cotizantes (afiliados aportantes o activos) está cointegrado con la tasa de retorno real de los portafolios y con la población ocupada, variable ésta que constituye la fuente natural de los cotizantes al sistema privado de pensiones, en el cual se concentra nuestro trabajo, y que opera además como variable de escala. Los resultados no permiten rechazar la hipótesis de que los coeficientes asociados a la tasa de retorno y a la población ocupada son iguales para todas las *AFP*. El valor promedio del fondo, utilizado como *proxy* de otras variables como la capacidad

que tiene cada administradora de difundir su producto y de contactar potenciales afiliados, no resultó significativo.

Dado el resultado de que la rentabilidad mínima afecta la rentabilidad de las *AFP*, la implicación de este ejercicio es que una manera de aumentar el número de afiliados activos al sistema de ahorro individual es liberando a las *AFP* de generar retornos mínimos, para que puedan cautivar más ahorradores y se estimule la competencia, y de regulaciones que impidan maximizar el retorno dado un nivel de riesgo permisible. Otra manera de obtener una mayor rentabilidad es modificando la forma en que cobran las comisiones de administración: en lugar de ser un porcentaje sobre los aportes mensuales como prescriben las normas actualmente, deberían ser una fracción del retorno que obtengan sobre el portafolio.

Referencias

Arena de Mesa, R., J. Behrman, y D. Bravo, 2004, Characteristics of and Determinants of the Density of Contributions in a Private Social Security System, *Working Paper*, Dirección de Impuestos, Ministerio de Hacienda, Chile.

Barr, N., 2000, Reforming Pensions: Myths, Truths, and Policy Choices, *IMF Working Paper*, 139.

Beck, N. y J. N. Katz, 1995, What to Do (and Not to Do) With Time-series Cross section Data, *American Political Science Review*, 89(3), 634-647.

Bonilla. R., 2001, Pensiones: en busca de la equidad, *Cuadernos de Economía*, Vol. XX, No. 34, 307-335.

Clavijo, S., 2002, *Sostenibilidad pensional y gasto social*, Alfaomega: Bogotá.

Doornik J.A., 1998, Approximations to the Asymptotic Distribution of Cointegration Tests, *Journal of Economic Surveys*, 12, 573-593.

Echeverry J.C., A. Escobar, C. Merchán, G. Piraquive, y M. Santa María, 2001, Elementos para el debate sobre una nueva reforma pensional en Colombia, *Archivos de Macroeconomía*, DNP.

Engle, R.F. y C.W.J. Granger, 1987, Co-Integration and Error Correction: Representation, Estimation and Testing, *Econometrica* 55, 251-276.

Flórez, L.A., 2006, Generación y flujo de ahorro en inversión: análisis de la coyuntura internacional, *Perfil de Coyuntura Económica*, No. 6, Universidad de Antioquia - CIE, próximo a aparecer.

GRECO, 2002, *El crecimiento económico colombiano en el siglo XX*, Fondo de Cultura Económica: Bogotá.

Groen, J.J.J, 2002, Cointegration and the Monetary Exchange Rate Model Revisited, *Oxford Bulletin of Economics and Statistics*, Blackwell Publishing, Vol. 64 No.4, 361-380.

Groen, J.J.J. y F. Kleibergen, 2003, Likelihood-Based Cointegration Analysis in Panels of Vector Error Correction Models, *Journal of Business & Economic Statistics*, Vol.21 No.2. 295-318.

Hadri, K. y R. Larsson, 2005, Testing for Stationarity in Heterogeneous Panel Data Where the Time Dimension is Finite. *The Econometrics Journal*, Vol. 8, No. 1, 55-69.

Hadri, K., 2000, Testing for Stationarity in Heterogeneous Panel Data. *The Econometrics Journal*, 3:148-161.

Im, K.S; Pesaran, M.H. y Shin, Y., 2003, Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, Vol. 115, No. 1, pp. 53-74, July.

Jara, D., C. Gómez y A. Pardo, 2006, Análisis de eficiencia de los portafolios pensionales obligatorios en Colombia, *Ensayos sobre Política Económica*, ESPE, No. 49, próximo a aparecer.

Johansen, S., 1991, Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59, 1551-1580.

Johansen, S., 1996, *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford, 2nd edition.

Kao, C., 1999, Spurious Regression and Residual-Based Tests for Cointegration in Panel Data. *Journal of Econometrics*, 90, 1-44.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. y Y. Shin, 1992, Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 91-115.

Larsson, R., J. Lyhagen y M. Lothgren, 2001, Likelihood-Based Cointegration Tests in Heterogeneous Panels. *Econometrics Journal*, 4, 109-142.

Maddala, G.S., Wu, S., 1999, A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

McCoskey, S., & Kao, C., 1998, A Residual-Based Test of the Null of Cointegration in Panel Data. *Econometric Reviews*, 17, 57-84.

Pedroni, P., 1995, Panel Cointegration: Asymptotics and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis, *Working Paper*, Department of Economics, No. 95-013, Indiana University.

Pedroni, P., 1996, Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity. Working paper, Department of Economics, No. 96-20, Indiana University.

Pesaran, H., 2003, A Simple Panel Unit Root Test in the Presence of Cross Section Dependence, September 2003, Revised January 2005, *Cambridge University DAE Working Paper 0346*.

Posada, C.E., 1998, La tasa de interés: el caso colombiano del siglo XX (1905-1997), *Ensayos sobre política económica*, No. 33, 5-60.

Sánchez. F., 1998, *El ahorro en Colombia*, TM Editores: Bogotá.

Urrutia, M., 1996, Primeros efectos de las reformas al sector financiero, Notas Editoriales, *Revista Banco de la República*, enero.

Wooldridge, J. M., 2002, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, 282-283.

Anexo 1. Pruebas de raíz unitaria para datos de panel

1. Prueba de raíz unitaria para datos de panel de Hadri (2000).

Hadri (2000) propone una prueba de raíz unitaria para datos de panel basada en una generalización de la prueba *KPSS* de Kwiatkowski et al. (1992), en la cual se considera el siguiente proceso para la serie y_{it} :

$$y_{i,t} = r_{i,t} + \varepsilon_{i,t} \quad (\text{A1.1})$$

$$r_{i,t} = r_{i,t-1} + u_{i,t} \quad (\text{A1.2})$$

donde $i=1, \dots, N$; $t=1, \dots, T$; ε_{it} y u_{it} son mutuamente independientes tales que $\varepsilon_{it} \xrightarrow{iid} N(0, \sigma_\varepsilon^2)$ y $u_{it} \xrightarrow{iid} N(0, \sigma_u^2)$.

Al igual que Kwiatkowski et al. (1992), Hadri (2000) considera la posibilidad de incluir una tendencia determinística; en este caso la ecuación (A1.1) es reemplazada por:

$$y_{i,t} = r_{i,t} + \beta_i t + \varepsilon_{i,t} \quad (\text{A1.3})$$

La hipótesis nula de esta prueba, $H_0 : \sigma_u^2 = 0$, implica que las series $\{y_{it}\}_{i=1, \dots, N}$ son estacionarias.

Hadri (2000) propone la siguiente estadística *LM* para evaluar esta prueba:

$$LM_j = \frac{1}{N} \sum_{i=1}^N KPSS_{j,i}; \quad j = \mu, \tau \quad (\text{A1.4})$$

donde $KPSS_{j,i}$ representa la estadística *KPSS*, no panel, para la variable y del individuo i . El primer subíndice indica el tipo de componente determinístico utilizado, $KPSS_{\mu,i}$ está asociado al modelo (A1.1) y $KPSS_{\tau,i}$ al modelo (A1.3).

2. Prueba de raíz unitaria para datos de panel de Pesaran (2003)

La literatura tradicional sobre pruebas de raíz unitaria para datos de panel, como Im, Pesaran y Shin (2003), Maddala y Wu (1999) y Hadri (2000), entre otros, supone que las series de tiempo individuales del panel están independientemente distribuidas entre ellas. Pesaran (2003) utiliza una metodología que relaja este supuesto. Este autor propone una estadística basada en regresiones similares a las de Dickey y Fuller, en contexto panel, adicionando rezagos de promedios seccionales cruzados de los niveles y primeras diferencias de las series individuales. Estas regresiones son de la siguiente forma:

$$\Delta y_{it} = \phi_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + \sum_{j=0}^p d_{ij} \Delta \bar{y}_{t-j} + \sum_{j=1}^p \delta_{ij} \Delta y_{i,t-j} + \varepsilon_{it}; \quad i=1, \dots, N, \quad t=p+2, \dots, T \quad (\text{A1.5})$$

donde $\bar{y}_t = \frac{1}{N} \sum_{j=1}^N y_{jt}$ y ϕ_i representa componentes determinísticos, que al igual que en la prueba de Dickey y Fuller pueden ser: ninguno, constante o constante y tendencia.

Pesaran (2003) evalúa la existencia de raíz unitaria en la serie $\{y_{it}\}$ en un contexto de datos de panel con N individuos y T observaciones en el tiempo, $(H_0 : y_{it} \sim I(1), i=1, \dots, N)$, utilizando la siguiente estadística:

$$CIPS = \frac{1}{N} \sum_{i=1}^N CADF_i^* \quad (\text{A1.6})$$

$$\text{donde: } CADF_i^* = \begin{cases} t_i & \text{si } -k_1 < t_i < k_2 \\ -k_1 & \text{si } t_i \leq -k_1 \\ k_2 & \text{si } t_i \geq k_2 \end{cases}, \quad t_i = \frac{\hat{b}_i}{\sqrt{V(\hat{b}_i)}}, \quad \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} -E(CADF_i) - \Phi^{-1}\left(\frac{\varepsilon}{2}\right) \sqrt{V(CADF_i)} \\ E(CADF_i) + \Phi^{-1}\left(1 - \frac{\varepsilon}{2}\right) \sqrt{V(CADF_i)} \end{bmatrix},$$

donde $\Phi(\cdot)$ es la función de distribución de una variable aleatoria normal estándar, ε es un número positivo lo suficientemente pequeño (por ejemplo, 1×10^{-6}) y b_i corresponde al coeficiente especificado en (A1.5).

Los valores críticos de la estadística $CIPS$ al igual que los valores esperados y varianzas de $CADF$ son tabulados en Pesaran (2003). Estos valores, $[E(CADF), V(CADF)]$ y los valores críticos de $CIPS$, dependen de los componentes determinísticos incluidos en la prueba, y de los valores de N y T .

Tabla A1.1. Raíces unitarias panel utilizando rezagos y componentes determinísticos de ADF individual

Variable	MW Ho:I(1)		
	Valor-p	Prueba	Resultado
Rentabilidad	0.191	15.996	I(1)
Rentabilidad media	0.087	19.059	I(1)
Activos	0.625	9.890	I(1)
Valor promedio	0.581	10.388	I(1)
	CIPS Ho:(1)		
	Valor crítico	Prueba	Resultado
Rentabilidad	-2.320	0.145	I(1)
Rentabilidad media	-2.830	-1.366	I(1)
Activos	-2.830	-1.815	I(1)
Valor promedio	-2.830	-2.241	I(1)
	HADRI Ho:I(0)		
	Valor-p	Prueba	Resultado
Rentabilidad	<0.0001	11.700	I(1)
Rentabilidad media	<0.0001	14.922	I(1)
Activos	<0.0001	19.448	I(1)
Valor promedio	<0.0001	25.206	I(1)

Tabla A1.2. Raíces unitarias panel utilizando el mismo número de rezagos

Variable	MW Ho:I(1)		
	Valor-p	Prueba	Resultado
Rentabilidad	0.090	18.922	I(1)
Rentabilidad media	0.071	19.793	I(1)
Activos	0.664	9.439	I(1)
Valor promedio	0.74	8.547	I(1)
	CIPS Ho:(1)		
	Valor Crítico	Prueba	Resultado
Rentabilidad	-2.320	0.431	I(1)
Rentabilidad media	-2.830	-1.295	I(1)
Activos	-2.830	-2.143	I(1)
Valor promedio	-2.830	-2.263	I(1)
	HADRI Ho:I(0)		
	Valor-p	Prueba	Resultado
Rentabilidad	<0.0001	11.700	I(1)
Rentabilidad media	<0.0001	14.922	I(1)
Activos	<0.0001	19.448	I(1)
Valor promedio	<0.0001	25.206	I(1)

Nota: la rentabilidad real se calcula como $(1 + \text{rentabilidad efectiva anual}) / (1 + \Pi_t) - 1$; rentabilidad efectiva anual, es la rentabilidad efectiva anual del portafolio a 36 meses; Π_t = inflación de 36 meses efectiva anual. Activos representa el logaritmo natural de la cantidad de afiliados aportantes activos a cada AFP. Valor promedio representa el logaritmo natural del valor promedio del fondo por afiliado a precios constantes de diciembre de 1998 (para calcular el valor del fondo se tomo el valor al último día de cada mes incluyendo rendimientos).

Anexo 2. Distribución de las estadísticas LR de las pruebas de rango para la cointegración panel de Groen y Kleibergen

En la Tabla A2.1 se presentan las distribuciones límite de las estadísticas consideradas en la Tabla 1 del texto para los diferentes componentes determinísticos considerados por Groen y Kleibergen (2003), **GK**. Estos resultados generalizan los de Johansen (1991) para el caso de cointegración panel²³.

Tabla A2.1 Pruebas de cointegración panel de Groen y Kleibergen (2003)

Prueba	Hipótesis nula	Estadística	Distribución Límite
Π_{B1}^{\bullet} vs. Π_A^{\bullet}	$H_0 : rango(\Pi_{B1,i}^{\bullet}) = rango(\alpha_i \beta_i') = r \quad \forall i$	$LR(\Pi_{B1}^{\bullet}(r) \Pi_A^{\bullet}) = T \left[\ln \Omega(\Pi_{B1}^{\bullet}(r)) - \ln \Omega(\Pi_A^{\bullet}) \right]$	$\sum_{i=1}^N tr(Q_{1,i})$
Π_{B2}^{\bullet} vs. Π_A^{\bullet}	$H_0 : rango(\Pi_{B2,i}^{\bullet}) = rango(\alpha_i \beta_i') = r \quad \forall i$	$LR(\Pi_{B2}^{\bullet}(r) \Pi_A^{\bullet}) = T \left[\ln \Omega(\Pi_{B2}^{\bullet}(r)) - \ln \Omega(\Pi_A^{\bullet}) \right]$	$\sum_{i=1}^N tr(Q_{2,i})$
Π_{C1}^{\bullet} vs. Π_A^{\bullet}	$H_0 : rango(\Pi_{C1,i}^{\bullet}) = rango(\alpha_i \beta_i') = r \quad \forall i$	$LR(\Pi_{C1}^{\bullet}(r) \Pi_A^{\bullet}) = T \left[\ln \Omega(\Pi_{C1}^{\bullet}(r)) - \ln \Omega(\Pi_A^{\bullet}) \right]$	$\chi^2(r(N-1)(k-r)) + \sum_{i=1}^N tr(Q_{1,i})$
Π_{C2}^{\bullet} vs. Π_A^{\bullet}	$H_0 : rango(\Pi_{C2,i}^{\bullet}) = rango(\alpha_i \beta_i') = r \quad \forall i$	$LR(\Pi_{C2}^{\bullet}(r) \Pi_A^{\bullet}) = T \left[\ln \Omega(\Pi_{C2}^{\bullet}(r)) - \ln \Omega(\Pi_A^{\bullet}) \right]$	$\chi^2(r(N-1)(k-r)) + \sum_{i=1}^N tr(Q_{2,i})$
Π_{C3}^{\bullet} vs. Π_A^{\bullet}	$H_0 : rango(\Pi_{C3,i}^{\bullet}) = rango(\alpha_i \beta_i') = r \quad \forall i$	$LR(\Pi_{C3}^{\bullet}(r) \Pi_A^{\bullet}) = T \left[\ln \Omega(\Pi_{C3}^{\bullet}(r)) - \ln \Omega(\Pi_A^{\bullet}) \right]$	$\chi^2(r(N-1)(k-r+1)) + \sum_{i=1}^N tr(Q_{2,i})$

En la Tabla A2.1., la distribución límite que aparece en la última columna de la derecha se obtiene como:

²³ GK prueban que la distribución asintótica de las pruebas LR asociadas al rango de cointegración panel con diferentes componentes determinísticos y diferentes grados de heterogeneidad de los β^1 s , equivale a la suma de N distribuciones límite de la estadística de la traza de Johansen (1991).

$$Q_{1,i} = \int dB_{k-r,i} S_i' \left(\int S_i S_i' \right)^{-1} \int S_i dB_{k-r,i}', \text{ donde:}$$

$$S_i(t) = \begin{pmatrix} B_{k-r-1,i}(t) - \int_0^1 B_{k-r-1,i}(t) dt \\ t - \frac{1}{2} \end{pmatrix}, \quad 0 \leq t \leq 1$$

$$Q_{2,t} = \int dB_{k-r,i} S_i' \left(\int S_i S_i' \right)^{-1} \int S_i dB_{k-r,i}' \text{ con:}$$

$$S_i(t) = \begin{pmatrix} B_{k-r-1,i} dt \\ 1 \end{pmatrix}$$

$B_{k-r,i}$ es un movimiento Browniano de dimensión $(k-r)$ para el individuo i con una matriz de covarianzas igual a la identidad.

Metodología de Doornik (1998)

En el caso de cointegración individual (no panel), Doornik (1998) aproxima las distribuciones de las estadísticas LR , asociadas a las pruebas de rango de cointegración de Johansen (1991), mediante una distribución Gamma con el mismo valor esperado y varianza de las variables aleatorias cuya distribución se desea aproximar. Por lo tanto, los valores críticos de significancia o valores- p asociados a la distribución de estas estadísticas son generados a partir de los valores- p de una variable aleatoria que sigue dicha distribución Gamma. Mediante simulaciones Monte Carlo, Doornik (1998) encuentra las siguientes aproximaciones para los valores esperados y varianzas asociados a la distribución de las estadísticas LR (denotadas como Z) de las pruebas de rango de cointegración:

$$E(Z) \approx \begin{cases} 2m^2 - m + 0.07 + 0.07 I_{\{m=1\}} & \text{para } H_z \\ 2m^2 + 2.01m + 0.06 I_{\{m=1\}} + 0.05 I_{\{m=2\}} & \text{para } H_c \\ 2m^2 + 1.05m - 1.55 - 0.5 I_{\{m=1\}} - 0.23 I_{\{m=2\}} & \text{para } H_{lc} \\ 2m^2 + 4.05m + 0.5 - 0.23 I_{\{m=1\}} - 0.07 I_{\{m=2\}} & \text{para } H_l \end{cases} \quad (\text{A2.1})$$

$$V(Z) \approx \begin{cases} 3m^2 - 0.33m - 0.55 & \text{para } H_z \\ 3m^2 + 3.60m + 0.75 - 0.40 I_{\{m=1\}} - 0.30 I_{\{m=2\}} & \text{para } H_c \\ 3m^2 + 1.80m - 2.80 I_{\{m=1\}} - 1.10 I_{\{m=2\}} & \text{para } H_{lc} \\ 3m^2 + 5.70m + 3.20 - 1.30 I_{\{m=1\}} - 0.50 I_{\{m=2\}} & \text{para } H_l \end{cases} \quad (\text{A2.2})$$

donde:

$$I_{\{m=a\}} = \begin{cases} 1 & \text{si } m = a \\ 0 & \text{si } m \neq a \end{cases}$$

y m corresponde al número de variables endógenas (k) menos el número de vectores de cointegración (r), H_z , H_c , H_{lc} y H_l indican los componentes determinísticos incluidos en el modelo VEC^{24} .

Debido a que la prueba de Doornik se basa en la comparación de (A2.1) y (A2.2) con los valores esperados y varianzas de una variable aleatoria con distribución Gamma, es conveniente recordar los dos primeros momentos de esta distribución. Una variable aleatoria X con distribución Gamma tiene la siguiente función de distribución con parámetros a y b :

$$\Gamma(x, a, b) = \int_0^x \frac{a^b}{\Gamma(b)} t^{b-1} e^{-at} dt, \quad x > 0, b > 0, a > 0$$

donde $\Gamma(b)$ representa la función Gamma y es definida como $\Gamma(b) = \int_0^\infty t^{b-1} e^{-t} dt$. Bajo esta

distribución $E(x) = \frac{b}{a}$ y $V(x) = \frac{b}{a^2}$; por lo tanto, igualando los valores esperados y varianzas de la estadística LR y de la distribución Gamma se obtienen los parámetros para esta última distribución:

$$a = \frac{E(LR)}{V(LR)}, \quad b = \frac{E^2(LR)}{V(LR)}$$

En conclusión, la prueba de cointegración para modelos individuales (no panel) está basada en la estadística de la traza propuesta por Johansen. Doornik (1998) sugiere que una aproximación a la distribución de estas estadísticas LR mediante el uso de una distribución Gamma con el valor esperado y varianza especificados en (A2.1) y (A2.2).

²⁴ En términos de los componentes determinísticos especificados en el Tabla 2 del texto se tiene la siguiente notación: $H_z = \delta_1 = \delta_2$, $H_c = \delta_1 = \alpha\mu'$, $\delta_2 = 0$, $H_{lc} = \delta_1$ no restringido, $\delta_2 = 0$, y $H_l = \delta_1$ no restringido, $\delta_2 = \alpha\mu'$

Metodología de Groen (2002)

En la Tabla A2.2 se presentan las aproximaciones de las distribuciones de las estadísticas LR utilizadas en la cointegración panel (Tabla A2.1). Groen (2002) propone estas aproximaciones basado en los resultados de Doornik (1998).

Tabla A2.2. Aproximación Gamma de las distribuciones de las estadísticas LR de Groen y Kleibergen (2003)

$$(LR \xrightarrow{A} \Gamma(a, b))^{25}$$

Estadística	a	b
$LR(\Pi_{b1}^*(r) \Pi_A^*)$	$\frac{E[tr(Q_1)]}{V[tr(Q_1)]}$	$\frac{NE^2[tr(Q_1)]}{V[tr(Q_1)]}$
$LR(\Pi_{b2}^*(r) \Pi_A^*)$	$\frac{E[tr(Q_2)]}{V[tr(Q_2)]}$	$\frac{NE^2[tr(Q_2)]}{V[tr(Q_2)]}$
$LR(\Pi_{c1}^*(r) \Pi_A^*)$	$\frac{r(N-1)(k-r) + NE[tr(Q_1)]}{2r(N-1)(k-r) + NV[tr(Q_1)]}$	$\frac{(r(N-1)(k-r) + NE[tr(Q_1)])^2}{2r(N-1)(k-r) + NV[tr(Q_1)]}$
$LR(\Pi_{c2}^*(r) \Pi_A^*)$	$\frac{r(N-1)(k-r) + NE[tr(Q_2)]}{2r(N-1)(k-r) + NV[tr(Q_2)]}$	$\frac{(r(N-1)(k-r) + NE[tr(Q_2)])^2}{2r(N-1)(k-r) + NV[tr(Q_2)]}$
$LR(\Pi_{c3}^*(r) \Pi_A^*)$	$\frac{r(N-1)(k-r+1) + NE[tr(Q_2)]}{2r(N-1)(k-r+1) + NV[tr(Q_2)]}$	$\frac{(r(N-1)(k-r+1) + NE[tr(Q_2)])^2}{2r(N-1)(k-r+1) + NV[tr(Q_2)]}$

Anexo 3. Metodología de cointegración panel de Groen y Kleibergen (GK).

Como se mencionó en el texto principal, la ecuación (12) en forma compacta se puede escribir como:

$$\Delta Y = Y_{-1}\Pi + W\Gamma + U \quad U \xrightarrow{iid} N(0, \Omega) \quad (A3.1)$$

con:

$$\Delta Y = \begin{bmatrix} \Delta Y'_1 \\ \vdots \\ \Delta Y'_T \end{bmatrix}; \quad Y_{-1} = \begin{bmatrix} Y'_0 \\ \vdots \\ Y'_{T-1} \end{bmatrix}; \quad \Pi = \begin{bmatrix} \Pi'_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \Pi'_N \end{bmatrix};$$

²⁵ El valor esperado y la varianza de las trazas de Q_1 y Q_2 son tomados de los resultados especificados en (A2.1) y (A2.2) para los componentes determinísticos correspondientes; r es al rango de cointegración, N es número de individuos y k corresponde al número de variables incluidas en el modelo.

$$W = \begin{bmatrix} W_1' \\ \vdots \\ W_T' \end{bmatrix}; \quad \Gamma = \begin{bmatrix} \Gamma_1' & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & \Gamma_N' \end{bmatrix}; \quad U = \begin{bmatrix} \varepsilon_1' \\ \vdots \\ \varepsilon_T' \end{bmatrix}$$

En este caso, el logaritmo de la función de verosimilitud es el siguiente²⁶:

$$l(\Pi, \Gamma, \Omega) = -\frac{NKT}{2} \ln(2\pi) - \frac{T}{2} \ln|\Omega| - \frac{1}{2} \text{vec}(U)' (\Omega^{-1} \otimes I_T) \text{vec}(U) \quad (\text{A3.2})$$

GK muestran que el estimador *ML* es equivalente en términos asintóticos a un estimador *GMM* con la siguiente función objetivo:

$$G = \text{vec}(\Delta Y - Y_{-1}\Pi)' M_W (I_{Nk} \otimes Y_{-1}) [(I_{Nk} \otimes Y_{-1})' M_W (I_{Nk} \otimes Y_{-1})]^{-1} (I_{Nk} \otimes Y_{-1})' M_W \text{vec}(\Delta Y - Y_{-1}\Pi) \quad (\text{A3.3})$$

donde²⁷:

$$M_W = (\Omega^{-1} \otimes I_T) - (\Omega^{-1} \otimes W) H_{SURE} [H_{SURE}' (\Omega^{-1} \otimes W'W) H_{SURE}]^{-1} H_{SURE}' (\Omega^{-1} \otimes W')$$

$$H_{SURE} = [(e_1 \otimes I_k) \otimes H_1, \dots, (e_N \otimes I_k) \otimes H_N]$$

$$H_1 = \left[\begin{array}{c} \left(\begin{array}{cc} 0_{q_i} \otimes 0'_{p_i} \\ I_{p_i} \end{array} \right) \\ \left(0_{P-p_i-q_i} \otimes 0'_{p_i} \right) \end{array} \right] \otimes I_k; \quad i = 1, \dots, N; \quad 0_K = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{k \times 1}$$

$$q_1 = 0; \quad q_i = (N-1)p; \quad P = Np$$

y e_i es el i -ésimo vector unitario de dimensión N .

Para el modelo B_2 con $\delta_{2,i}=0$, especificado en la Tabla 2²⁸, *GK* sugieren como método de estimación el procedimiento iterativo que se describe a continuación:

²⁶ En la ecuación (A3.2), el símbolo \otimes denota el producto Kronecker. El operador *vec* transforma una matriz en un vector, apilando las columnas de una matriz una debajo de otra.

²⁷ Adicionalmente, *GK* muestran que la expresión (A3.2) puede ser expresada en términos de la función objetivo (A3.3). La expresión (A3.3) corresponde a la forma general de la función objetivo, por lo tanto, la inclusión de componentes determinísticos puede variar las matrices de diseño de este modelo. Más adelante se muestran los cambios para el modelo B_2 de la Tabla 2 del texto.

²⁸ En el modelo B_2 la función objetivo *GMM* es igual a la especificada en (A3.3), en la cual se adiciona el componente determinístico correspondiente. Por lo tanto, se debe reemplazar la matriz de Y_{-1} por $Z_{-1} = [Y_{-1} \quad \iota]$, donde ι representa un vector de 1's de dimensión T . la matriz Π_{B_2} es la siguiente:

0: Construcción de estimaciones iniciales de Ω y α 's.

Un estimador inicial y consistente de Ω puede ser obtenido a partir de la siguiente expresión:

$$\hat{\Omega}^0 (\hat{\Pi}_{B_2}) = (\hat{\Omega}_{ij}^0)_{i,j=1,\dots,N}$$

donde $\hat{\Omega}_{ij}^0 = \frac{1}{T} \sum_{t=1}^T \hat{\eta}_{it} \hat{\eta}_{jt}'$ y $\{\hat{\eta}_{ij}\}_{t=1,\dots,T}$ representan los residuales del modelo *VEC* sin restringir para el individuo i , estimado por separado.

Partiendo de estas estimaciones por individuo en forma y estableciendo normalizaciones sobre los vectores de cointegración, también se pueden obtener estimaciones iniciales de $\alpha_1, \dots, \alpha_N$.

1. En esta etapa se construye un estimador para los parámetros β 's, el cual corresponde al argumento que minimiza la función objetivo *GMM*, condicionado a las estimaciones de la etapa anterior:

$$\begin{bmatrix} \text{vec} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\mu}_1 \\ \vdots \\ \hat{\beta}_N \\ \hat{\mu}_N \end{pmatrix} \end{bmatrix} = \left[\hat{E}'_{SURE} (I_{Nk} \otimes Z_{-1})' \hat{M}_W (I_{Nk} \otimes Z_{-1}) \hat{E}_{SURE} \right]^{-1} \hat{E}'_{SURE} (I_{Nk} \otimes Z_{-1})' \hat{M}_W \text{vec}(\Delta Y)$$

donde:

$$\hat{E}_{SURE} = \left[(e_1 \otimes \hat{\alpha}_1) \otimes \begin{pmatrix} e_1 \otimes I_k & 0_{Nk} \\ 0_k & 1 \end{pmatrix}, \dots, (e_N \otimes \hat{\alpha}_N) \otimes \begin{pmatrix} e_N \otimes I_k & 0_{Nk} \\ 0_k & 1 \end{pmatrix} \right]$$

$$Z_{-1} = [Y_{-1} \quad \iota_T]$$

ι_T es un vector de 1's de dimensión T y 0_A es un vector de ceros de dimensión A .

2. Construcción de un estimador de Ω dados los estimadores de α 's y β 's anteriores:

$$\hat{\Omega} = \frac{1}{T} (\Delta Y - Y_{-1} \Pi_{B_2} - W \hat{\Gamma})' (\Delta Y - Y_{-1} \Pi_{B_2} - W \hat{\Gamma})$$

donde $\hat{\Gamma}$ corresponde al estimador *OLS* de la regresión de $(\Delta Y - Y_{-1} \hat{\Pi}_{B_2})$ contra W , $\hat{\Pi}_{B_2}$ es una función de $\hat{\alpha}$'s y $\hat{\beta}$'s de las etapas anteriores.

$$\Pi_{B_2} = \begin{bmatrix} \alpha_1 \beta_1' & 0 & \dots & 0 & 0 \\ 0 & & \ddots & & 0 \\ 0 & 0 & \dots & 0 & \alpha_N \beta_N' \\ \mu_1 \alpha_1' & \mu_2 \alpha_2' & \dots & & \mu_N \alpha_N' \end{bmatrix}$$

con μ_i , $i = 1, \dots, N$ un vector fila de dimensión r .

3. Construcción de un estimador de los parámetros α 's, el cual es obtenido al minimizar la función objetivo *GMM* condicionada a los estimadores anteriores de β 's y Ω :

$$\begin{bmatrix} \text{vec}(\hat{\alpha}_1) \\ \vdots \\ \text{vec}(\hat{\alpha}_N) \end{bmatrix} = \left[\hat{\phi}'_{B,SURE} (I_{Nk} \otimes Z_{-1})' \hat{M}_W (I_{Nk} \otimes Z_{-1}) \hat{\phi}_{B,SURE} \right]^{-1} \hat{\phi}'_{SURE} (I_{Nk} \otimes Z_{-1})' \hat{M}_W \text{vec}(\Delta Y)$$

donde:

$$\hat{\phi}_{B,SURE} = \left[(e_1 \otimes I_k) \otimes \left(\begin{pmatrix} e_1 \otimes I_K & \mathbf{0}_{Nk} \\ \mathbf{0}'_k & 1 \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\mu}_1 \end{pmatrix} \right), \dots, (e_N \otimes I_k) \otimes \left(\begin{pmatrix} e_N \otimes I_k & \mathbf{0}_{Nk} \\ \mathbf{0}'_K & 1 \end{pmatrix} \begin{pmatrix} \hat{\beta}_N \\ \hat{\mu}_N \end{pmatrix} \right) \right]$$

4. Cuando la función objetivo y las estimaciones no han convergido se retorna a la etapa 1.

Para los otros tipos de modelos especificados en la Tabla 2 del texto, también se utiliza el anterior procedimiento iterativo de estimación. Sin embargo, las matrices de diseño y las matrices E_{SURE} y $\phi_{B,SURE}$ deben ser reemplazadas por unas apropiadas para estos casos.