



SUPERINTENDENCIA  
VALORES Y SEGUROS

## **Serie Documentos de Trabajo**

Superintendencia de Valores y Seguros  
Santiago – Chile

Documento de Trabajo N° 11

### **Un nuevo índice de precios Para rentas vitalicias**

**Salvador Valdés - Nicolás Castro - Arístides Torche**

**Abril 2010**



SUPERINTENDENCIA  
VALORES Y SEGUROS

La Serie Documentos de Trabajo publicada por la Superintendencia de Valores y Seguros tiene como propósito difundir trabajos de investigación aplicada, desarrollados por profesionales de esta institución o delegados a investigadores externos, académicos y participantes del mercado. De esta manera, nos comprometemos a abrir un espacio para la discusión académica de temas relevantes para el desarrollo y perfeccionamiento de los mercados de Valores y Seguros.

Los trabajos presentados en esta serie corresponden a versiones en progreso, donde serán bienvenidos comentarios adicionales. Toda la información contenida en éstos, así como su análisis y conclusiones, es de exclusiva responsabilidad de su (s) autor (es) y no reflejan necesariamente la opinión de la Superintendencia de Valores y Seguros.

The main objective of the Working Paper Series published by the Superintendencia of Securities and Insurance is to share applied research studies, conducted by our staff or entrusted to outside researchers, with scholars and market participants. Thus, we are committed to open a space for academic discussion on relevant topics for the development and improvement of the securities and insurance markets.

The papers included in these series are work in progress and further comments are mostly welcomed. All the information, as well as the analysis and conclusions of these papers, are exclusively those of the author(s) and do not necessarily reflect the opinion of the Superintendencia of Securities and Insurance.

**Documentos de Trabajo**  
**Superintendencia de Valores y Seguros, Santiago-Chile**  
**Avda. Lib. Bernardo O'Higgins 1449**  
[www.svs.cl](http://www.svs.cl)

# *Un nuevo índice de precios para rentas vitalicias<sup>1</sup>*

por  
Salvador Valdés<sup>2</sup>  
Nicolás Castro<sup>3</sup>  
Aristides Torche<sup>4</sup>

## Resumen

Este estudio presenta nuevos índices de precio para las rentas vitalicias previsionales. Se describe en detalle la metodología para su construcción y se muestra cómo implementarlos a través de una aplicación piloto para rentas vitalicias de vejez. La propuesta genera subíndices, que los asesores de los afiliados pueden usar directamente para buscar y comparar entre Compañías de Seguros de Vida. Los índices propuestos también generan información financiera útil para la gestión de los fondos de pensiones para afiliados activos mayores de 50 años de edad, quienes en su gran mayoría están en los multifondos D y E.

El estudio analiza en primer término las “tasas de interés promedio” que son las estadísticas sobre las rentas vitalicias publicadas desde 1988. Se muestra que dichas cifras presentan debilidades porque omiten el efecto de los cambios en la curva de rendimiento que preservan su nivel promedio y además tienden a generar una distorsión perceptual al sugerir que determinados cambios son pequeños al expresarse en forma porcentual cuando en realidad su impacto financiero es grande. Finalmente se indica como estas desventajas son superadas por el nuevo índice.

---

<sup>1</sup> La Superintendencia de Valores y Seguros de Chile (SVS) financió parte de este estudio, a través del contrato “Estudio para Analizar la Representatividad de las Tasas de Venta de Rentas Vitalicias Previsionales que Publica la SVS”, firmado el 21 de diciembre de 2009. La División de Estudios de la SVS entregó la base de datos sobre pólizas de rentas vitalicias utilizada aquí. La empresa *LVA Índices* nos dio acceso a sus datos sobre la curva de rendimientos estimada para cada día hábil del período en que dichas pólizas fueron contratadas. Los autores expresan su reconocimiento a las autoridades de ambas entidades. Las opiniones vertidas en este estudio son de exclusiva responsabilidad de los autores y no reflejan la opinión de la SVS ni de *LVA Índices*.

<sup>2</sup> Profesor Titular del Instituto de Economía, Pontificia U. Católica de Chile.

<sup>3</sup> Magister en Economía, del Instituto de Economía, Pontificia U. Católica de Chile.

<sup>4</sup> Profesor Titular del Instituto de Economía, Pontificia U. Católica de Chile.

## *Índice de Materias*

	<i>página</i>
<b>1. <i>Introducción</i></b>	<b>3</b>
<b>2. <i>Evaluación de la tasa media de las rentas vitalicias</i></b>	
2.1. Descripción de la información entregada hoy	
2.2. Problemas de la tasa de interés promedio.	
2.3. Cambios no paralelos en la curva de rendimientos	
2.4. Efectos composición	
<b>3. <i>Propuesta de un nuevo índice de precios para las rentas vitalicias</i></b>	
3.1. Conceptos base	
3.2. Criterios para definir celdas y subíndices	
<b>4. <i>Definición de la variable dependiente y Optimización de subgrupos por medio del método CART</i></b>	
4.1. Metodología para construir los índices a publicar	
<b>5. <i>Resultados de la implementación piloto del nuevo índice</i></b>	
<b>6. <i>Conclusiones y recomendaciones</i></b>	
<b>7. <i>Bibliografía</i></b>	
<b><i>Anexo 1: Descripción estadística del método CART</i></b>	
<b><i>Anexo 2: Estudio de la monotonidad de las variables prima, adelanto y atraso.</i></b>	
<b><i>Anexo 3: Grupos CART definitivos</i></b>	

## 1. Introducción

El estudio parte con un análisis de la utilidad y representatividad de las tasas de interés media de las rentas vitalicias previsionales, que desde 1988 reporta la SVS cada mes, en la sección Estadísticas de su página web. Se trata de un promedio mensual de la “tasa interna de retorno” de los contratos o pólizas de rentas vitalicias previsionales cerrados durante cada mes.<sup>5</sup>

Un problema conocido de la tasa de interna de retorno es que es una medida defectuosa de la sensibilidad del precio de una renta vitalicia frente a cambios en la curva de rendimiento (esta es una curva que contiene las tasas de interés de mercado a distintos plazos), cuando dichos cambios son en parte transitorios. Por ejemplo, si la tasa de interés de mercado revierte a una media fija, un aumento hoy en la tasa de interés de mercado genera un efecto mayor en el precio de las rentas vitalicias de menor duración que en aquellas de mayor duración. Esta diferencia no es recogida por la tasa interna de retorno.

Se evalúa empíricamente la representatividad de las cifras publicadas por la SVS, en esta dimensión.<sup>6</sup> Para ello se utiliza una base de datos sobre los contratos de renta vitalicia previsional de *vejez* anticipada y normal (este informe no analiza el caso de las pensiones de invalidez, ni las de orfandad). Se encuentra que los cambios en la forma de la curva de rendimiento han sido cuantitativamente importantes, y que el sesgo resultante ha impactado de modo más intenso a las pólizas de menor plazo que a las de mayor plazo. Ello ha creado con frecuencia situaciones donde el indicador que publica la SVS sube cuando debería bajar, y viceversa.

La tasa interna de retorno también hace aparecer como irrelevantes diferencias de pensión que son relevantes. Por ejemplo, pocas personas estimarían que la diferencia entre una tasa de 3,67% y otra de 3,88% es grande, pues la diferencia es de unos pocos decimales. Sin embargo, la pensión resultante puede diferir en 4%, lo cual para una pensión de \$300 mil, significa \$12.000 cada mes, posiblemente durante 15 años que es una esperanza de vida habitual al jubilar. Esa diferencia sumaría \$2,2 millones. La información actual minimiza las diferencias de precio a ojos del afiliado y asesores previsionales, lo cual reduce sus usos posibles.

A continuación se inicia el estudio de lo que es el objetivo fundamental del informe: la construcción de un indicador del precio de las rentas vitalicias previsionales, que enriquezca la información disponible para afiliados mayores de 50 años sin necesidad de pasar por el SCOMP, y que no esté sujeto a los problemas de la tasa interna de retorno.

---

<sup>5</sup> La SVS declara que es la “tasa anual de interés promedio implícita”.

<sup>6</sup> También evalúa si existe un efecto composición, pero se demuestra que por construcción el TIR está libre de él.

Sin embargo, proponer un nuevo índice de precios de las rentas vitalicias previsionales exige definir subíndices y ponderadores. Se presenta una discusión teórica profunda y se propone utilizar el método estadístico CART para la construcción de dichos índices.

Finalmente, el índice propuesto se aplica en forma piloto, a la base de datos de rentas vitalicias de vejez, con el fin de comprobar su factibilidad. Los resultados comprueban que los nuevos subíndices son factibles de implementar y se reportan en este estudio.

El nuevo índice es fácil de entender y de usar para proyectar pensiones, sin necesidad de pasar por el SCOMP. En efecto, su relación con la suma disponible para pensionarse, permite conocer el monto de la pensión que podría lograrse con un determinado ahorro (prima) o a la inversa, el ahorro total requerido para lograr una pensión de un monto determinado. Creemos que la entrega continua de estos nuevos índices enriquecería la información con la que cuentan asesores previsionales, afiliados y otros actores del mercado previsional.

Estos sub índices, también entregan información financiera útil para perfeccionar la gestión de los fondos de pensiones orientados hacia afiliados activos, pero mayores de 50 años de edad, quienes en su gran mayoría están en los multifondos D y E. En efecto, con la información que entregan, los fondos de pensiones podrían proyectar los precios futuros de las rentas vitalicias que deberán comprar sus afiliados, sin necesidad de pasar por el SCOMP a cada afiliado. Esa información permitiría ajustar las carteras de los fondos D y E para minimizar un riesgo financiero importante para esos afiliados: el riesgo de que el precio de compra de una renta vitalicia resulte alto justo cuando ellos deseen jubilar.

## **2. Evaluación de la tasa media de las rentas vitalicias**

### **2.1 Descripción de la información entregada hoy**

Desde 1988, la SVS exige a aquellas Compañías de Seguros de Vida (CSV) que venden rentas vitalicias previsionales, la entrega de un conjunto de datos para cada contrato individual. Esos datos incluyen la “prima” o precio de esa renta vitalicia, y la tasa interna de retorno (TIR) de ese contrato.

La SVS agrupa estos contratos en cinco familias:

1. “Vejez”, si la edad del causante al momento de firmar el contrato es igual o mayor que su edad legal de pensión (65 hombres, 60 mujeres),
2. “Vejez anticipada”, si la edad del causante al momento de firmar el contrato es menor que su edad legal de pensión,
3. “Invalidez parcial”, en el caso previsto en la ley (D.L. 3.500)

4. “Invalidez total”, en el caso previsto en la ley (D.L. 3.500)
5. “Sobrevivencia”, si la póliza fue contratada por un beneficiario de un afiliado activo fallecido. Esto incluye orfandad y otros casos menos frecuentes. Las pensiones de sobrevivencia contratadas por afiliados antes de fallecer se encuentran registradas en la categoría de vejez o de invalidez del causante.

Luego, para cada uno de estos cinco grupos, que indizamos por  $J = \{j: 1, 2 \dots 5\}$ , y para cada mes  $t$ , la SVS calcula y reporta la “tasa de interés media” ( $TIM_{jt}$ ), calculada a partir de los datos para los contratos individuales, indizados por  $i$ , de la siguiente forma:

$$TIM_{jt} \equiv \sum_{i \in j} \beta_{it} \cdot TIR_{it} \quad \text{donde } \beta_{it} \equiv \frac{Prima_{it}}{\sum_{i \in j} Prima_{it}} \quad (1)$$

Se ve que en cada grupo  $j$ , la tasa de interés media es un promedio de tasas internas de retorno de los contratos  $i$  pertenecientes a dicho grupo, ponderado según el tamaño de las primas de cada uno de dichos contratos individuales  $i$ .

Por último, la SVS construye la tasa de interés media *general* para el mes  $t$ , combinando las tasas de interés medias para estos cinco grupos. Usa las siguientes fórmulas:

$$TIMG_t \equiv \sum_{j \in J} \alpha_{jt} \cdot TIM_{jt} \quad \text{donde } \alpha_{jt} \equiv \frac{\sum_{i \in j} Prima_{it}}{\sum_{j \in J} \sum_{i \in j} Prima_{it}} \quad (2)$$

Como se observa, en esta etapa final el promedio también está ponderado según el tamaño de las primas en cada grupo.

Hasta ahora la SVS no ha publicado indicadores de tasa de interés promedio para otras agrupaciones posibles, algunas de las cuales tienen interés para los asesores previsionales, como el período de pago garantizado y el número de años de adelantamiento respecto a la edad de vejez normal. También podría interesar agrupar por género y tamaño de prima.

## 2.2 Problemas de la tasa de interés promedio.

La tasa de interés promedio puede introducir distorsiones perceptuales. Por ejemplo, pocas personas creerían que la diferencia entre una tasa TIR de 3,61% y otra de 3,88% es grande y relevante, pues la diferencia se mide en decimales y se ve pequeña. En este ejemplo, es apenas 0,27%, es decir 0,0027.

Sin embargo, la pensión resultante puede diferir en 4%<sup>7</sup>. Esto significa, para una pensión de \$300 mil al mes, una diferencia de \$12.000 cada mes durante 15 años (entre los 65 y los 80 años de edad, suponiendo que fallece a esta última edad). Esta diferencia

<sup>7</sup> Si la renta vitalicia tiene una “duración” de 15 años, entonces el efecto en el precio y en la pensión sería  $15 \times 0,27\% = 4,05\%$ .

suma \$2,2 millones sin descontar. La cifra de \$2,2 millones se aprecia como grande y significativa, en contraste con la cifra de 0,0027. Es decir, el uso de decimales pequeños en la tasa de interés promedio puede inducir a percibir como nimias a diferencias que son altamente relevantes en la realidad.

Expresando el punto en forma más general, la presentación como tasa de interés anualizada sobre UF maximiza la abstracción y reduce el universo de asesores y afiliados que comprende y atribuye realidad a las estadísticas que publicadas.

Desde el año 2004 existe en Chile el Sistema de Consultas y Ofertas de Montos de Pensión (SCOMP), cuya función es transmitir solicitudes y ofertas de montos de pensión, y entregar comparaciones entre las distintas modalidades de pensión que puede adquirir cada afiliado individual. Por ley, cualquier afiliado al sistema previsional que desee pensionarse debe obtener en el SCOMP una primera ronda de ofertas o montos de pensión que recibiría con distintas Compañías.

La información entregada por el SCOMP compite con la información entregada por la estadística de tasas de interés promedio, por la atención de afiliados y asesores previsionales. Se observa que la tasa de interés promedio es más difícil de interpretar que el SCOMP, pues éste último consiste en montos de pensión (\$ por mes).

También es poco probable que las Compañías de Seguros de Vida utilicen las estadísticas de tasa de interés promedio. En efecto, por el lado comercial, las Compañías no utilizarían la información de la tasa de interés promedio para evaluar sus estrategias de venta, pues saben que ella no recoge los cambios en la forma de la curva de rendimientos. Por el lado de la tarificación, las Compañías tampoco la usan pues requieren tomar en cuenta variables futuras o actuales, como el nivel de calce, la rentabilidad exigida por los accionistas, la estructura de los contratos que se están vendiendo (en términos de duración, flujos, etc.) y la proyección de la rentabilidad futura de sus inversiones. En cambio, la tasa de interés media constituye información pasada, que es obsoleta en comparación a la usada, y es demasiado agregada.

Podría pensarse que las estadísticas actuales serían útiles para estudios regulatorios, al constituir un indicador indirecto de los cambios y las tendencias en los precios de las rentas vitalicias previsionales. Sin embargo, la tasa de interés promedio no mide correctamente los cambios de precio ni sus tendencias en el tiempo, como demuestra la subsección siguiente.

Por último, debemos referirnos a la ponderación por tamaño de primas, que subyace a la estadística actual. La ponderación por tamaño de prima no es de interés directo para ningún afiliado en particular, pues cada uno compra *una sola* renta vitalicia. La ponderación según tamaño de primas podría tener algún interés para las Compañías de Seguros de Vida (CSV), pues ellas deben constituir reservas matemáticas en proporción al volumen en dinero de las primas, y no en proporción al número de pólizas. Sin

embargo, las CSV usan modelos mucho más afinados, de su propiedad, para definir sus políticas de reservas.

### 2.3 Cambios no paralelos en la curva de rendimientos

La tasa de interés media está basada en el concepto financiero “tasa interna de retorno” (TIR), aplicado a los flujos de caja esperados de cada contrato de renta vitalicia, como indica la ecuación (1).<sup>8</sup> La tasa interna de retorno se define como aquella tasa de descuento *hipotética y constante según plazo*, que igualaría el valor presente esperado de los flujos de caja con el precio de la renta vitalicia. Es decir,

$$TIR_{it} \equiv \text{raíz real de } 0 = -PRIMA_{it} + \sum_{h=0}^{12 \cdot (110 - \text{edad}_{i,t})} \frac{s_{h+t,i} \cdot p_{h+t,i}}{(1 + TIR_i)^h} + CM \cdot \sum_{h=0}^{12 \cdot (110 - \text{edad}_{i,t})} \frac{s_{h+t,i} \cdot f_{h+t,h+t+1}^i}{(1 + TIR_i)^h} \quad (3)$$

Donde:  $h$  = plazo contado a partir de la fecha  $t$ ;  $s_{h+t,i}$  = probabilidad del beneficiario  $i$  — dadas todas sus características, en particular,  $\text{edad}_{i,t}$ , su edad al momento  $t$  — de sobrevivir hasta la fecha  $t+h$ , dado que estaba vivo en la fecha  $t$ ;  $f_{h+t,h+t+1}^i$  = probabilidad del beneficiario del contrato  $i$  firmado en  $t$  de fallecer entre la fecha  $h+t$  y  $h+t+1$ , condicional a que está vivo en  $h+t$ .<sup>9</sup> Por convención, las tablas de mortalidad se calibran para cumplir que  $s_{110,i} = 0$ , es decir para que la probabilidad de sobrepasar los 110 años de edad sea cero;  $p_{h+t,i}$  = pensión pagadera por el contrato  $i$  en la fecha  $t+h$ ;  $CM$  = cuota mortuoria para el causante (15 UF).

El problema radica en los adjetivos “*hipotética y constante según plazo*”. Las tasas de descuento observadas realmente en el mercado financiero en la fecha  $t$ , son diferentes para distintos horizontes  $h$ , una vez eliminado el riesgo de crédito. Esta variación según el horizonte  $h$  de pago da origen a la “curva de rendimientos libre de riesgo (de crédito)”. Esta curva es un registro del rendimiento de mercado en función del plazo de la inversión. La curva de rendimientos está dada por los rendimientos medios (*yields*) según plazo,  $y_t(h)$ , definidos así:<sup>10</sup>

$$y_t(h) \equiv \left( P_t^0(h) \right)^{1/h} - 1 \quad (4)$$

donde  $P_t^0(h)$  es el precio de un bono seguro de valor de \$1, cuyo único pago ocurre  $h$  años después de  $t$ . Es clave insistir en que los  $y_t(h)$  no son hipotéticos, sino objetivos, obtenidos del mercado financiero, donde ocurren transacciones reales.

<sup>8</sup> La SVS declara que es la “tasa anual de interés promedio implícita”.

<sup>9</sup> Se puede demostrar que  $f_{h+t,h+t+1}^i = \frac{s_{t+h,i} - s_{t+h+1,i}}{s_{t+h,i}}$ .

<sup>10</sup> Ver Campbell, Lo y Mackinley (1997) *Financial Econometrics*, cap 10 y 11.

Empíricamente, se observa que la curva de rendimientos cambia a lo largo del tiempo. Es inusual que sea plana, y si lo es, ello no dura más que algunas semanas. Promediando en el tiempo, esta curva es creciente, con una asíntota finita. El valor del rendimiento  $y_i(h)$  es más estable para plazos largos, que para plazos cortos, porque el extremo corto está influenciado por la cambiante política monetaria del Banco Central y la situación de liquidez financiera. Los bancos centrales prefieren intervenir en plazos cortos con el fin de minimizar los efectos de la política monetaria sobre las decisiones de inversión de las empresas, cuyo plazo está dado por la durabilidad de los activos productivos, que tiende a ser largo.

Es sabido que la tasa de interna de retorno es una medida defectuosa de la sensibilidad del precio de una renta vitalicia frente a cambios en la curva de rendimiento cuando dichos cambios son “no paralelos”, o no permanentes. Por ejemplo, si la tasa de interés de mercado de corto plazo revierte a una media fija y estable en el tiempo, entonces un aumento hoy en la tasa de interés de mercado de corto plazo, que no cambia el nivel esperado para el largo plazo de dicha tasa de interés, eleva la curva de rendimiento en forma transitoria (esta es una modalidad de *no paralelismo*). En este ejemplo, que es realista, un aumento hoy en la tasa de interés de mercado de corto plazo genera un efecto mayor en el precio de las rentas vitalicias de menor duración que en aquellas de mayor duración. Este efecto *no es recogido* por la TIR.

A continuación medimos la magnitud que ha tenido en Chile esta falencia de la TIR, para las pólizas de rentas vitalicias previsionales. Para ello usaremos los precios de las pólizas de manera de recoger la forma de la curva de rendimientos  $y_i(h)$  y sus variaciones.

La *magnitud* de la *pensión* mensual depende de la prima que paga el afiliado a la Compañía de Seguros de Vida (CSV) que elija, menos el costo de la cuota mortuoria que recibirá el afiliado, y del *precio* de la renta vitalicia *unitaria*.

$$Pensión_i \equiv \frac{PRIMA_i - VECM_i}{PRVU_i} \quad (5a)$$

donde la *Pensión* se mide en UF por mes, la *PRIMA* (precio de la renta vitalicia) y el *VECM* (valor presente de la cuota mortuoria) se miden en UF, y *PRVU* se mide en meses. Al invertir esta ecuación se define el “precio de la renta vitalicia unitaria”:

$$PRVU_i \equiv \frac{PRIMA_i - VECM_i}{Pensión_i} \quad (\text{meses}) \quad (5b)$$

donde:

$Pensión_i$  = magnitud de la pensión mensual que el afiliado  $i$  habría obtenido si hubiera usado su ahorro previsional acumulado para adquirir una renta vitalicia en el mes  $t$ . Este es un valor objetivo observado en el último mes con información disponible, que llamamos  $t$ . Es un valor histórico, que sólo registra las realizaciones de procesos estocásticos.

$PRIMA_i$  = precio total de la renta vitalicia, adquirida en el mes  $t$  por el afiliado a una Cía. de Seguros de Vida. En un caso normal se cumpliría que  $PRIMA_{it} \equiv VC_t \cdot N_{it}$ , donde  $VC_t$  es el valor cuota en el mes  $t$  del multifondo donde el afiliado mantenía sus ahorros, y  $N_{it}$  es el número de cuotas que ha ahorrado el afiliado  $i$  hasta la fecha  $t$  de la compra. En otros casos se agregarían los recursos que el afiliado obtiene de un Bono de Reconocimiento por cotizaciones realizadas al sistema antiguo y el ahorro previsional voluntario (APV).

$VECM_{it}$  = valor esperado de la “cuota mortuoria” legal, descontado al presente. El monto de cuota mortuoria que fija la ley es de 15 UF, pagadas solamente al causante.

$PRVU_{it}$  = “precio” o “prima unitaria” por cada 1 UF de pensión vitalicia que podría haber adquirido el afiliado  $i$  en el mes  $t$ , excluyendo el beneficio de la cuota mortuoria

El precio  $PRVU_t$  tiene dos propiedades atractivas:

- 1) Es fácil de entender y de usar: cuando se usa para dividir cualquier suma que se destine a adquirir una pensión, entrega la pensión mensual que se podría haber obtenido en  $t$  de haber comprado una renta vitalicia.
- 2) Recoge la forma de la curva de rendimientos  $y_t(h)$  de mercado. Esto ocurre simplemente por tratarse de los valores efectivos de la  $PRIMA_i$  y de la  $Pensión_i$  expresados a precios de mercado: ambos valores están influenciados por los precios efectivos de los bonos a distintos plazos. En esta parte del cálculo de  $PRVU_{it}$  no interfieren hipótesis como que la curva de rendimiento podría haber sido plana.

Debe destacarse que el precio  $PRVU_{it}$  es medido directamente para cada póliza a partir de los datos  $PRIMA_i$  y  $Pensión_i$ . Como ambos son medidos **sin error**,  $PRVU_{it}$  no tiene fuentes de error originados en estos ítems.

El único número adicional requerido es  $VECM_{it}$ . Empíricamente,  $VECM_t$  varía entre 11 y 7 UF.<sup>11</sup> Como el valor mediano de  $PRIMA_t$  es de 1.900 UF en el universo de los contratos analizados para 2005-2008, se deduce que  $VECM_t$  tiene una baja incidencia, inferior al 0,5% en el caso mediano. Así, un error de 1% en la medición  $VECM_t$  tiene una influencia de 0,05% sobre el valor del  $PRVU_{it}$  obtenido a partir de la ecuación (5b). Luego, no se debe sobredimensionar los errores identificados en la discusión que sigue.

En general, existen varios métodos para calcular  $VECM_{it}$ , y sus errores inciden algo sobre  $PRVU_{it}$ . Esos métodos difieren según el origen de la tabla de mortalidad utilizada para el mes en que se firmó cada contrato, y según el origen de la curva de rendimientos de los bonos libres de riesgo para el mes ya indicado. Algunas fuentes posibles de datos son:

<sup>11</sup> Este valor no depende del período de pago garantizado, ni del monto de la prima, ni del diferimiento entre contratación e inicio. Sólo depende de la tabla de mortalidad individual del causante, que depende de la edad y sexo al contratar, y de las tasas de interés (curva de rendimiento), que varían en el tiempo.

1. Curva de rendimientos de los bonos libres de riesgo. Un posible origen es la estimación privada de una o varias CSV, otra es la estimación proporcionada por alguna empresa especializada en proveer información sobre rendimientos de bonos (*LVA Índices*, Riskamerica, otras) que dependen de los modelos de proyección que son propiedad intelectual de dichas empresas. Una tercera fuente posible es una que no está protegida por derechos de propiedad intelectual, y que es 100% objetiva, aunque contiene sesgos, como es la propia tasa interna de retorno (*TIR*) del contrato  $i$ , informada por la CSV vendedora a la SVS. Este informe recomienda usar la tercera opción para calcular el valor presente de la cuota mortuoria, aunque tenga sesgos. Los motivos son, primero, que la incidencia de los errores en *VECM* es muy pequeña, del orden de 0,05%, por lo que el costo de ese error es bajo. Segundo, hay costos de que una estadística pública dependa de opiniones protegidas por propiedad intelectual, como ocurre en las otras alternativas. Esta dependencia sería también más cara de operar para el fisco, si la SVS tuviera que pagar algún honorario a uno o más proveedores de información sobre la curva de rendimientos.
2. Tabla de mortalidad. Aquí las fuentes posibles son la tabla oficial<sup>12</sup>, y una o más tablas entregadas por actuarios de algunas CSV. Estas últimas son opiniones privadas, guardadas celosamente. En el futuro podría surgir una opción intermedia dada por opiniones de empresas especializadas, obtenidas a partir de los precios de los nuevos “bonos de mortalidad” que JP Morgan empezó a vender “over the counter”, es decir en forma privada, hace pocos años en Gran Bretaña. Nuevamente, como no parece conveniente que una estadística pública dependa de opiniones protegidas por propiedad intelectual, se recomienda usar la tabla de mortalidad oficial.

En este informe aplicamos estas recomendaciones en forma parcial. Si bien usamos la tabla de mortalidad oficial, nuestros gráficos y cálculos usan la curva de rendimientos para el mes respectivo que obtuvimos a partir de los datos que nos proporcionó amablemente la empresa *LVA Índices*. Esta elección fue realizada por la mayor calidad de estos datos por sobre el *TIR* de cada contrato, porque para estos efectos *LVA Índices* nos dio acceso a esta información ejerciendo sus derechos de propiedad intelectual. Con todo, insistimos en recomendar a cualquier repartición pública que aplique las recomendaciones de nuevos índices contenidas en este informe, usar el *TIR* del respectivo contrato para valorar la cuota mortuoria. Desde luego, las cifras entregadas en este informe no permiten inferir las curvas de rendimientos que proporcionó por *LVA Índices*, para preservar la confidencialidad de esa información.

Ahora volvemos a la discusión de nuestro objetivo, que es medir el sesgo propio del cálculo de cualquier *TIR*, que es la imposición de una curva de rendimientos plana. Para cada contrato de la muestra construimos dos variables:  $PRVU_i$  y  $1/TIR_i$ . Luego, se

---

<sup>12</sup> Las tablas oficiales de Chile incluyen un factor de mejoramiento de la mortalidad desde 2007.

agruparon y promediaron estas variables para todas las pólizas firmadas en el mismo mes  $t$ , creando las series temporales  $\overline{PRVU}_t$  y  $\overline{1/TIR}_t$ .

Si las estadísticas que actualmente se publican sobre el  $\overline{TIR}_t$  recogieran en forma insesgada y completa los cambios en la forma de la curva de rendimientos, entonces la correlación entre  $\overline{PRVU}_t$  y  $\overline{1/TIR}_t$  debería ser cercana a 1. En efecto, si todos los cambios en la curva de rendimientos fueran cambios estrictamente paralelos según plazo, entonces cuando sube la tasa de interés a un plazo  $h$ , sube en igual cantidad a todos los demás plazos. Eso haría subir al  $TIR_i$  en la misma cantidad, con lo cual  $1/TIR_i$  (y por ende  $\overline{1/TIR}_t$ ) caerían correlativamente. Por el otro lado, un aumento de ese tipo en la tasa de interés permitiría a las CSV pagar una *Pensión* $_i$  mayor a cambio de la misma *PRIMA* $_i$ , hecho que haría caer proporcionalmente a  $PRVU_i$ , según indica la definición (5b). Por ello, frente a ese tipo de aumento en la tasa de interés, tanto  $\overline{PRVU}_t$  como  $\overline{1/TIR}_t$  deberían caer. Ocurre lo inverso frente a una caída en las tasas de interés. De esta forma, queda establecido que habría una correlación cercana a 1 entre  $\overline{PRVU}_t$  y  $\overline{1/TIR}_t$  si todos los cambios en la curva de rendimientos fueran estrictamente paralelos según plazo. La correlación no sería igual a 1 porque la relación entre ambas variables no necesariamente es lineal.

En cambio, si la curva de rendimientos presentara con frecuencia cambios *no paralelos*, podría ocurrir que en algunos meses,  $\overline{PRVU}_t$  aumente mientras  $\overline{1/TIR}_t$  cae, o viceversa. Con ello, la correlación entre  $\overline{PRVU}_t$  y  $\overline{1/TIR}_t$  podría ser bastante inferior a la unidad. Nuestros resultados empíricos para esta correlación están en el Cuadro 1:

*Cuadro N° 1: Correlación mensual entre  $\overline{PRVU}_t$  y  $\overline{1/TIR}_t$  para cuatro grupos de pólizas*

<i>Grupo de pólizas</i>	<i>Número de Pólizas</i>	<i>Correlación (<math>1/TIR_i; PRVU_i</math>)</i>
RV de Vejez Anticipada	15.298	0,82
RV de Vejez Normal	21.032	0,56
RV Individuales	16.429	0,50
RV Conjuntas	19.901	0,87

Total contratos: 36.330. Parte de las RV de Vejez Anticipada son Individuales y el resto son Conjuntas. Lo mismo ocurre con las RV de Vejez Normal.

Puede observarse que la correlación es muy inferior a 1 - apenas 0,50 y 0,56 - para aquellos grupos de rentas vitalicias de menor duración, o sea de menor madurez promedio, que son Vejez Normal e Individuales. En cambio, las pólizas de mayor duración, que son Vejez Anticipada y Conjuntas, exhiben correlaciones de 0,82 y 0,87, que son superiores, aunque todavía bastante inferiores a 1.

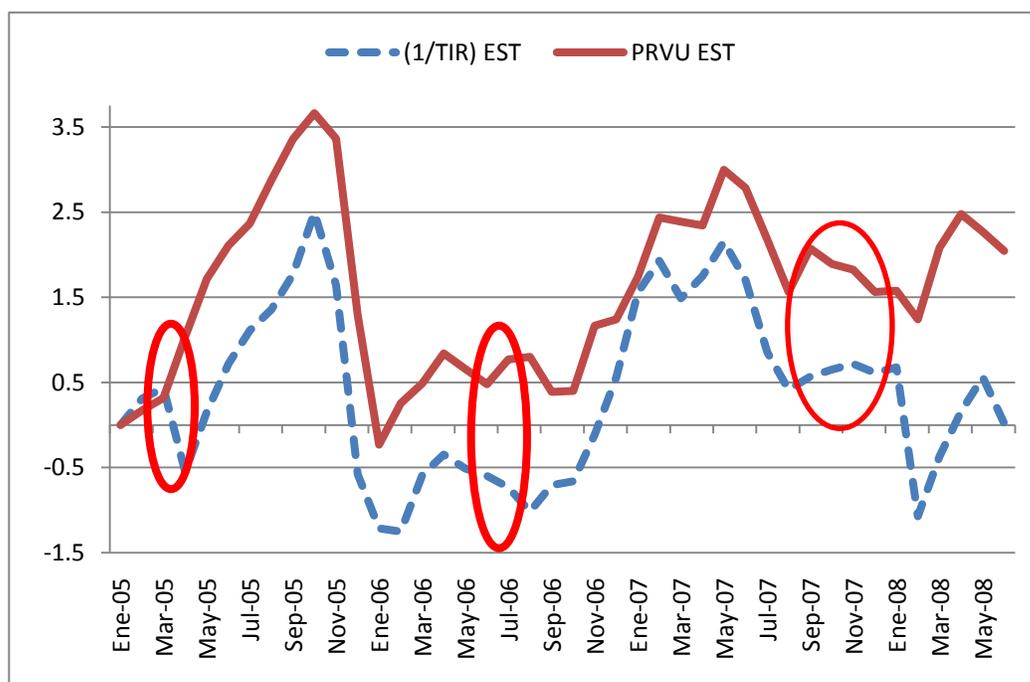
Este resultado empírico demuestra que en la realidad, la parte de la curva de rendimientos que más varía es aquella cercana al presente ( $h < 10$  años), mientras que la parte más lejana de la curva ( $h > 10$  años) varía mucho menos. Como se indicó, esta

diferencia se debe a que el Banco Central interviene con su política monetaria sobre las tasas de interés más cortas solamente.

Los siguientes gráficos destacan con un círculo aquellos meses en que  $\overline{PRVU}_t$  aumentó mientras  $\overline{1/TIR}_t$  cayó, o viceversa. Para simplificar, los gráficos se presentan solamente para los casos de rentas vitalicias de vejez normal y vejez anticipada, aunque al observar el comportamiento de las rentas vitalicias individuales y conjuntas los resultados son completamente análogos.

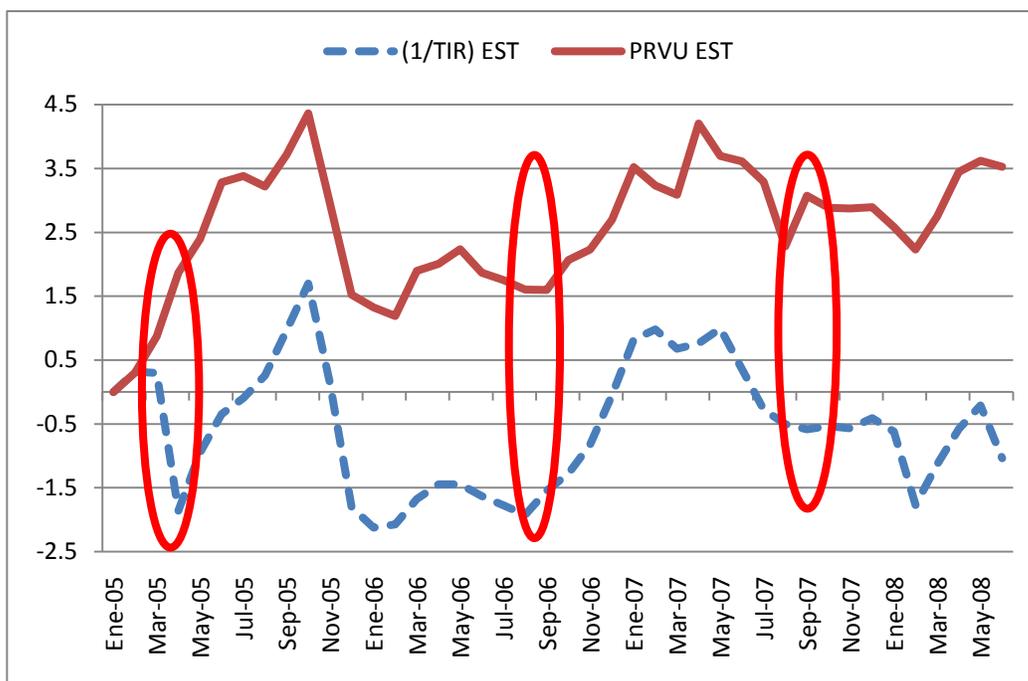
Los gráficos están “estandarizados” de la siguiente manera: primero, ambas curvas parten en el mismo nivel base o “cero”, con el fin de mostrar si se acumulan desviaciones en el tiempo o no; segundo, la unidad de medida es la desviación estándar de la serie, y por tanto el eje vertical mide el número de desviaciones estándar en que la variable se desvía en ese mes, respecto al valor promedio de la variable en el período completo.

Gráfico N° 1: Valores estandarizados de  $\overline{PRVU}_t$  y  $\overline{1/TIR}_t$  para Vejez Anticipada



Nota: Las series se estandarizaron con respecto al valor correspondiente de enero de 2005.

Gráfico N° 2: Valores estandarizados de  $\overline{PRVU}_t$  y  $\overline{1/TIR}_t$  para Vejez Normal



Nota: Las series se estandarizaron con respecto al valor correspondiente de enero de 2005.

La evidencia de los gráficos 1 y 2, sumada a la del Cuadro N° 1, es contundente: el índice publicado actualmente entrega información errónea en forma frecuente, en dos dimensiones:

- El nivel del índice actual se desvía en forma permanente, en algunos casos, del nivel del índice correcto; y
- A veces el índice actual sugiere que el precio de las rentas vitalicias está subiendo cuando en realidad está bajando, y viceversa (ver los meses marcados en círculos).

Estos errores son producto del conocido hecho de que las *TIR* no representan los precios reales del mercado financiero. En cambio, los *PRVU* sí los recogen, por definición. Estos errores pueden haber inducido a algunos asesores previsionales a realizar recomendaciones erradas.

## 2.4 Efecto composición

Se observa en las fórmulas (1) y (2) de página 5 que se usan en la actualidad, que los ponderadores  $\beta_{it}$  y  $\alpha_{jt}$  pueden variar a lo largo del tiempo, mes a mes. Ello ocurre en la práctica si cambia la composición de las nuevas pólizas contratadas. Este cambio abre la posibilidad de que la serie de tasa de interés promedio esté sesgada por el conocido

“efecto composición”. En esta subsección exploramos los factores que controlan la magnitud de este efecto. El tema merece una explicación detallada porque es importante para entender la propuesta de este estudio, presentada en la sección 3.

El “efecto composición” existe cuando se cumplen los siguientes tres requisitos:

- Los niveles o valores de los  $TIM_{jt}$  difieren entre los grupos  $j$ .
- Existe una tendencia temporal en la participación  $\alpha_{jt}$  del grupo  $j$ , sea al aumento o a la disminución (la tendencia temporal en la participación de los demás grupos, debe tener la dirección opuesta). Este requisito no se cumple si la participación  $\alpha_{jt}$  del grupo  $j$  es constante en el tiempo.
- Existe covarianza negativa entre la evolución en el tiempo de la  $TIM_{jt}$  y la tendencia temporal de la participación  $\alpha_{jt}$  de ese mismo grupo  $j$ .

Por ejemplo, si en todos y cada uno de los grupos  $j$  la  $TIM_{jt}$  aumenta, pero también cambia la composición entre grupos de manera tal de *aumentar* la participación  $\alpha_j$  de aquellos grupos con *menor nivel* de  $TIM_{jt}$ , entonces puede ocurrir que el promedio general  $TIMG_t$  *disminuya*, a pesar de que en todos y cada uno de los grupos la tasa de interés media *aumentó*. Si ello ocurriera, la variación en el índice reportado actualmente estaría entregando información errónea.

En general, existen dos métodos para evitar el efecto composición: evitar que se cumpla el requisito (a), o evitar que se cumpla el requisito (b).

Los índices de precios habituales, como el IPC, evitan que se cumpla el requisito (b). Para ello mantienen *fija en el tiempo* la “cantidad” de cada uno de los bienes considerados, según las cantidades observadas en un “período base”. Por ese medio, logran que los ponderadores  $\alpha_j$  queden fijos en el valor base  $\alpha_{j0}$ .

Otra forma de evitar el efecto composición es evitar que se cumpla el requisito (a). Para ello se debe redefinir la variable observada para asegurar que su nivel promedio difiera muy poco entre los grupos  $j$ . Mostramos ahora que la *TIR posee esta propiedad por construcción*.

Tomando la definición (3) de la *TIR*, y dividiendo a ambos lados por la pensión inicial (pagada al causante en el primer mes  $t$ ), que es  $p_{ti}$ , se obtiene:

$$TIR_{it} \equiv \text{raíz real de } 0 = - \left( \frac{PRIMA_t}{p_{ti}} \right) + \sum_{h=0}^{12 \cdot (110 - edad_{i,t})} \frac{S_{t+h,i} \cdot (P_{t+h,i} / P_{ti})}{(1 + TIR_i)^h} + \left( \frac{CM}{p_{ti}} \right) \cdot \sum_{h=0}^{12 \cdot (110 - edad_{i,t})} \frac{S_{t+h,i} \cdot f_{h+t,h+t+1}^i}{(1 + TIR_i)^h} \quad (3b)$$

La ecuación (3b) revela que la *TIR* controla por *dos* efectos de cantidad a la vez: primero, controla por el tamaño de la pensión inicial. Segundo, controla por el tipo de contrato, pues aquellas pólizas que entregan pagos por más periodos - por ejemplo, debido a que la pensión es anticipada y se inició a una edad más temprana - exhiben al

mismo tiempo dos efectos: (i) una mayor razón ( $PRIMA/p_{ii}$ ), y (ii) términos mayores en la primera sumatoria del lado derecho. Como estos dos ajustes tienden a compensarse, permiten que la  $TIR_i$  alcance un nivel similar para todos los tipos de póliza.

Como la  $TIR$  posee esta propiedad por construcción, que asegura el incumplimiento del requisito (a), en la práctica queda libre de los efectos composición.

Para terminar esta sección, se evalúa si ello es así en la práctica, para las pólizas de rentas vitalicias previsionales de la muestra de contratos. Se comparó la evolución en el tiempo de un índice de tasa de interés media ( $TIM$ ) calculado con ponderadores fijos en el tiempo, con la evolución de otro índice que utiliza ponderadores variables que siguen la cambiante composición de la muestra (como el  $TIMG_t$  usado actualmente, ver ecuación (2)).

En este ejercicio, medimos el efecto composición que afectó al índice “Vejez Anticipada” (“VA”) en el período 2005-2008. En cada ilustración, dividimos los contratos de vejez anticipada en dos grupos. En el primer intento, los dos grupos fueron aquellas pólizas suscritas con 5 ó menos años de anticipación a la edad legal de vejez normal ( $a = n^\circ$  años de anticipación), y las demás pólizas, que fueron suscritas con más de 5 años de anticipación ( $a > 5$ ). Para los hombres, estar en el primer grupo requiere haber suscrito el contrato antes de los 60 años de edad, y para las mujeres requiere haberlo hecho antes de los 55 años. Las fórmulas que definen el promedio general corregido son las siguientes:

$$(TIMG_t^{VA})_{corregido} \equiv \alpha_{0,a \leq 5} \cdot TIM_t^{VA,a \leq 5} + (1 - \alpha_{0,a \leq 5}) \cdot TIM_t^{VA,a > 5} \quad (6)$$

$$\text{donde} \quad \alpha_{0,a \leq 5} \equiv \frac{\sum_{i \in \{a \leq 5, VA\}} Prima_{i0}}{\sum_{i \in \{todas a, VA\}} Prima_{i0}} \quad (7)$$

$$TIM_t^{VA,a \leq 5} = \sum_{i \in \{a \leq 5, VA\}} \beta_{it} \cdot TIM_{it} \quad \text{donde} \quad \beta_{it} \equiv \frac{Prima_{it}}{\sum_{i \in \{a \leq 5, VA\}} Prima_{it}} \quad (8)$$

La ecuación (7) muestra que los ponderadores  $\alpha$  son constantes en el tiempo, pues se fijan en un período base, que fue elegido como el mes en que se inicia nuestra muestra de datos, enero de 2005.

Con todo, los ponderadores  $\beta$  siguen siendo variables en el tiempo, como muestra la ecuación (8). Por ello, el procedimiento mostrado aquí sólo elimina la primera “capa” del efecto composición. Éste podría seguir presente al interior de los subgrupos definidos según el número de años de anticipación de la pensión. A medida que se profundiza en subgrupos cada vez más finos, la última “capa” de subgrupos continúa expuesta al efecto composición.

Sin embargo, si los subgrupos no son seleccionados al azar, sino son elegidos para maximizar la similaridad entre los contratos que quedan en un mismo subgrupo, la

situación es distinta: Esa misma similaridad impide que se cumpla el requisito (a) para la existencia del efecto composición, que es que “los niveles o valores de la variable difieren entre los grupos  $j$ ”. Así, el efecto composición sería eliminado por completo si se subdividiera en subgrupos suficientemente homogéneos en su interior respecto de la variable medida. Esta es la idea central del proceso realizado en la sección 3.

Al aplicar a la base de datos de contratos de vejez anticipada la división entre grupos con más o menos de cinco años de anticipación, este estudio encontró que la diferencia entre el promedio corregido y el promedio con ponderador variable (utilizado actualmente) alcanza a sólo 1 punto base, es decir 0,01% anual en unidades de *TIR*. Probamos también con otras dos divisiones entre dos grupos y la diferencia fue muy pequeña, también. Esto confirma que el efecto composición es ínfimo para el caso de la *TIR*.

Como se comentó, este resultado se explica por el hecho de que, por construcción, la *TIR* es similar entre distintos tipos de contrato. Esto impide que el efecto composición opere, pues se incumple el requisito (a) para que exista.

#### *Conclusión de la Evaluación de la Tasa Media de las Rentas Vitalicias*

Hemos encontrado que las críticas a la tasa de interés promedio (*TIM*), indicador que se publica actualmente, son de peso. Primero, la tasa media tiene poca utilidad para los distintos integrantes del mercado. Segundo, la *TIM* impone artificialmente la hipótesis de que los cambios en la curva de rendimiento son paralelos, pero la evidencia revela que ocurren numerosos episodios de cambios no paralelos. Ello puede haber inducido a los afiliados y a los asesores previsionales a tomar decisiones erradas. También reducen la validez de los estudios académicos y regulatorios que utilicen esta estadística. Estas críticas justifican explorar un perfeccionamiento a los índices que publica la SVS.

Por otra parte, se reconoció que la *TIR*, que es la base de la *TIM*, evita los sesgos del efecto composición. Por ello, para superar a la *TIM* es necesario que el nuevo índice que se proponga posea también el atributo de estar libre del efecto composición. Como se explicó, esto exige elegir los grupos  $j$  de manera de maximizar la similitud entre el valor del indicador de los contratos que quedan en un mismo subgrupo. Esta es la idea central del proceso realizado en la sección 3.

La ganancia de pasar a un nuevo indicador provendrá de entregar información correcta cada vez que la curva de rendimientos cambie de forma no paralela.

### 3. Propuesta de un nuevo índice de precios para las rentas vitalicias

#### 3.1 Conceptos base

El afiliado activo que proyecta jubilarse dentro de algunos años no accede todavía al SCOMP. Sin embargo, está interesado en proyectar la *magnitud* de la *pensión* mensual futura. Ella dependerá de cuánto haya acumulado, es decir de la prima que esté dispuesto a pagar a la CSV, menos el costo en términos del valor presente esperado de la cuota mortuoria que corresponda. Depende también del *precio* de la renta vitalicia *unitaria* que proyecta para el futuro.

$$PRVU_i \equiv \frac{PRIMA_i - VECM_i}{Pensión_i} \quad (\text{PRVU medido en meses})^{13} \quad (5b)$$

donde:

$Pensión_i$  = magnitud de la pensión mensual que el afiliado  $i$  habría obtenido en el mes  $t$  si hubiera usado cierto stock de ahorro para adquirir una renta vitalicia.

$PRIMA_i$  = precio total de la renta vitalicia, adquirida en el mes  $t$  por el afiliado  $i$  a una Cía. de Seguros de Vida.

$VECM_{it}$  = valor esperado de la “cuota mortuoria” legal, descontado al presente. El monto de cuota mortuoria que fija la ley es de 15 UF, pagadas solamente al causante.

$PRVU_{it}$  = “precio” o “prima unitaria” por cada 1 UF de pensión vitalicia que podría haber adquirido el afiliado  $i$  en el mes  $t$ , excluyendo el beneficio de la cuota mortuoria. Se mide en meses.

Otra forma de explicar  $PRVU_{it}$  parte por definir el “valor actual neto” (VAN) que obtiene el afiliado al contratar una póliza de renta vitalicia, pero usando los rendimientos de mercado para descontar los flujos, no una tasa hipotética:

$$VAN_{it} \equiv -PRIMA_{it} + \sum_{h=0}^{12 \cdot (110 - edad_{i,t})} \frac{S_{t+h,i} \cdot P_{t+h,i}}{(1 + y_t(h))^h} + CM \cdot \sum_{h=0}^{12 \cdot (110 - edad_{i,t})} \frac{S_{t+h,i} \cdot f_{h+t,h+t+1}^i}{(1 + y_t(h))^h} \quad (9a)$$

Cuando dividimos esta ecuación por la pensión inicial,  $p_{ti}$ , se obtiene:

$$\frac{VAN_{it}}{P_{ti}} \equiv -\left(\frac{PRIMA_{it}}{P_{ti}}\right) + \sum_{h=0}^{12 \cdot (110 - edad_{i,t})} \frac{S_{t+h,i} \cdot (P_{t+h,i} / P_{ti})}{(1 + y_t(h))^h} + \left(\frac{CM}{P_{ti}}\right) \cdot \sum_{h=0}^{12 \cdot (110 - edad_{i,t})} \frac{S_{t+h,i} \cdot f_{h+t,h+t+1}^i}{(1 + y_t(h))^h} \quad (9b)$$

<sup>13</sup> Recordemos que el valor que tuvo este precio en el pasado se obtiene directamente a partir de los datos de cada póliza vendida en el pasado.

Es interesante destacar que la prima incluye el valor presente de las pensiones recibidas por el afiliado más el valor presente de su cuota mortuoria, es decir las dos sumatorias de la derecha de la igualdad y además los costos de administración, los impuestos y el margen de utilidad de la CSV (normal y sobrenormal). Por lo tanto el VAN será negativo y tan más negativo cuanto mayores sean dichos gastos. En otras palabras el VAN puede considerarse una medida con signo contrario de los costos de operación de la CSV

Reordenando el lado izquierdo para replicar la forma (5b), se obtiene:

$$(9b) \quad \left( \frac{PRIMA_i - CM \cdot \sum_{h=0}^{12 \cdot (110 - edad_{i,t})} \frac{s_{t+h,i} \cdot f_{h+t,h+t+1}^i}{(1 + y_t(h))^h}}{P_{ii}} \right) \equiv \sum_{h=0}^{12 \cdot (110 - edad_{i,t})} \frac{s_{t+h,i} \cdot (P_{t+hi} / P_{ii})}{(1 + y_t(h))^h} - \left( \frac{VAN_{it}}{P_{ii}} \right) \equiv PRVU_{it}$$

El lado derecho muestra que el  $PRVU_{it}$  también puede ser definido como el valor presente esperado de los flujos de caja a recibir por el afiliado, menos el VAN del contrato para el afiliado, todo por unidad de pensión inicial. En otros términos y como nos dice la intuición, cuando los costos de administración, los impuestos y el margen de utilidad de la CSV (normal y sobrenormal) hacen que el VAN sea más negativo, es decir que el afiliado reciba menos dinero de lo que paga en esperanza, entonces el  $PRVU_{it}$  aumenta y la renta vitalicia es “más cara”.

El precio  $PRVU_t$  tiene dos propiedades atractivas:

- 1) Es fácil de entender y de usar para proyectar pensiones, sin necesidad de pasar por el SCOMP: entrega la pensión que se podría haber obtenido, cuando se usa para dividir cualquier suma que se destine a adquirir una pensión. Por ejemplo, si un afiliado piensa ahorrar \$44 millones netos del valor de la cuota mortuoria que es muy pequeño (inferior a 11 UF en valor presente) y proyecta que el  $PRVU_t$  para su tipo de póliza será 183 meses, entonces su pensión mensual proyectada es  $\$44.000.000/183 = \$240.437$  /mes mientras viva. Los asesores previsionales apreciarían acceder a un precio  $PRVU_t$  que sea fácil de explicar al afiliado. Esto a su vez facilita la comparación de precios entre CSV y ofrece una herramienta adicional potente para estos efectos.
- 2) El precio  $PRVU_t$  recoge correctamente la forma de la curva de rendimientos  $y_t(h)$  que estuvo vigente en el mercado en la fecha  $t$ . Es decir, evita hacer supuestos que contradicen la evidencia, como que la curva de rendimientos es plana y tiene un valor fijo para todos los horizontes  $h$ .

Este estudio evalúa la factibilidad de que la SVS reporte índices de precios como el  $PRVU_t$ , pues es útil para la toma de decisiones del afiliado y sus asesores previsionales, y no contiene los sesgos ó errores de la tasa de interés media (basada en la TIR).

### 3.2 Criterios para definir subíndices o “celdas”

Se analizó anteriormente que para reducir el error de composición era necesario agrupar los contratos de rentas vitalicias en grupos que incluyeran aquellas rentas más homogéneas en términos del sus PRVU. Pero además es conveniente que estos grupos configuren agrupaciones útiles para las personas que están interesadas en contratar dichas rentas. Por estas razones, los criterios para definir los subíndices constituyen un punto importante en la construcción del índice propuesto.

Los criterios que se han considerado son los siguientes:

1. Criterio Institucional: En el pasado la SVS definió ciertos grupos o subíndices, para los que ha publicado la tasa de interés media (*TIM*) con frecuencia mensual desde el año 1988. El criterio continuidad institucional consiste en mantener los grupos definidos con anterioridad por la SVS.

Para el caso de la muestra de contratos utilizada en este estudio, este criterio entrega solamente *dos* grupos institucionales: pensión por vejez normal y pensión por vejez anticipada. Más adelante, si se adoptaran las recomendaciones de este estudio piloto, el criterio institucional podría agregar más grupos y subíndices, como los de pensión por invalidez parcial y por sobrevivencia. Una vez siniestrados, estos afiliados o beneficiarios tienen un “aporte adicional” de magnitud conocida, pero todavía deben elegir algunos aspectos de la pensión y elegir entre retiro programado y renta vitalicia, para lo cual puede ser útil contar con información de precios que no exija pasar por el SCOMP.<sup>14</sup>

2. Criterio de evitar los sesgos del efecto composición

Una vez que se ha aceptado la existencia de dos grupos mayores: rentas normales y rentas adelantadas, se deberán construir subíndices en cada caso con el objeto de proteger al precio  $PRVU_t$  de los efectos composición. La propuesta es protegerlo en dos niveles:

- a) Al interior de cada subíndice (grupos de contratos que se agrupan en una celda particular), lograr suficiente similitud entre contratos para evitar que se cumpla el requisito (a) para la existencia del efecto composición.
- b) Entre subíndices, evitar que se cumpla el requisito (b) para la existencia del efecto composición, que es que “Existe una tendencia temporal en la participación  $\alpha_{jt}$  del grupo  $j$ ”, por medio de fijar los ponderadores  $\alpha_{jt}$  en los valores  $\alpha_{j0}$  de un período base.

---

<sup>14</sup> En el caso de los beneficiarios de pensión por invalidez total, hay menos elecciones y por tanto menos necesidad de información.

Este criterio nos impulsa a agrupar los contratos de manera de *minimizar las diferencias al interior de los grupos*, y de maximizar las diferencias promedio entre grupos. Ello conducirá al uso del método CART.

### 3. Criterio de evitar sesgos por cambios estructurales conocidos

Una reforma legal aprobada en 2004 elevó gradualmente los requisitos legales para acceder a la pensión anticipada en Chile. Ese aumento gradual se materializó durante la mayor parte de la muestra de contratos que utiliza este informe (enero 2005 hasta junio 2008).

Proponemos el criterio de que el índice propuesto sea lo más representativo posible para el período *post*-muestral. Con ese fin, elegimos el período base en la parte *final* de nuestra muestra. En concreto, se proponen los ponderadores  $\alpha_{j0}$  que corresponden a los contratos informados en los últimos doce meses de la muestra, junio 2007 a junio 2008.

La reforma introducida por la ley 20.255, en cuanto obliga por primera vez a las mujeres a contratar pensiones de sobrevivencia para su cónyuge, genera un cambio estructural adicional. Este no es recogido en este informe, porque la base de datos disponible cuenta apenas con 30 contratos de ese tipo, número que no es representativo. Sugerimos incorporar este nuevo grupo al índice general de precios de rentas vitalicias sólo cuando se hayan acumulado al menos unos 1000 contratos del nuevo tipo.

### 4. Criterio de representatividad mensual de cada subíndice.

Uno de los objetivos clave es proporcionar información con periodicidad *mensual* (o más frecuente si fuera posible). Ello es indispensable para incorporar con celeridad los efectos de los cambios en la tasa de interés de mercado. Así, esta información será útil para los asesores previsionales y los afiliados.

La periodicidad mensual también es indispensable para que las series puedan ser utilizadas para optimizar las políticas de inversión de los fondos de pensiones orientados a servir afiliados activos mayores de 50 años y por tanto más próximos a iniciar una pensión de vejez, y para fondos que sirven a pensionados por retiro programado. Esto se discute en más profundidad en el próximo acápite.

Este criterio nos lleva a limitar la subdivisión en grupos, de modo que en los meses con menos contratos, cada subgrupo incluya al menos 10 contratos. Nos parece que esta es la cifra mínima para que la ley de los grandes números pueda operar al interior de cada subgrupo y así el promedio mensual de  $\overline{PRVU}_{jt}$  logre un grado adecuado de representatividad, en todos los subgrupos  $j$  y en todos los meses  $t$ .

Por otra parte, un exceso de información dificulta la toma de decisiones y conduce a la parálisis decisional, según indica la literatura reciente del comportamiento. Si hubiera cientos de subíndices, casi todos los asesores previsionales encontrarían excesivo el esfuerzo requerido para usarlos.

Además, si existiera una relación estrecha entre muchos de estos subíndices, el asesor ganaría muy poca información nueva accediendo a cientos de subíndices, por oposición a unos pocos subíndices representativos. Por eso, cuando existe suficiente similitud entre los *PRVU* de distintos tipos de póliza, es adecuado fusionarlas y representarlas con un solo índice combinado.

##### 5 . Criterio de utilización para la gestión financiera de los multifondos

Un índice de precios para las rentas vitalicias sería de gran utilidad para optimizar la gestión financiera de los fondos de pensiones que prestan servicio a los afiliados activos, cuya edad sugiere que están a menos de 15 años a contratar una pensión, y también para los afiliados ya pensionados por retiro programado<sup>15</sup>. Los gestores de fondos de pensiones requieren de información barata, que no exija pasar a sus afiliados por el SCOMP.

El afiliado consume una pensión mensual, no el saldo de su cuenta individual. Por ello, el afiliado se beneficia si la gestión financiera toma en cuenta las correlaciones de las variaciones futuras en el precio de la renta vitalicia unitaria, con las variaciones también futuras del valor cuota en su actual fondo de pensiones. Es posible y conveniente ajustar la selección de instrumentos financieros para favorecer aquellos cuyo precio futuro covaría en forma ventajosa con el precio futuro de la renta vitalicia (Valdés-Prieto y Castro, 2010).

La ecuación (5) indica que el precio de la renta vitalicia unitaria afecta en forma *multiplicativa* – y no aditiva – la magnitud de la pensión adquirida.

$$Pensión_i \equiv (PRIMA_i - VECM_i) \cdot \frac{1}{PRVU_i} \quad (5c)$$

Sin embargo, la representación multiplicativa dificulta el análisis financiero, si las variables aleatorias tienen una distribución que no cumple con la propiedad de que la combinación de dos variables aleatorias con igual distribución tenga

---

<sup>15</sup> Hay que recalcar el punto que el índice propuesto es útil para la gestión de los Administradores de Fondos de Pensiones y no para procesos de tarificación de contratos por parte de las Compañías de Seguros de Vida.

una distribución de esa misma familia.<sup>16</sup> Para recuperar una representación aditiva que facilite el análisis financiero, se requiere transformar la ecuación (5b) por medio de logaritmos:

$$\ln(\text{Pensión}_i) \equiv \ln(\text{PRIMA}_i - \text{VECM}_i) - \ln(\text{PRVU}_i) \quad (10)$$

Por esto, el criterio de que el índice de precios de rentas vitalicias pueda ser utilizado para la gestión financiera hace conveniente que el índice publicado pueda ser transformado a un costo moderado en un índice del logaritmo de dicho precio. Esa transformación sería realizada por las empresas que realizan la gestión financiera, es decir por cada AFP, no por la SVS. Esta transformación sólo puede tener costos bajos y ser oportuna si la *agrupación* de contratos en subíndices o grupos está hecha sobre la base del logaritmo del precio de la renta vitalicia unitaria de cada contrato.

#### 6. Criterio de destacar atributos de interés para definir la política pública.

Lo ideal sería que los subíndices propuestos también entreguen, información sobre atributos de interés para definir la política pública, en particular:

- a. El tamaño de la prima, que es un indicador de la riqueza laboral del individuo en el ciclo de vida. Puede ser útil determinar en qué medida el precio de las rentas vitalicias aumenta (o disminuye) con el ingreso en el ciclo de vida del afiliado. Una buena medida del ingreso en el ciclo de vida es la acumulación de ahorro previsional, que se manifiesta en el tamaño de la prima (Ver también el Anexo 2);
- b. Género del causante. Interesa saber si el precio unitario de las rentas vitalicias es mayor para las mujeres o para los hombres, controlando por la diferencia en expectativa de vida; y
- c. La presencia de un heredero obligatorio, como ocurre en las rentas vitalicias conjuntas, versus el caso de renta vitalicia individual, donde no es obligatorio dejar herencia.

Sin embargo, este criterio podría obligar a multiplicar muchas veces el número de subgrupos. Ello reduciría la representatividad mensual de algunos subíndices y dañaría las propiedades del índice en cuanto a los criterios 4 y 5. Nos parece preferible satisfacer el objetivo N° 6 de otro modo: que la SVS genere índices agregados separados cuando desee satisfacer las demandas de la política pública. La SVS está en condiciones de hacerlo porque posee información sobre el universo de pólizas firmadas en Chile.

---

<sup>16</sup> Esta es la propiedad “reproductiva”. Por ejemplo, la suma de normal también es normal, y el producto de lognormales también es lognormal. Sin embargo, la suma de lognormales NO ES lognormal, y tampoco normal, sino tiene una distribución de alta complejidad que impide el análisis financiero.

### 3.3 Definición de la variable dependiente y Optimización de subgrupos por medio del método CART

De acuerdo a los criterios indicados, el problema de agrupación consiste en identificar aquellos grupos de contratos que minimicen la variabilidad al interior de cada uno de los grupos y maximizan la diferencia promedio entre grupos. Esto debe hacerse sin que el número de contratos que queda en los grupos menos numerosos sea tan bajo como para comprometer la representatividad del promedio mensual del precio  $PRVU_t$ .

Una forma de abordar este problema de agrupación es partir desde aquel “grupo universo” que unifica todos los contratos, y luego *dividirlo* en subgrupos de acuerdo a algún criterio basado en el precio que nos interesa, que es el  $\ln(PRVU_t)$ . Este método puede ser calificado de *top-down* porque divide grupos formando subgrupos menores.

Existe un procedimiento estadístico reconocido para realizar esta división, llamado *Classification and Regression Trees* (CART), donde la parte que nos interesa es la de *Regression* (no aplicaremos la parte de *Classification*).<sup>17</sup>

#### *Definición de la variable dependiente*

La aplicación del método CART requiere previamente definir la variable dependiente. En nuestro caso se pueden considerar dos candidatos: le  $PRVU$  y su logaritmo. En el punto 5 de la sección precedente se mostró la conveniencia de la especificación logarítmica. Por otra parte la conveniencia de vincularlo con el Money Worth Ratio para el análisis de la gestión de fondos para afiliados mayores de 50 años también sugiere su especificación logarítmica. En efecto, el precio de una renta vitalicia unitaria  $PRVU$  es una combinación de dos aspectos:

- La cantidad de UFs esperadas de hoy que entrega esa póliza, ajustando por mortalidad y por tasas de interés. Una renta vitalicia unitaria puede contener cantidades distintas de UF esperadas de hoy, a pesar de pagar 1 UF al mes mientras viva el causante. En efecto, una renta vitalicia que paga 1 UF al mes paga más UFs esperadas de hoy *en total* cuando el afiliado contrata la pensión siendo más joven, porque paga por más tiempo. También paga más UFs esperadas de hoy en total si el causante tiene cónyuge, porque también pagará sobrevivencia a la cónyuge, o si la póliza tiene un período de pago garantizado positivo y así en varias otras dimensiones; y
- El precio por cada UF esperada de hoy que entrega la póliza.

Esta descomposición es lograda por el indicador conocido como *Money's Worth Ratio*. Este se define como:

---

<sup>17</sup> La referencia base es Breiman, Friedman, Olshen y Stone (1984) *Classification and Regression Trees* published by [Wadsworth Press](http://www-wadsworth.com). <http://www-stat.stanford.edu/software/cart/index.html>. Conviene operar CART con STATA, bajando de la web un módulo de acceso gratuito. Alguien desarrolló la rutina, y la puso bajo el nombre CART.

$$MWR_{it} \equiv E_t FCRV_i / (PRIMA_{it} - VECM_{it}) \quad (R.1)$$

donde se define al valor presente esperado de los flujos de caja totales como  $E_t FCRV_i \equiv E_t FCRVU_i \cdot Pensión_{it}$ , en que  $E_t FCRVU_i$  es el valor presente esperado de los flujos de caja de una pensión unitaria excluyendo la cuota mortuoria.

Recordando que ya se definió al precio total de la renta vitalicia como  $PRIMA_{it} \equiv PRVU_{it} \cdot Pensión_{it} + VECM_{it}$ , se deduce de estas definiciones que:

$$PRVU_{it} \equiv E_t FCRVU_i \times \frac{1}{MWR_{it}} \quad (R.2)$$

El precio de interés para el afiliado es el producto de los dos términos del lado derecho de la ecuación (R.2):

a)  $E_t FCRVU_i$  que es el valor presente esperado de los flujos de caja de una pensión “unitaria” que paga 1 UF al mes. Este varía de un mes a otro, siguiendo las variaciones de las tasas de interés de mercado y las tendencias de la tabla de mortalidad. Las variaciones de la tasa de interés dependen del nivel y de la forma de la curva de rendimientos de bonos. Estas variaciones pueden afectar en grado diferente a rentas vitalicias de diferente plazo promedio, lo cual depende de varios atributos del contrato. Por esto ese término tiene el subíndice  $i$ . La tabla de mortalidad exhibe una tendencia denominada “factor de mejoramiento”, que es estable a lo largo de los meses y varía sólo a nivel de décadas, excepto cuando ocurren importantes avances médicos o epidemias. La tabla de mortalidad depende de la edad del afiliado y de sus beneficiarios potenciales, y por ello también depende del contrato.

b)  $MWR_t$ , depende de la organización de la industria – en especial de los métodos de marketing de las CSV y de búsqueda de los afiliados-, de los costos de transacción e impuestos, de la carga efectiva de las regulaciones de solvencia y de otro tipo que deban cumplir las CSV, de la entrada de nuevas CSV, de las fusiones en la medida que afecten el grado de competencia entre ellas, y de la selección adversa no reflejada en las tablas de mortalidad para rentistas vitalicios.

Esta descomposición es valiosa para los afiliados mayores de 50 años. En efecto, el término  $E_t FCRVU_i$  covaría con el precio de los bonos y otros activos donde está invertido dicho fondo de pensiones. La covarianza indicada puede ser aprovechada por los gestores de fondos con el fin de reducir el riesgo de la pensión y de subir su valor promedio entre distintos estados posibles (Valdés-Prieto y Castro, 2010). En cambio, el  $MWR_t$  puede covariar de manera diferente con el precio de los activos donde está invertido el fondo de pensiones. Los datos sobre esta segunda covarianza pueden ser utilizados por el gestor de fondos para optimizar la cartera de los afiliados.

Para los efectos de este informe, la ecuación (R.2) revela que una eventual descomposición ulterior del precio de la renta vitalicia unitaria PRVU en sus factores causantes, entre ellos el *Money's Worth Ratio*, también es *multiplicativa*. Ello sólo puede ser tratado en forma sencilla pasando a una formulación aditiva, lo cual requiere pasar a logaritmos.

Por otro lado, la variable *PRVU* es más decidora y fácil de entender para los usuarios del índice. Ello facilita su masificación y uso más generalizado por parte de éstos, lo que sin duda es un objetivo muy importante de todo índice. Para satisfacer ambos criterios, nuestra propuesta es un compromiso entre el uso de la variable  $\ln(PRVU)$ , que es aquella más relevante desde el punto de vista financiero, pues permite la construcción de índices que pueden ser utilizados por los gestores de fondos sin necesidad de pasar por redefinir el índice desde cero, aplicando el método CART.

Al mismo tiempo, el índice precios de rentas vitalicias se expresará finalmente en términos de *PRVU* sin logaritmos, para facilitar su comprensión por parte de sus usuarios: asesores previsionales y adquirentes de rentas vitalicias.

Debe destacarse que la propiedad optimizante del CART de minimizar la heterogeneidad intra grupos (subíndices) se mantiene para transformaciones monótonas de la variable utilizada para definir el índice y en este caso *PRVU* es una transformación monótona del  $\ln PRVU$ .

#### *Descripción básica del método CART*

El método CART tiene por objeto agrupar ciertas unidades de análisis que en nuestro caso son los contratos individuales de rentas vitalicias de tal forma que la variabilidad de una variable objetivo sea mínima en cada uno de dichos grupos. En nuestro caso, la variable objetivo, llamada también variable indicadora o dependiente es el log de PRVU. Para realizar esta agrupación, se cuenta con información de varias otras variables de atributos de los contratos que estarían relacionadas con la variable objetivo. En nuestro caso dichas variables son: tasas de interés, prima pagada, período garantizado, período diferido, edad del cónyuge, género del comprador, y años de atraso o de adelanto de la pensión en relación a la edad legal para pensionarse.

En términos algebraicos, el método consiste en identificar cortes para las variables de atributos que, permitan definir celdas que agrupen aquellos contratos con la menor variabilidad posible para el log de los PRVU.

El CART es un método de análisis sistemático y secuencial de datos que analiza uno por uno todos los atributos disponibles y forma grupos dicotómicos con ellos a los que aplica un criterio de para definir el corte óptimo. Para precisar este punto supóngase un atributo  $a_1$  que tiene sólo tres valores 1, 5 y 6. Se parte con el valor 1 y se definen los

dos grupos  $a_1 \leq 1$  versus  $a_1 > 1$ . En el primer grupo se incluyen todos los contratos que se caracterizan porque en este atributo su valor es  $\leq 1$  y en el otro grupo los restantes contratos. Se define una medida de variabilidad por ejemplo la varianza y se calcula su valor para la variable objetivo ( $\ln PRVU$ ) de todos los contratos de la población en estudio. Este valor se compara con una medida promedio de la variabilidad de los grupos  $a_1 \leq 1$  y  $a_1 > 1$  obteniéndose una medida de la reducción de variabilidad para el corte  $a_1 \leq 1$  versus  $a_1 > 1$ . Este procedimiento se repite para los cortes  $a_1 \leq 5$  vs  $a_1 > 5$  y  $a_1 \leq 6$  vs  $a_1 > 6$ . Aquel corte que genere una mayor reducción de variabilidad determina el corte óptimo para la variable  $a_1$ . Este proceso se repite para todas las otras variables  $a_2, a_3$  etc. y se elige para partir, aquella que genere la mayor reducción de variabilidad. Se tiene entonces el primer corte y por lo tanto los dos primeros grupos. Luego este proceso se repite independientemente para cada uno de los dos grupos iniciales generándose nuevos cortes, que en general pueden ser diferentes, para cada uno de ellos. El análisis termina cuando la reducción de variabilidad de la variable dependiente entre una iteración y la anterior es inferior a un umbral mínimo determinado por el investigador. En el anexo N° 1 se presenta el desarrollo formal de este proceso.

La descripción anterior muestra la importancia de dos criterios para que este mecanismo de cortes dicotómicos conduzca a un resultado óptimo: la monotonicidad de la relación entre los atributos y la variable dependiente y el umbral mínimo.

Para visualizar la importancia de la monotonicidad supóngase que la relación entre al variable dependiente  $D$  y un atributo “ $a$ ” sea cuadrática es decir  $D = a^2$  entonces los valores  $a=3$  y  $a = -3$  se asocian con el mismo valor de  $D = 9$ , por lo tanto cuando el CART genere sus grupos sobre la base de dicotomías entonces los grupos asociados al corte  $a \leq -3$  ( $a \leq -3$  vs  $a > -3$ ) y aquellos grupos asociados con el corte  $a \leq 3$  ( $a \leq 3$  vs  $a > 3$ ) van a ser prácticamente indistinguibles imposibilitando la selección de los grupos finales.

La monotonicidad de los atributos respecto del  $\ln PRVU$  será estudiada en la sección siguiente

El umbral mínimo también es importante para determinar el número de celdas que entregue el programa. En efecto si se desea que la variabilidad dentro de cada celda sea muy pequeña, en el extremo nula, entonces cada grupo va a estar constituido por muy pocos contratos y en el extremo uno solo. Por lo tanto, en este caso, el número de celdas será igual al de contratos analizados. Para disponer de un número de subgrupos razonable será necesario permitir una cierta variabilidad definida por el umbral mínimo. Este punto será estudiado en la sección sobre construcción del índice propuesto.

Finalmente es interesante destacar que para respetar el criterio institucional mencionado primero en la sección 1, el método CART se aplicará *por separado* a las rentas vitalicias de vejez anticipadas, y al conjunto de las rentas vitalicias de vejez normal y postergadas.

### Revisión del cumplimiento del requisito de Monotonicidad

Corresponde ahora, revisar el requisito de monotonicidad:

- Atributos dicotómicos, que siempre son monotónicos:
  - carácter conjunto versus individual de la renta vitalicia. Siempre el precio de una renta vitalicia unitaria ( $PRVU_t$ ) es mayor para una renta conjunta, cuando los demás atributos se mantienen constantes. Luego, hay monotonicidad.
  - Género del causante, en el caso de un contrato individual. El precio de la renta vitalicia unitaria ( $PRVU_t$ ) es siempre mayor para la mujer, cuando los demás atributos —como la edad— se mantienen constantes. Esta es la contracara del hecho de que en todo el mundo las mujeres tienen mayor esperanza de vida que los hombres a cada edad, situación envidiada por éstos últimos. En todo caso, se cumple la monotonicidad.
- Atributos discretos con tres o más valores (también pueden ser representados con una variable continua):
  - Período de pago garantizado: los valores son 0, 5, 10, 15 y 20 años. Mientras mayor sea el período garantizado, es mayor el precio de la renta vitalicia unitaria. Luego, hay monotonicidad.
  - Diferimiento desde la compra de la póliza hasta su primer pago: los valores son 0, 12, 24 y 36 meses. A mayor diferimiento, siempre es menor el precio de la renta vitalicia unitaria. Luego, también hay monotonicidad.
- Atributos continuos:
  - Edad del *cónyuge* al momento de contratar la pensión: Siempre el precio de la renta vitalicia unitaria ( $PRVU_t$ ) es menor cuando la edad *del cónyuge* es mayor, cuando los demás atributos se mantienen constantes. En el caso de aquellas pólizas individuales, donde no hay *cónyuge*, se redefine la edad del *cónyuge* como igual a 110 años, pues así la tabla de mortalidad le asigna una probabilidad de sobrevivir igual a cero. Esto reconoce la cercanía de la renta individual con el caso de renta conjunta donde el *cónyuge* tiene 100 ó 105 años de edad.
  - Anticipación o adelanto de la compra de la póliza respecto a la edad normal del *causante*. Esta variable se define como:

$$a_{it} \equiv \left( \begin{array}{c} \text{edad de pensión normal} \\ \text{del causante } i \end{array} \right) - \left( \begin{array}{c} \text{edad } E \text{ del causante} \\ \text{que contrata en } t \end{array} \right) \quad (13)$$

En este Informe, se aplica el método CART por separado a las pólizas de vejez anticipada, y a las pólizas por vejez normal. Esto implica que la variable  $a_{it}$  siempre tendrá signos positivos en la primera muestra, y

siempre tendrá signos negativos en la segunda muestra. Para que en ambos casos la variable  $a$  sea positiva, se emplea la variable atraso en el caso de las rentas normales que se define como el valor negativo de  $a_{it}$ .

Siempre el precio de la renta vitalicia unitaria ( $PRVU_t$ ) es mayor cuando la anticipación  $a_{it}$  es mayor (o cuando es menos negativa), si los demás atributos se mantienen constantes. Por eso hay monotonicidad.

Debido a que la edad legal de pensión normal fue constante en la muestra (60 mujeres, 65 hombres), la variable  $a_{it}$  está perfectamente correlacionada con la *edad del causante al contratar*, una vez dado el género del causante. Preferimos a la variable  $a_{it}$  por sobre la *edad del causante*, porque ofrece al método CART la libertad de agrupar pólizas de hombres y mujeres, si resultan tener una anticipación similar.

- Tamaño de prima. En una entrevista con una importante CSV de la plaza se nos sugirió que el tamaño de prima podría mostrar un efecto no monotónico, sobre el precio de la renta vitalicia unitaria la  $PRVU_t$ . En efecto, se sabe que las pólizas pequeñas son relativamente más afectadas por los costos fijos, que incluyen el costo de administración y pagos, la comercialización y el desembolso de 15 UF por cuota mortuoria. Por otro lado, se nos informó que las pólizas más grandes serían afectadas en mayor grado por la mayor esperanza de vida de los grupos de mayores ingresos, explicada a su vez en su mayor inversión en salud tanto en la vida activa como en la vejez.

Si bien ambos efectos pueden existir, podría ocurrir que uno de ellos domine en todo el rango de valores de la prima. Si hubiera dominancia de este tipo, entonces habría monotonicidad y se podría usar el método CART sin correcciones.

Para evaluar la hipótesis de que el tamaño de prima presenta una relación no monotónica con  $\ln(PRVI_i)$ , se corrió la regresión múltiple que permitió estudiar el comportamiento de dicha variable dependiente con respecto a la prima cuando se controla por los otros atributos considerados. Los resultados se presentan en el Anexo 2. La conclusión es que para fines del CART el tamaño de prima también se comporta monotónicamente.

### *Metodología de aplicación del CART y grupos finales*

El método CART emplea un conjunto de variables para construir los grupos finales. Obviamente, las variables consideradas son aquellas determinantes de la variable dependiente. En este caso los atributos considerados son: el género del asegurado, la prima pagada, el periodo garantizado, el periodo diferido que se haya pactado, el adelanto o atraso con que la persona se pensionó<sup>18</sup>, las tasas de interés y un conjunto de variables de organización industrial que tienen una dimensión temporal. Ejemplos de estas últimas son los métodos de marketing de las CSV (que varían en el tiempo), la carga efectiva de las regulaciones de solvencia y de otro tipo que deban cumplir las CSV, la entrada de nuevas CSV y las fusiones en la medida que afecten el grado de competencia entre ellas.

En la regresión empleada para determinar la monotonía de la prima (Anexo 2), las variables de interés fueron descritas por el valor esperado de la renta vitalicia unitaria neta del valor de la cuota mortuoria y aquellas otras de organización industrial por variables mudas temporales (una por cada uno de los 42 meses que incluye la muestra). No es posible incluir estas variables entre los clasificadores de los grupos construidos por el CART puesto que no son públicas o muestran los cambios históricos y por lo tanto reflejan un desarrollo pasado.

Fue necesario diseñar una metodología en que bastara el uso de las variables género del asegurado, prima pagada, periodo garantizado, periodo diferido que se haya pactado y el adelanto o atraso con que la persona se pensionó, todas ellas públicas, para caracterizar adecuadamente el PRVU. Para resolver este problema, se decidió seguir un esquema secuencial en que se empleó el CART para calcular los grupos mensualmente pues de esta forma las variables de interés y de organización industrial si bien desconocidas, permanecen constantes y por lo tanto las variaciones del PRVU se asocian directamente con aquellas mostradas por las variables públicas. El resultado es un conjunto de clasificaciones, una para cada uno de los meses considerados. Al respecto, se decidió concentrarse sólo en los primeros 36 de los 42 meses de que se dispone de información, con el objeto de dejar los 6 meses restantes para verificación de los resultados fuera de muestra.

El paso siguiente consistió formar grupos de clasificaciones agrupando todos aquellos meses en que las clasificaciones mostraran el mismo patrón, definido por la igualdad de las variables de clasificación de las celdas y por la similitud de los puntos de corte. Después de un análisis sistemático basado en una inspección visual de las variables de corte se constató que era posible considerar dos grandes grupos en la base de contratos normales: uno que incluía a 21 de los 36 meses<sup>19</sup> y un segundo que consideraba otros 4

---

<sup>18</sup> El adelanto o atraso se define como la diferencia entre la edad legal para pensionarse menos la edad efectiva de compra de la pensión vitalicia. Las variables se computan de tal manera que siempre el resultado sea positivo.

<sup>19</sup> Los meses considerados son: 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 20, 21, 25, 26, 27, 28, 29, 30, 34, 35.

meses<sup>20</sup>. Las restantes clasificaciones no parecieron mostrar similitud significativa y fueron dejadas de lado. Para el caso de los contratos adelantados también se pudo configurar dos grupos: uno constituido por 15 meses<sup>21</sup> y el segundo, por otros 5 meses<sup>22</sup>.

El paso siguiente consistió en construir nuevas clasificaciones en función de estos grupos. Así, se aplicó el CART a los 21 meses del primer grupo y se configuró una nueva clasificación llamada grupo A, otro tanto se hizo con los 4 meses del segundo grupo generando el llamado grupo B. Las dos agrupaciones de la base adelantada generaron también dos grupos. Sin embargo, en este caso, ellos fueron idénticos y por lo tanto se procedió a agregar los 20 meses y construir una agrupación con la totalidad que se denominó grupo C y representa la clasificación de la base de contratos adelantados. Las características de estas clasificaciones se presentan en el Anexo 3

### *Clasificación definitiva*

Los contratos normales comprados en cada uno de los 36 meses se ordenaron según la clasificación del grupo A primero y luego del grupo B y se calculó para cada uno de ellos, la varianza de log (PRVU). Luego dichas varianzas se promediaron dentro de cada mes y finalmente dentro de todos los meses para disponer de una cifra de la variabilidad total dentro de muestra para la clasificación A y B. Estos cálculos se repitieron para los 6 meses finales y luego para los 42 meses y se compararon. El cuadro 2 muestra la desviación estándar de cada grupo para los distintos períodos. Para las rentas vitalicias de vejez normal, la variabilidad del grupo B resultó la menor para todos los períodos, pero como las diferencias no son significativas —en ningún caso la desviación estándar del grupo A es mayor a 2,9% la del grupo B— y el grupo A es más estable, se eligió este como representativo de toda dicha base.

Los resultados de variabilidad para las rentas vitalicias de vejez anticipada se reportan también el cuadro 2, aunque como ya se mencionó, en este caso el resultado del CART arrojó solamente un grupo representativo en toda la muestra.

*Cuadro N° 2. Desviación Estándar grupos CART*

	RV Normales		RV Anticipadas
	Grupo A Desv. Est.	Grupo B Desv. Est.	Grupo C Desv. Est.
Muestra 36 Meses	0.0827	0.0803	0.0662
Muestra Ultimos 6 Meses	0.0793	0.0791	0.0657
Muestra Total	0.0818	0.0800	0.0662

<sup>20</sup> Los meses considerados son: 2, 17, 24, 31.

<sup>21</sup> Los meses considerados son: 5, 10, 13, 14, 16, 19, 21, 22, 24, 25, 28, 29, 30, 36.

<sup>22</sup> Los meses considerados son: 15, 20, 27, 32, 34.

La comparación de variabilidad junto con el criterio de representatividad de los grupos definidos por el CART, permiten definir los grupos para los cuales se calcularán los índices. Las características de estos grupos se presentan más adelante en la sección 4.

Antes de pasar a los resultados se debe mencionar que el CART exige que se coloque un umbral de mejoramiento de la variabilidad entre iteración. Dicho umbral permite definir el número de celdas que entregará el programa. En este caso se eligió un  $C = 0.07$  que permitió dividir la base en cuatro grupos. El cuadro 3 indica que la variabilidad al interior de cada grupo, medida por el coeficiente de variación de  $\ln PRVU$ , es siempre inferior a 2 % con una sola excepción en que alcanza al 2,26%. Esto muestra que los valores de la variable independiente ( $\ln PRVU$ ) son lo suficientemente similares como para que se pueda decir que el error de composición asociado a la variabilidad intracelda es prácticamente inexistente.

*Cuadro N° 3. Variabilidad de  $\ln(PRVU)$  al interior de los subgrupos*

Subgrupo	RV Normales				RV Anticipadas			
	N	Desv Estand	MediaY	Coef. Var. %	N	Desv Estand	MediaY	Coef.Var %
1	1066	0,11	5,09	2,16	2193	0,06	5,32	1,17
2	2995	0,08	5,22	1,50	950	0,07	5,23	1,40
3	859	0,09	5,20	1,71	683	0,06	5,41	1,07
4	3440	0,06	5,33	1,18	3548	0,07	5,40	1,37

### 3.4 Metodología para la construcción de los índices parciales de valor del PRVU

El resultado de la etapa anterior son 4 celdas para vejez anticipada y 4 celdas para vejez normal. En cada celda se calcula – y se propone publicar- un valor mensual para el subíndice respectivo, calculado de la siguiente manera:

$$IPRVU_{jt} \equiv \exp\left(\frac{1}{N_{jt}} \sum_{v_i \in j \wedge \epsilon t} \ln(PRUVU_{it})\right) \quad (14a)$$

Este subíndice mensual está expresado en unidades naturales de  $PRVU$  (meses), gracias a la operación de exponenciación, que cancela el efecto del logaritmo natural. Otra forma de verlo es que cada  $IPRVU$  es un promedio geométrico de los  $PRVU_i$  al interior de la celda o grupo.

El subíndice  $IPRVU_{jt}$  es utilizable directamente por un asesor previsional, como divisor de un stock de ahorro acumulado que determina la magnitud de la pensión que se habría obtenido en caso de haber adquirido una pensión tipo  $j$  en el mes  $t$ .

Además, el nuevo subíndice representa un indicador del precio de las rentas vitalicias que está libre de los problemas de la tasa interna de retorno encontrados en la *TIM*.

Observamos que el  $N_{jt}$  que se utiliza para calcular  $IPRVU_{jt}$  corresponde al número de pólizas vendidas en el período (mes)  $t$  que se ubican en la celda  $j$ . Esto no expone al subíndice a los sesgos del efecto composición, porque por construcción, el método CART asegura que las pólizas en el grupo  $j$  son muy similares entre sí en cuanto al valor de la variable indicadora, que es lo que controla el efecto composición.

Por su parte, un gestor financiero puede recuperar el valor logarítmico de éste subíndice por medio de aplicar logaritmo a la cifra publicada, que es  $IPRVU_{jt}$ :

$$\ln(IPRVU_{jt}) \equiv \left( \frac{1}{N_{jt}} \sum_{\forall i \in j \wedge \epsilon t} \ln(PRVU_{it}) \right) \quad (14b)$$

El lado derecho es justamente el dato que necesita el gestor financiero, como se explicó anteriormente.

La SVS también debe construir un índice *general* de los precios de las rentas vitalicias. Para esta etapa, proponemos construir un índice general de precios de las rentas vitalicias unitarias del tipo vejez anticipada, a partir de los  $IPRVU_{jt}$  de las cuatro (en general del número  $J_{VA}$  de) celdas de vejez anticipada. Lo mismo se aplica, *mutatis mutandi*, para construir el índice general de precios de las rentas por vejez normal.

Para evitar el efecto composición, requerimos que los ponderadores para las distintas celdas no varíen a lo largo del tiempo, sino que sean los ponderadores obtenidos del período base, que denominamos  $\alpha_{j0}$ . Recomendamos que este período base sea el de los *últimos* doce meses de la muestra con el fin de facilitar su uso en los períodos posteriores al término de la muestra.

Proponemos actualizar el período base cada cinco años. Esa actualización implica usar nuevamente el método CART o similar para redefinir la identidad las celdas  $J$  y sus ponderadores  $\alpha_{j0}$ .

Proponemos usar un promedio lineal para obtener el índice general de precios de las rentas vitalicias unitarias del tipo vejez anticipada (igual para vejez normal), a partir de los  $IPRVU_{jt}$  de las  $J_{VA}$  celdas de vejez anticipada:

$$IGPRVU_t^{VA} \equiv \sum_{\forall j \in J_{VA}} \alpha_{j0} \cdot IPRVU_{jt} \quad (15)$$

Por último, debemos proponer una forma funcional para los ponderadores  $\alpha_{j0}$ . Por las razones mencionadas en la sección 2.1, no recomendamos ponderar según tamaño de prima. Para reiterar:

- a) Cada afiliado compra una sola póliza y una familia a lo más es beneficiaria de dos pólizas; además, aquellas pólizas muy grandes, como aquellas que superan 10.000 UF, no afectan el precio de las pólizas que enfrenta la familia mediana o la familia promedio.
- b) Las CSV no utilizan este índice para sus decisiones financieras, porque cuentan con índices más afinados.

Se propone ponderar según el número de pólizas en cada celda. Esta es la medida más adecuada cuando cada afiliado adquiere una sola póliza. Es decir, proponemos:

$$\alpha_{j0} \equiv \frac{N_{j0}}{\sum_{k \in J_{VA}} N_{k0}} \quad (16)$$

Notamos que  $N_{jt}$ , el número de observaciones en el mes  $t$  en cada celda, no entra en esta fórmula. Esto protege al índice general del efecto composición.

#### 4. Resultados de la implementación piloto del nuevo índice propuesto

Los grupos obtenidos después de la aplicación del CART fueron los siguientes:

*Cuadro N° 4: Identidad de los grupos o subíndices que resultaron más relevantes*

N°	<i>Vejez Anticipada</i>			<i>Vejez Normal</i>		
	Atributos $k$ del subgrupo $j$ (los atributos omitidos se aplican a todos)	$\sum_t N_{jt}$	$\alpha_{j0}$ (%)	Atributos $k$ del subgrupo $j$ (los atributos omitidos se aplican a todos)	$\sum_t N_{jt}$	$\alpha_{j0}$ (%)
1	Hombre, RV Conjunta, Adelanto inferior a 5,79 años	281	6,91	Hombre, Individual	445	5,26
2	Hombre, RV Individual, Adelanto inferior a 5,79 años	1610	39,60	Hombre, Conjunta	2846	33,62
3	Mujer, Adelanto inferior a 5,79 años	410	10,08	Mujer, Atraso pensión superior a 4 años	1237	14,61
4	Adelanto superior o igual a 5,79 años	1765	43,41	Mujer, Atraso inferior a 4 años	3938	46,52

Nota: Los  $\alpha_{j0}$  y los  $\sum_t N_{jt}$  reportados se calculan sólo con el número de contratos de los últimos 12 meses de la muestra, es decir con  $31 \leq t \leq 42$  (Julio 2007 a junio 2008).

Este cuadro revela que la agrupación según el número de años de anticipación de la pensión es de gran importancia. El género del causante de la RV, así como también el hecho de que la RV sea individual o conjunta también aparecen como características importantes para definir grupos. En cambio, la agrupación según otras características

del contrato como períodos de pagos garantizados, períodos de pagos diferidos, y el monto de la prima no parecen ser relevante.

Como se desarrolló en la sección 3.4, una vez construidos los grupos se procede a calcular los subíndices respectivos. Para los dos casos de renta vitalicia de vejez normal y de vejez anticipada, los grupos encontrados fueron 4, por lo que a continuación se presentan un total de 8 subíndices. Los gráficos 3 y 4 muestran la evolución de los subíndices de vejez normal y vejez anticipada respectivamente. En los dos casos se observan diferencias importantes en los niveles de los distintos subgrupos lo que corrobora la necesidad de tener ponderadores fijos para evitar que se cumpla el requisito (b) del efecto de composición. Para el caso de las rentas vitalicias de vejez normal por ejemplo, las mayores diferencias en los niveles se encuentran entre el subgrupo de hombres con contrato individual (subgrupo 1) y el de mujeres con postergación de pensión de menos de 4 años (subgrupo 3). La diferencia entre el *IPRVU* de estos dos subgrupos llega a ser de 60 meses en diciembre de 2006. Para las rentas vitalicias de vejez anticipada en tanto, los grupos más disímiles son el de hombres con RV individual (subgrupo 1), y el de mujeres con adelanto inferior a 5,79 años (subgrupo 3). En este caso la diferencia llega a ser de 50 meses en enero de 2007.

Gráfico N° 3: Subíndices Mensuales PRVU Vejez Normal

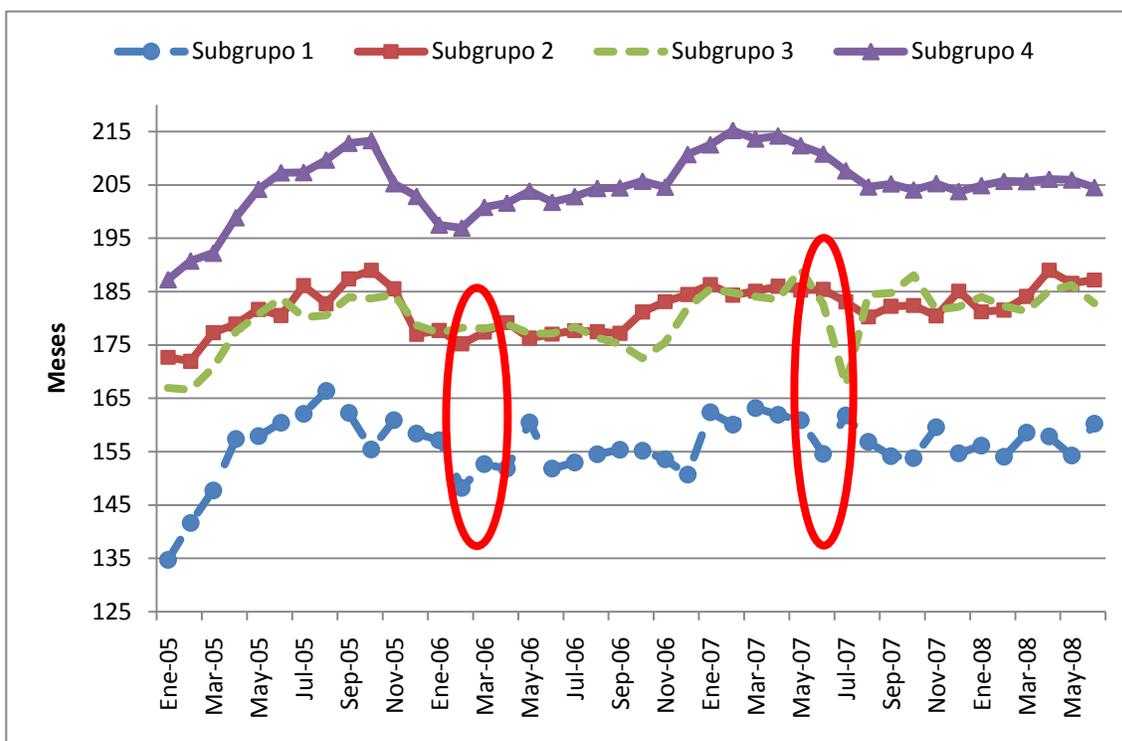
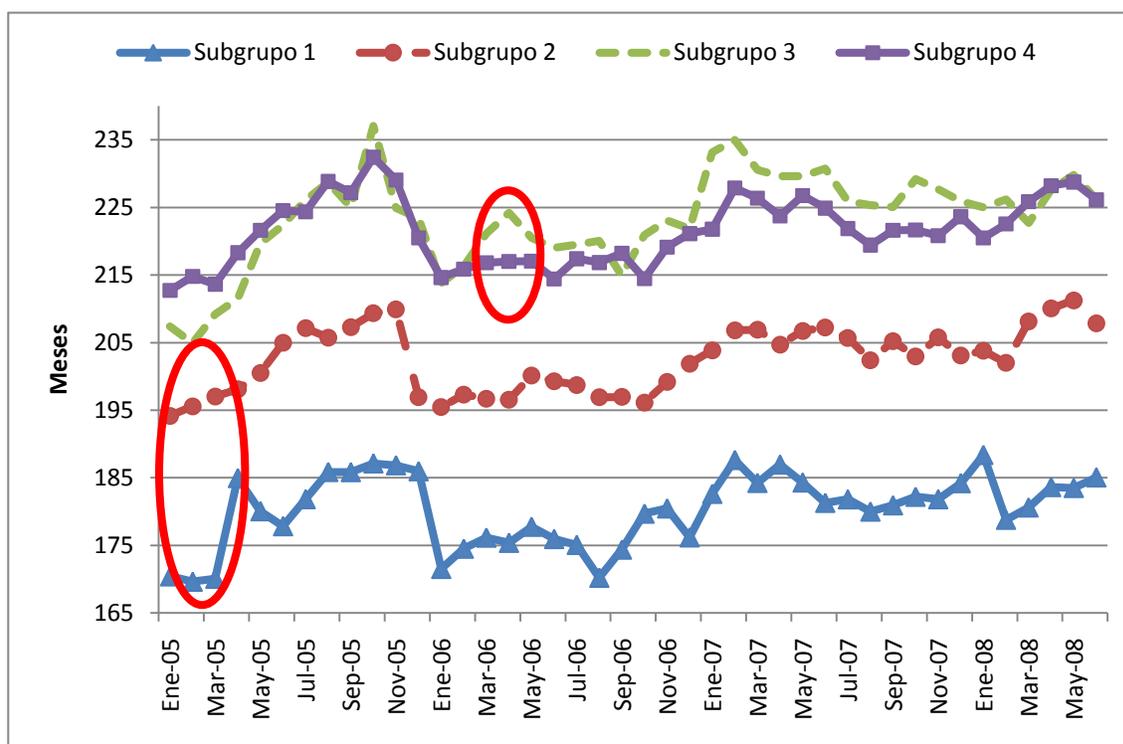


Gráfico N° 4 Subíndices mensuales PRVU Vejez Anticipada



Además de las diferencias encontradas en los niveles de los subíndices —lo que en parte justifica su cómputo—, las correlaciones obtenidas muestran que estos difieren en su comportamiento en el tiempo. Los cuadros 5 y 6 corresponden a las matrices de correlación para los subgrupos de rentas vitalicias de vejez normal y de rentas vitalicias de vejez anticipada respectivamente. Las correlaciones entre los subgrupos son más bajas que la unidad. Encontramos una correlación promedio de 0,68 para los subgrupos de rentas vitalicias normales y de 0,78 para aquellos de rentas vitalicias anticipadas. Los círculos en los gráficos 3 y 4 ilustran las bajas correlaciones existentes entre algunos subgrupos pues muestran casos en que ciertos subíndices subieron fuertemente mientras otros lo hicieron en menor cuantía o incluso llegaron a bajar

Las diferencias encontradas en los niveles del IPRVU, junto con la existencia de correlaciones distintas a 1 (y en ocasiones bastante más bajas), justifican los subgrupos creados. Los subíndices no sólo aportan información respecto del monto en que difiere en un mes en particular el PRVU entre distintos tipos de contratos, sino que ilustran las distintas tendencias que dichos precios pueden experimentar a lo largo del tiempo.

*Cuadro N° 5. Matriz de Correlaciones Subgrupos Rentas Vitalicias Normales*

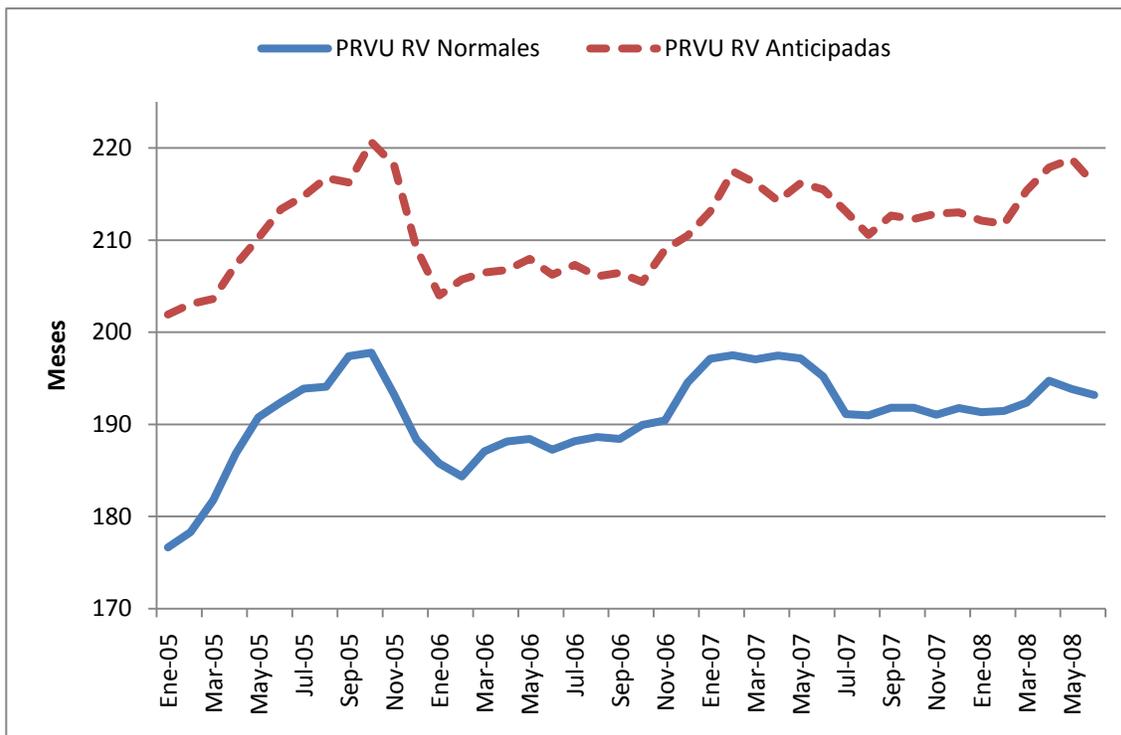
	Subgrupo 1	Subgrupo 2	Subgrupo 3	Subgrupo 4
Subgrupo 1	1			
Subgrupo 2	0.61	1		
Subgrupo 3	0.55	0.69	1	
Subgrupo 4	0.76	0.79	0.68	1

*Cuadro N° 6. Matriz de Correlaciones Subgrupos Rentas Vitalicias Anticipadas*

	Subgrupo 1	Subgrupo 2	Subgrupo 3	Subgrupo 4
Subgrupo 1	1			
Subgrupo 2	0.73	1		
Subgrupo 3	0.76	0.77	1	
Subgrupo 4	0.78	0.92	0.78	1

Dado los subíndices calculados, el índice general de precios para rentas vitalicias de vejez normal y de vejez anticipada se calcula con la metodología presentada en la sección 3.4. El gráfico 5 muestra los índices.

*Gráfico N° 5: Índices Mensuales PRVU agregado Vejez Normal y Vejez anticipada*



Los resultados indican que los niveles de los índices de PRVU difieren en promedio en 20,4 meses, con un máximo de 25,3 meses en enero de 2005. Las pensiones obtenidas de manera anticipada son sostenidamente más caras que aquellas obtenidas de acuerdo a la edad legal de jubilación. Este resultado no es nuevo y se debe principalmente a la mayor duración esperada del primer tipo de contrato. Sin embargo los índices propuestos en este estudio permiten conocer el monto de esta diferencia entre otras cosas.

Por último, la correlación encontrada entre los dos índices en el período estudiado es de 0,88. Esta correlación es alta en relación a las encontradas entre los distintos subgrupos, aunque sigue siendo menor a 1.

## 5. Conclusiones y recomendaciones

Este estudio propuso una metodología rigurosa para la confección de un indicador del precio de contratos de rentas vitalicias previsionales. Dicha metodología busca eliminar la existencia de un sesgo en este indicador debido al efecto composición que suele afectar a los índices de precios. Para estos efectos, se utilizó el método CART para la conformación de subgrupos relativamente homogéneos, y se definieron ponderadores fijos para cada uno de estos subgrupos.

El índice construido es robusto a cambios en la forma de la curva de rendimientos. A diferencia de la *TIM* (actualmente calculada por la SVS), el índice presentado recoge la estructura de tasas efectiva en cada momento del tiempo. Esto se debe al cálculo del índice de manera directa, tomando en cuenta la prima cancelada por el causante y la pensión obtenida mediante el contrato, y sin hacer supuestos acerca de la estructura de tasas aplicada a la tarificación del contrato.

Los resultados del método CART son la conformación de 4 subgrupos para contratos de rentas vitalicias de vejez normal y otros 4 para contratos de rentas vitalicias de vejez anticipada.

Para el período abarcado entre enero de 2005 y junio de 2008, se elaboraron índices para cada uno de los 8 subgrupos. Los resultados evidencian diferencias importantes en los precios cancelados por las rentas vitalicias en los distintos subgrupos.

Además, los subíndices muestran ciertas diferencias a lo largo del tiempo, lo que se ve reflejado en correlaciones distintas a 1, y en algunos casos bastante menores.

El índice de precios presentado complementa las estadísticas existentes hoy en día en el mercado de las rentas vitalicias. Su uso se ve facilitado por su naturaleza simple e intuitiva, quedando al alcance de asesores previsionales, y del público en general.

Una utilización nueva corresponde a la que puedan hacer administradores de fondos de pensión al incluir el precio de los contratos de rentas vitalicias en sus decisiones

de inversión y proteger al cotizante del riesgo de precio que enfrentan a la hora de comprar una pensión. Esta función del índice es novedosa y requiere un desarrollo formal<sup>23</sup>.

Se recomienda extender la metodología propuesta a las pólizas de renta vitalicia por invalidez y sobrevivencia. También se recomienda difundir a la comunidad de asesores previsionales, gestores financieros y la opinión pública en general, cómo se utilizan los nuevos índices propuestos y las ventajas de hacerlo.

---

<sup>23</sup> Ver Valdés-Prieto, Castro (2010) para una discusión detallada.

## 7. Bibliografía

Breiman, Friedman, Olshen y Stone (1984) *Classification and Regression Trees*, Wadsworth Press, Belmont, California.

Breiman, Leo (1996), "Bagging Predictors", *Machine Learning*, 26, 123-140.

Bühlmann Peter y Bin Yu. (2002), "Analysing Bagging", *Annals of Statistics*, 30, 927-961.

Croux, C. , K. Joossens y A. Lemmens (2007) "Trimmed Bagging", Decision Sciences and Information Management (KBI) Working Paper 0721, Department of Applied Economics, Katholieke Universiteit Leuven.

Imbens, Guido W. y Tony Lancaster (1996), "Efficient Estimation and Stratified Sampling", *Journal of Econometrics*, 74, 289-318.

Lemmens, A. y C. Croux (2006) "Bagging and Boosting Classification Trees to Predict Churn", *Journal of Marketing Research*, también Research Report OR 0361, Department of Applied Economics, Katholieke Universiteit Leuven.

Valdés-Prieto S., Castro N. (2010) "Annuity-price risk and optimal portfolios before annuity purchase", presentado en la Jornada Anual de Investigación de la Facultad de Ciencias Económicas y Administrativas de la Pontificia Universidad Católica de Chile, Enero 2010.

Zhang, P. , Z. Zhang<sup>1</sup>, A. Li, and Y. Shi (2008) "Global and Local Bagging Approach for Classifying Noisy Dataset", *International Journal of Software and Informatics*, Vol.2, No.2, December 2008, pp. 181-197.

## **Anexo 1: Descripción estadística del método CART**

El método CART (*Classification and Regression Trees*) es un procedimiento estadístico que puede ser usado con dos objetivos distintos.

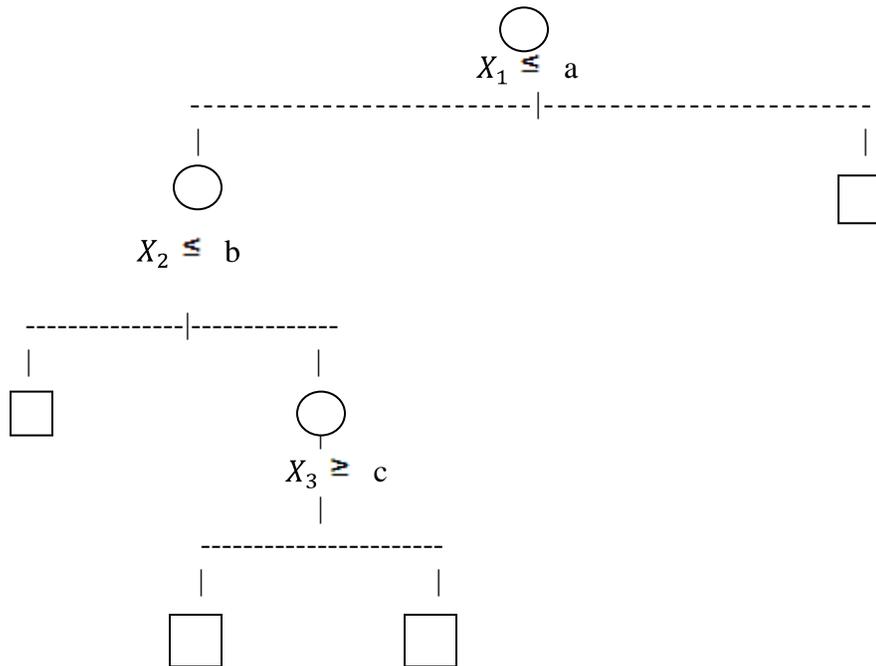
En un caso existe una variable de interés que tiene varios valores. Por ejemplo, una variable que mide la calidad de aire y admite los valores “normal”, “preemergencia” y “emergencia” que están definidos con claridad. Además se cuenta con un conjunto de variables que son observables y contribuyen a determinar la calidad del aire. En este caso, el objetivo consiste en identificar el conjunto de valores de estas últimas variables que permitirían predecir en “la mejor forma posible” si la calidad del aire (la variable de interés) es normal, preemergencia o emergencia.

En el otro caso, el método CART se emplea para identificar la división óptima de un conjunto de objetos en grupos que sean “óptimos”. En este caso cada objeto presenta un valor para la variable de interés –el logaritmo de PRVU en nuestro caso -. También aquí se cuenta con datos sobre otros atributos del objeto, denominados genéricamente  $X_1, X_2, \dots, X_K$ . En este caso, el método CART divide el conjunto de objetos en grupos que son “óptimos”, en el sentido de que minimizan la variabilidad de la variable de interés al interior de cada grupo.

En este informe se emplea el CART para resolver este segundo tipo de problema.

Para aclarar estas ideas, presentamos la notación que utiliza el CART a través de un ejemplo. Supóngase que el resultado de la aplicación del método puede ser descrito por el gráfico A1.1

Gráfico A1.1: Notación utilizada por el método CART



Los círculos denotan celdas o nodos intermedios, es decir que admiten nuevas clasificaciones o divisiones, mientras que los cuadrados denotan celdas o nodos terminales.

El gráfico muestra que se parte con una celda que incluye toda la población de objetos en estudio. El primer paso consiste en elegir un atributo o variable - en este ejemplo se eligió a  $X_1$  - y un punto de corte para dicha variable - en este ejemplo es "a", que divide a la población según si cada objeto presenta un valor  $X_1 \leq a$  ó  $X_1 > a$ . Los objetos con  $X_1 \leq a$  se agrupan en el nodo de la izquierda, y los demás objetos se agrupan a la derecha. Notemos que esta agrupación es independiente de los valores que los objetos presentan para las variables  $X_2, X_3, \dots, X_K$ .

El paso siguiente parte de la celda caracterizada por los valores  $X_1 \leq a$  y divide esa población en dos subgrupos. Para ello se elige a la variable  $X_2$  y al punto de corte b.

Supondremos por ahora que la población definida por  $X_1 > a$  no se vuelve a clasificar o dividir, debido al incumplimiento de un criterio que explicamos más adelante, y por tanto es terminal. También suponemos que la población de objetos donde  $X_2 \leq b$  es una celda que no admite nueva clasificación (también es terminal). Esta nueva celda queda definida por  $X_1 \leq a$  y  $X_2 \leq b$ , es decir por dos desigualdades y no por una.

Supondremos que la celda restante, donde  $X_2 > b$ , admite una nueva clasificación, y se elige para ello a la variable  $X_3$  y el punto de corte c. Finalmente suponemos que los

dos grupos que resultan de esta nueva clasificación no admiten nuevas divisiones y son también terminales.

Esta descripción deja en claro que la especificación del método CART debe responder las siguientes tres preguntas básicas:

¿Por qué se elige en primer término la variable  $X_1$ , luego  $X_2$  y finalmente  $X_3$ ? El orden podría ser otro.

b) ¿Cómo se eligen los cortes a, b y c, respectivamente?

c) ¿Cual es el criterio para establecer que una celda es terminal, donde una subdivisión adicional no es admisible?

También interesa calcular la variabilidad de la variable objetivo en cada una de las celdas terminales y en el conjunto de ellas. Las secciones que siguen responden estas preguntas.

*a. Elección de las variables para cortar y de los valores de corte*

El CART es un método de análisis sistemático y secuencial de datos que analiza uno por uno todos los atributos disponibles y forma grupos dicotómicos con ellos a los que aplica un criterio de para definir el corte óptimo. Para precisar este punto supóngase un atributo  $X_1$  que tiene sólo tres valores 1, 5 y 6. Se parte con el valor 1 y se definen los dos grupos  $X_1 \leq 1$  versus  $X_1 > 1$ . En el primer grupo se incluyen todos los contratos que se caracterizan porque en este atributo su valor es  $\leq 1$  y en el otro grupo los restantes contratos. Se define una medida de variabilidad y se calcula su valor para la variable objetivo ( $\ln PRVU$ ) de todos los contratos de la población en estudio. Este valor se compara con una medida promedio de la variabilidad de los grupos  $X_1 \leq 1$  y  $X_1 > 1$  obteniéndose una medida de la reducción de variabilidad para el corte  $X_1 \leq 1$  versus  $X_1 > 1$ . Este procedimiento se repite para los cortes  $X_1 \leq 5$  vs  $X_1 > 5$  y  $X_1 \leq 6$  vs  $X_1 > 6$ . El corte que maximiza la reducción de variabilidad respecto de la situación de base es el corte seleccionado para el atributo  $X_1$ .

Este proceso se repite para todas las otras variables  $X_2, X_3 \dots$  y se elige para partir, aquella variable que genere la mayor reducción de variabilidad. En definitiva, el método CART se define por seleccionar a aquél atributo, y su respectivo corte optimizado, que entrega la mayor reducción de variabilidad. Este primer corte define dos grupos. El método que se acaba de describir se repite independientemente para cada uno de dichos dos grupos. Esta etapa identifica nuevos atributos prioritarios y sus respectivos cortes, que en general pueden ser diferentes, para cada uno de estos grupos.

En términos más precisos, se parte de toda la población en estudio y se calcula una medida de variabilidad  $R$  de la variable objetivo. En nuestro caso, esa medida de variabilidad es la varianza de log PRVU para las pólizas del universo. Se tiene entonces

que para el universo inicial de pólizas la variabilidad es,  $R_1 = \frac{\sum(y_i - \bar{y})^2}{N}$ . Luego de definir el primer atributo y su corte óptimo se generan dos grupos  $g_1$  y  $g_2$  y se calcula la medida de variabilidad  $R$  para cada uno de ellos. Se tiene entonces:  $R_2^{g_1} = \frac{\sum(y_i - \bar{y}/g_1)^2}{N^{g_1}}$  y  $R_2^{g_2} = \frac{\sum(y_i - \bar{y}/g_2)^2}{N^{g_2}}$ . El sub índice 2 define los dos sub grupos surgidos de dividir la población original

Este proceso continua y para el caso del corte  $s$  (referido a una variable y aun valor particular de ella) en el nodo  $t$ , se define la expresión  $\Delta R(s,t)$ , que indica el cambio en variabilidad al dividir el nodo  $t$  por la variable y corte  $s$  en dos grupos  $g_1s$  y  $g_2s$ . La expresión  $\Delta R(s,t)$  se define como:

$$\Delta R(s, t) \equiv R_t^s - p_1 \cdot R_t^{g_1s} - p_2 \cdot R_t^{g_2s}$$

Donde  $R_t^s$  es la variabilidad (varianza) de la celda (nodo)  $t$ ,  $R_t^{g_1s}$  y  $R_t^{g_2s}$  son las variabilidades (varianzas) de los dos grupos asociados al corte  $s$  y  $p_1$  y  $p_2$  son factores de ponderación que indican la proporción de la población (contratos) del nodo  $t$  que se ubican en los grupos  $g_1s$  y  $g_2s$  respectivamente. La variable y el punto de corte  $s^*$  elegido en cada nodo  $t$  es aquel que en que se maximiza el  $\Delta R(s,t)$ . Es decir,

$$s^* = (k^*, X_t^*) = \operatorname{argmax}_{\{k, X_t\}} \Delta R(k, X_t; t)$$

En que  $k^*$  es el corte óptimo de la variable también óptima  $X_t^*$  para el nodo  $t$ .

Es interesante destacar que el atributo y el corte que maximizan  $\Delta R(s,t)$  son los mismos que *minimizan* la suma ( $p_1 R_t^{g_1s} + p_2 R_t^{g_2s}$ ), es decir son los mismos que minimizan la variabilidad al interior cada nueva celda. Es una identidad que ello también maximiza la diferencia entre los promedios de los valores de la variable dependiente *entre* celdas. Cuando se usa la varianza como medida de variabilidad, se aprovecha la propiedad de que la expresión  $\sum(Y_i - \theta(Y))^2$  es mínima cuando  $\theta(Y) = E(Y/X)$ , donde  $Y$  es la variable objetivo y  $E(Y/X)$  es la esperanza condicional de  $Y$  dados los atributos  $X$  considerados. En este trabajo la expresión  $E(Y/X)$  se asocia a la media de la celda respectiva, puesto que ella agrupa todos los contratos en que los  $X$  cumplen las propiedades que definen dicha celda y sólo ellos. La distribución condicional incorpora el efecto de los diferentes valores de los otros atributos.

Para completar esta sección, es necesario explicar la importancia de la monotonicidad de la relación entre cada atributo y la variable dependiente para un adecuado funcionamiento del CART. Para ilustrar este punto, supóngase que la relación entre la variable dependiente  $Y$  y uno de los atributos “ $X$ ” sea cuadrática, es decir que  $Y = X^2$ . Entonces los cortes  $a = 3$  y  $a = -3$  se asocian con el mismo valor de  $Y$ , que es 9. Esto implica que cuando el CART genere sus grupos sobre la base de dicotomías, entonces los grupos asociados al corte  $X \leq -3$  seleccionan los mismos objetos que los grupos asociados con el corte  $X \leq 3$ , estos grupos van a ser indistinguibles. Esto puede generar un obstáculo serio para la optimización del corte para el atributo  $X$ , impidiendo que el método CART opere.

*b. Criterio para definir celdas terminales*

El método CART por sí mismo no determina el número de celdas terminales. Esta es una decisión del investigador que surge de su conocimiento del tema y de otros criterios como factibilidad o aplicabilidad. El elemento básico es la variabilidad de la variable de interés dentro de cada celda (nodo). En efecto, si se desea que la variabilidad dentro de cada celda sea muy pequeña, en el extremo nula, entonces cada grupo va a estar constituido por un solo objeto y por lo tanto el número de celdas será igual al de objeto incluidos en el universo inicial de objetos, lo cual es trivial e inútil. Para obtener un número razonable de grupos donde la optimización ya indicada sea informativa, será necesario permitir una cierta variabilidad intra celda. La variable que controla dicha decisión es el umbral que el investigador coloque a  $\Delta R(s,t)$  para continuar iterando. Si dicho umbral es poco restrictivo, es decir si con un cambio en variabilidad alto basta, entonces el método CART se va a detener en pocas iteraciones. Por otra parte si el umbral es pequeño, el método va a continuar dividiendo celdas que en el caso anterior habrían quedado como terminales, y por lo tanto el número de clasificaciones finales va a incrementarse.

Existen varias formas de aplicar el umbral, sin embargo en problemas que pueden asociarse a un esquema de análisis de varianza se calcula una expresión algebraica que se puede interpretar como un  $R^2_t$  (cuadrado del coeficiente de regresión en la iteración t) y que se asocia con  $\Delta R(s,t)$ . El umbral U impuesto por el investigador se aplica de tal forma que  $R^2_t / R^2_{t+1} \leq 1+U$ . Es decir las iteraciones continúan si el  $R^2$  al pasar de t a t+1 se ha reducido en al menos U unidades. En este estudio se eligió  $U = 0,07$ .

*c. Cálculo del valor de la variable objetivo y de su variabilidad intraceldas*

El programa entrega varias medidas de estadística descriptiva de las celdas terminales. Entre ellas, el valor de  $E(Y/X)$ , es decir, la media de la variable dependiente (log PRVU) condicional a los valores de los atributos que definen dichas celdas. También entrega la varianza condicional, es decir  $\frac{1}{N_i-1} \sum (Y_i - E(y/X_i))^2$  para cada una de dichas celdas y sus tamaños muestrales. Con estos antecedentes se puede calcular el coeficiente de variación de cada celda y mostrar el grado de homogeneidad de la variable dependiente. En este estudio los coeficientes de variación de todos los nodos terminales son inferiores a 2,2%. Véase Cuadro N°3, página 31.

## Anexo 2: Estudio de la monotonicidad de las variables prima, adelanto y atraso.

La ecuación (R2) en página 24 del texto principal mostró la siguiente identidad:

$PRVU_{it} \equiv E_t FCRVU_i \times \frac{1}{MWR_{it}}$ , donde La variable  $E_t FCRVU_{it}$  es el valor presente de la renta vitalicia unitaria neta del valor presente de la cuota mortuoria, y  $MWR$  es el Money's Worth Ratio. Al aplicar logaritmo natural a esta identidad, se obtiene la relación aditiva siguiente:

$$\log(PRVU_{it}) \equiv \log(E_t FCRVU_{it}) + \log(1/MWR_{it})$$

A continuación proponemos un modelo reducido para las variables de la mano derecha, que pueda ser estimado a partir de los datos de las pólizas individuales.

Respecto del segundo término de la derecha, existen buenas razones para pensar que el  $MWR$  podría depender del género del asegurado, del monto o valor de la prima pagada, del periodo garantizado, del periodo diferido que se haya pactado y del adelanto (anticipación) o atraso con que la persona se pensionó<sup>24</sup>.

Nuestro modelo para el  $MWR$  agrega efectos fijos mensuales para cada uno de los 42 meses que incluye la base de datos de la SVS. Con estas variables se desea captar aquél conjunto de factores de organización de la industria que tiene una dimensión temporal. Así, estos efectos fijos captarían cambios en los métodos de marketing de las CSV y de búsqueda de los afiliados, en los costos de transacción e impuestos, en la carga efectiva de las regulaciones de solvencia y de otro tipo que deban cumplir las CSV, el efecto de entrada de nuevas CSV, las salidas y las fusiones en la medida que afecten el grado de competencia en precios, y eventuales acuerdos de precio y guerras de precios entre CSV.

Respecto del primer término de la derecha, sabemos que las tasas de interés (curva de rendimientos) intervienen solamente en su determinación, al igual que las tablas de mortalidad. El valor de  $\log(E_t FCRVU_{it})$  también depende del género del asegurado, del periodo garantizado, del periodo diferido que se haya pactado y del adelanto (anticipación) o atraso con que la persona se pensionó.

Sabemos que el monto de la prima pagada no afecta el valor de  $\log(E_t FCRVU_{it})$ , justamente por estar definido como el valor presente de una renta vitalicia *unitaria*. Del mismo modo, la curva de rendimientos y la tabla de mortalidad, no afectan el valor de  $\log(1/MWR_{it})$ .

Encontramos que no es necesario modelar aquí la influencia de los factores indicados sobre  $\log(E_t FCRVU_{it})$ , pues gracias al apoyo de *LVA Índices* contamos con datos sobre

<sup>24</sup> El adelanto o atraso se define como la diferencia entre la edad legal para pensionarse menos la edad efectiva de compra de la pensión vitalicia para personas de cada género. Las variables se definen de tal manera que siempre el resultado sea positivo.

la curva de rendimientos (la que ellas estiman como promedio para cada mes de las curvas diarias). También conocemos la tabla de mortalidad oficial (RV-2004), que varía en el tiempo. Esos datos nos permiten calcular el término  $\log(E_tFCRVU_{it})$ , en forma directa, aunque posiblemente con algún error. Por eso, se ha incluido directamente nuestra estimación para la variable  $\log(E_tFCRVU_{it})$  como variable explicativa de la regresión. Ella aparece con el nombre “lfaclnt\_mort” en las regresiones presentadas en el Cuadro A.1.

No debe interpretarse que la implementación del índice propuesto en este estudio requiera de datos sobre la curva de rendimientos. Estos datos se han utilizado solamente en este Anexo, y sólo para estimar esta regresión, que a su vez es necesaria sólo para verificar si se cumplen las condiciones de monotonicidad necesarias para aplicar el método CART. Los coeficientes de esta regresión no se utilizan para construir el índice propuesto.

#### *Aspectos econométricos*

En una regresión de este tipo no es conveniente incluir variables explicativas que sean colineales entre sí. En este caso, la *edad del causante* es colineal con la variable *adelanto o atraso* (respecto de la edad normal) cuando se controla además por género. Para resolver este punto, nos ha parecido mejor incluir sólo adelanto (anticipación) o atraso.

Otro aspecto es que una especificación lineal para la variable adelanto o atraso presenta el defecto de imponer el mismo ajuste por año de anticipación de la pensión cuando el afiliado adelanta un año que cuando lo hace en diez años. Esa igualdad podría contradecir el hecho de que la tabla de mortalidad no es lineal y además no admite que los efectos de clientela – que pueden modificar el margen operacional y el valor del *MWR* – sean no lineales en el número de años de anticipación o atraso. Para evitar un error de especificación, se flexibiliza la forma funcional agregando una variable explicativa más: la variable adelanto o atraso elevada al cuadrado.

La monotonicidad de  $\log(PRVU)$  respecto de la variable “prima” (tamaño de la prima) es controvertida, porque como se indicó en el texto principal, fuentes de la industria nos han informado que existen dos factores contrapuestos: el costo fijo por póliza de administración eleva  $\log(PRVU)$  a medida que el monto de la prima es menor. Por otro lado, las primas mayores son pagadas por personas de mayores ingresos en el ciclo de vida, que tienen mayor esperanza de vida, y ello eleva el costo y el  $\log(PRVU)$  a medida que el monto de la prima es mayor. Es posible que cada uno de estos factores domine en un cierto rango de tamaños de prima, haciendo que el monto de prima donde  $\log(PRVU)$  es mínimo (*ceteris paribus*) sea interior al rango de montos de prima, pero también es posible que uno de estos factores domine en todo el rango y el mínimo esté en un extremo del rango. Con el fin de permitir que los datos decidieran entre estas posibilidades, se usó una especificación flexible que incluye la variable prima en forma lineal, su cuadrado, su logaritmo y su inversa. Se estimó el modelo completo, se calculó

la derivada parcial del log del PRVU con respecto al monto de la prima considerando la especificación aludida y finalmente se verificó si el signo de esta derivada se mantenía constante para todo el recorrido de dicha variable.

La monotonía respecto de la variable “prima” se cumplió para la base de datos de las rentas normales y atrasadas, revelando que uno de los factores (el costo fijo por póliza de administración) domina en todo el rango de montos de prima.

También se encontró monotonía respecto de la variable “prima” para la base de datos de las rentas adelantadas, para montos de la prima inferiores a 20.700 UF (\$435 millones de pesos, 11 veces el monto de la prima mediana). Nuevamente, el costo fijo por póliza de administración domina en el rango de montos de prima hasta 20.700 UF. Como sólo seis pólizas adelantadas de la base de datos tienen primas superiores a 20.700 UF, se optó por eliminar estos 6 contratos de la base de datos. Esto es preferible porque la monotonía de las variables permite una interpretación mucho más clara de los resultados del CART. Al final de este anexo se entrega una tabla con las características de dichos contratos eliminados. En todo caso, una vez aplicado el CART se encontró que el monto de la prima nunca es una variable prioritaria para agrupar pólizas. Esto se debe a que el efecto del tamaño de la prima sobre  $\log(\text{PRVU})$  resulta ser pequeño económicamente, excepto para montos muy pequeños de prima, inferiores a 500 UF.

#### *La base de datos utilizada*

La base de datos entregada por la SVS incluía 49 mil contratos de vejez. Así, desde el principio esta base excluyó las pólizas originadas en invalidez y en sobrevivencia. Además, la base de datos de contratos de vejez, estaba incompleta en dos aspectos adicionales:

- En el caso de las pólizas con sobrevivientes no cónyuges (hijos menores de ciertas edades, hijos inválidos), no había datos sobre la edad de los sobrevivientes potenciales al momento de contratar. Como este atributo podría incidir significativamente sobre el PRVU, su ausencia nos obligó a excluirlas de la base de datos.
- Las rentas vitalicias conjuntas donde el causante es mujer eran voluntarias en el período de los datos (hasta junio de 2008). Esto explica que existieran sólo 30 contratos con estas características. Las pólizas voluntarias tienen un método de venta diferente que justifica excluirlas de la base.

La base limpiada de los contratos con los problemas mencionados en los dos párrafos precedentes alcanza a las 36.332 observaciones. Sin embargo en dos de ellos la pensión mensual reportada era igual a cero por lo que al calcular el PRVU de acuerdo a (9b) se obtenían valores indefinidos. Por esta razón dichos contratos fueron eliminados desde el comienzo y de allí en adelante se trabajó con una base de 36.330 observaciones. Como se indicó, para asegurar monotonía respecto del monto de la prima se eliminaron 6 contratos adelantados o anticipados más. Por lo tanto la base con que se realizaron los

estudios del CART tiene 36.324 contratos, divididos en 15.292 anticipados y 21.032 normales o atrasados.

#### *Resultados de la estimación*

El modelo se estimó por separado para las pólizas de vejez anticipada y las de vejez normal (que incluye a las postergadas), porque el criterio institucional nos exige aplicar el método CART por separado a ambas muestras. Por lo tanto, también debemos evaluar la monotonía separadamente para cada muestra. Los resultados se presentan en el cuadro N° A.1, siguiente.

Cuadro N° A.1: Regresión múltiple para determinar monotonicidad

VARIABLES	lprvu_Adel	lprvu_Atraso
lfacInt_mort	0.912***	0.984***
prima	-2.31e-06	1.98e-06**
prima2	6.25e-11	-0
lprima	0.000979	-0.0138***
prim_inv	20.24***	4.890**
per_gar	3.01e-05***	2.88e-05***
per_dif	-0.000561***	-0.000168***
edad_conyug	0.000145***	0.000484***
gen	-0.0143***	0.00588***
adelanto	0.00325***	
adelanto2	-4.11e-05***	
dmes2	-0.0171***	-0.0174***
dmes3	-0.0231***	-0.0137***
dmes4	-0.0167***	0.000887
dmes5	-0.0369***	-0.00947***
dmes6	-0.0521***	-0.0275***
dmes7	-0.0442***	-0.0175***
dmes8	-0.0571***	-0.0324***
dmes9	-0.0451***	-0.0162***
dmes10	0.0324***	0.0573***
dmes11	0.0408***	0.0597***
dmes12	-0.0187***	0.00426***
dmes13	-0.0668***	-0.0300***
dmes14	-0.0646***	-0.0287***
dmes15	-0.0496***	-0.0167***
dmes16	-0.0378***	-0.00747***
dmes17	-0.0284***	0.00296**
dmes18	-0.0203***	0.00973***
dmes19	-0.0195***	0.00571***
dmes20	-0.0553***	-0.0263***
dmes21	-0.0668***	-0.0297***
dmes22	-0.0744***	-0.0344***
dmes23	-0.0857***	-0.0476***
dmes24	-0.0734***	-0.0347***
dmes25	-0.0538***	-0.0185***
dmes26	-0.0325***	-0.00331**
dmes27	-0.0424***	-0.0127***
dmes28	-0.0398***	-0.00805***
dmes29	-0.00668***	0.0202***
dmes30	0.0168***	0.0419***
dmes31	-0.00346**	0.0224***
dmes32	-0.0214***	0.00778***
dmes33	-0.0340***	-0.00300**
dmes34	-0.0271***	-0.000363
dmes35	-0.0408***	-0.0119***
dmes36	-0.0350***	-0.000790
dmes37	-0.0287***	-0.000198
dmes38	-0.0505***	-0.0191***
dmes39	-0.0416***	-0.0174***
dmes40	-0.0293***	-0.00517***
dmes41	0.0280***	0.0476***
dmes42	0.0489***	0.0647***
atraso		-0.00155***
atraso2		-6.40e-06
Constant	0.459***	0.119***
Observations	15292	21032
R-squared	0.945	0.964

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Se observa en el Cuadro A.1 que para los contratos adelantados (anticipados), los coeficientes de la variable prima son marginalmente significativos y si se calculan con estimadores robustos ninguno lo es. El resto de las variables mantienen su significancia en este caso. Por otra parte un análisis numérico de la especificación cuadrática de la variable adelanto muestra que ella se comporta monótonicamente. Esto es lo requerido para la aplicación sencillo del CART. El caso de los contratos normales es similar pero con una diferencia: la variable prima mantiene su significancia aun cuando se calculen estimadores robustos. Como en el caso anterior, las restantes variables mantienen su significancia.

También se encuentra que los efectos fijos mensuales son casi todos significativos. Ello sugiere que la organización industrial de este mercado varió en dicho período y que no controlar por ello generaría coeficientes sesgados. Ello podría sesgar el efecto del monto de la prima.

El Cuadro N° A.2 muestra el comportamiento de las derivadas parciales del log del PRVU respecto de la prima, en las bases de contratos adelantados y normales respectivamente. Puede observarse que el signo se mantiene en todo su recorrido desde su valor mínimo hasta su máximo.

*Cuadro N° A.2: Valores mínimos y máximos de las funciones Derivada de la variable prima*

DERprim_	Obs	Mean	Std. Dev.	Min	Max
Base de datos					
Anticipadas	15292	-8.97e-06	7.20e-06	-.000388	-5.23e-08
Normales y Atr.	21032	-8.28e-06	6.91e-06	-.0004367	-3.82e-07

Concluimos que para todos los efectos prácticos, hay monotonidad en la relación de la variable PRIMA sobre el logaritmo del precio de la renta vitalicia unitaria ( $PRVU_t$ ). También concluimos que las variables adelanto y atraso aunque cuadráticas son monótonicas.

Debido al cumplimiento de la monotonidad en cada atributo  $k$  por separado, el método CART es aplicable en forma directa en ambas muestras (vejez anticipada y vejez normal), sin necesidad de definir variables auxiliares.

La tabla A.3 muestra las características de los 6 contratos eliminados de la base de datos de los contratos adelantados.

*Cuadro A.3: Contratos eliminados por tener primas mayores a las 20700 UF.*

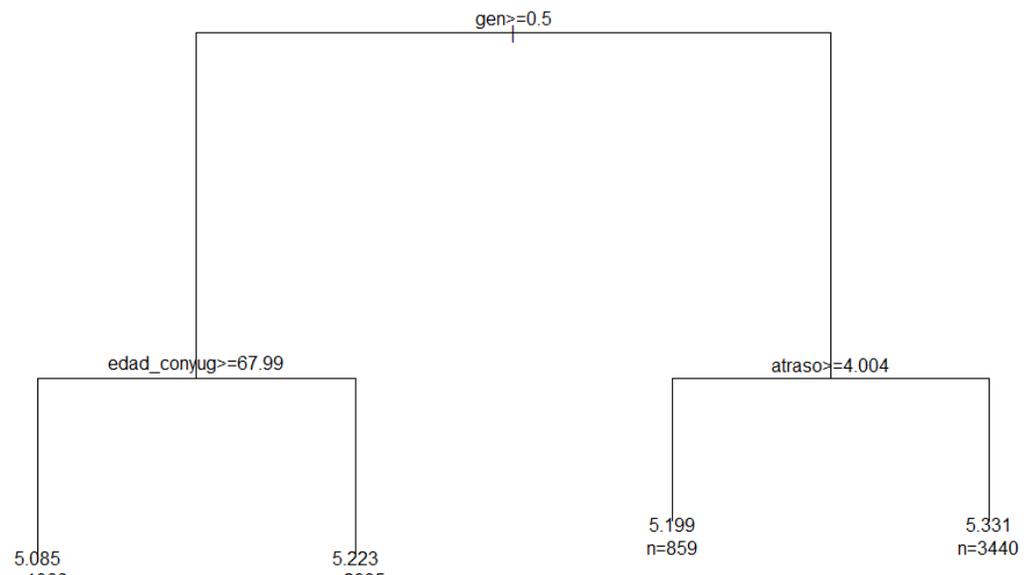
prima	per_dif	per_gar	edad_causante	edad_cony	genero	adelanto	lfacInt_mort
20.979	0	240	61	53	1	3,70	5,34
21.610	0	240	61	58	1	4,01	5,40
21.216	0	0	62	57	1	2,67	5,34
28.052	0	0	59	55	1	6,19	5,42
28.181	0	180	63	55	1	2,05	5,42
21.257	0	180	60	58	1	4,90	5,40

### Anexo 3: Grupos CART definitivos

#### 1. Base de datos de pólizas normales y atrasadas

En esta base se presentan dos grupos denotados por las letras A y B

##### a. GRUPO A

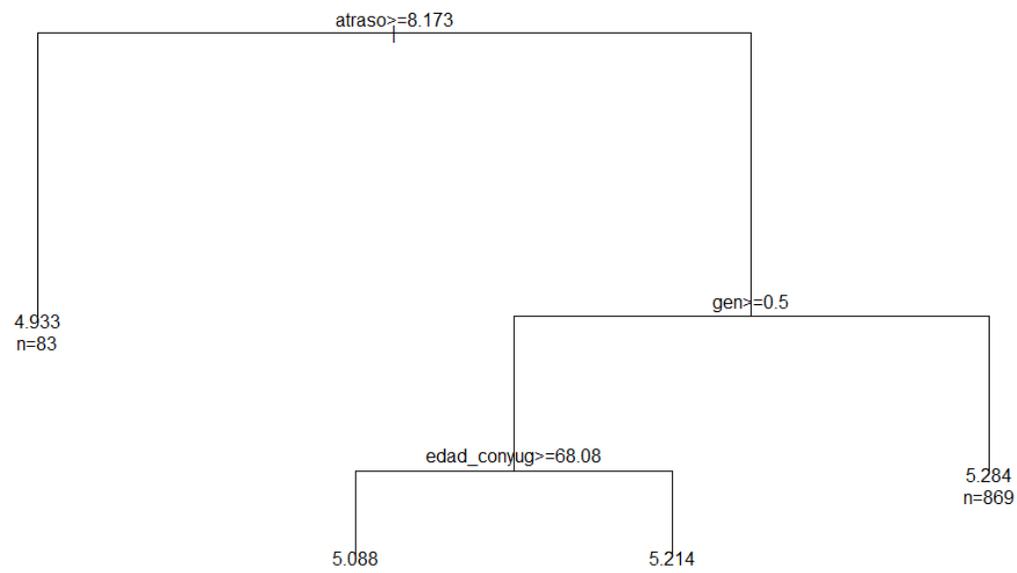


Meses utilizados para el CART: 21

4 grupos generados:

- i. Hombres, RV individual;
- ii. Hombres, RV conjunta;
- iii. Mujeres (RV individual), atraso  $\geq 4.004$  años;
- iv. Mujeres (RV individual), atraso  $< 4.004$  años.

## b. GRUPO B



Meses utilizados para el CART: 4

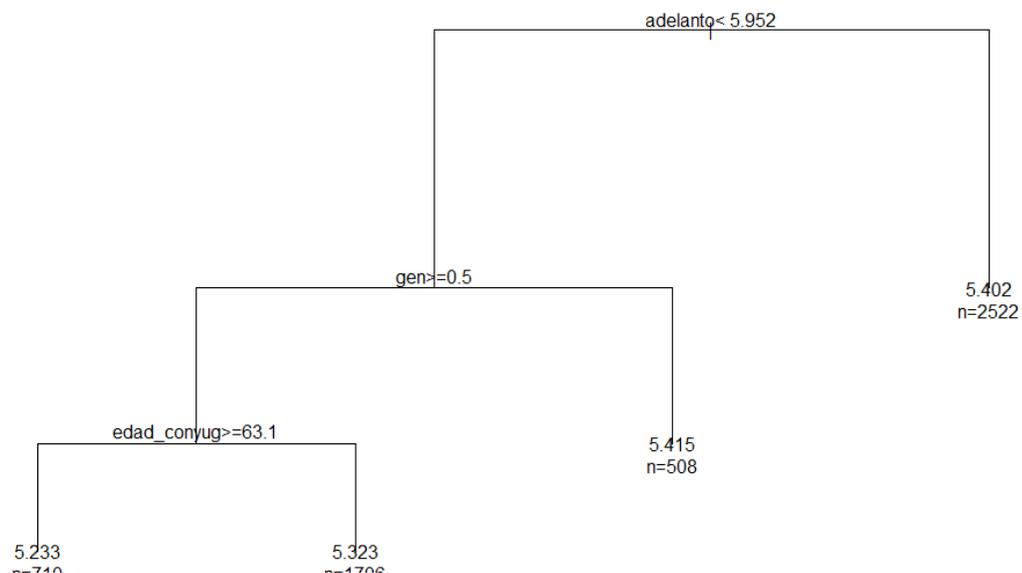
4 grupos generados:

- i. Atraso  $\geq 8.173$  años;
- ii. Atraso  $< 8.173$ , hombres, RV individual;
- iii. Atraso  $< 8.173$ , hombres, RV conjunta;
- iv. Atraso  $< 8.173$ , mujeres.

## 2. Base de datos de rentas anticipadas

En esta base también se presentan dos grupos identificados por A y B.

### a. GRUPO A

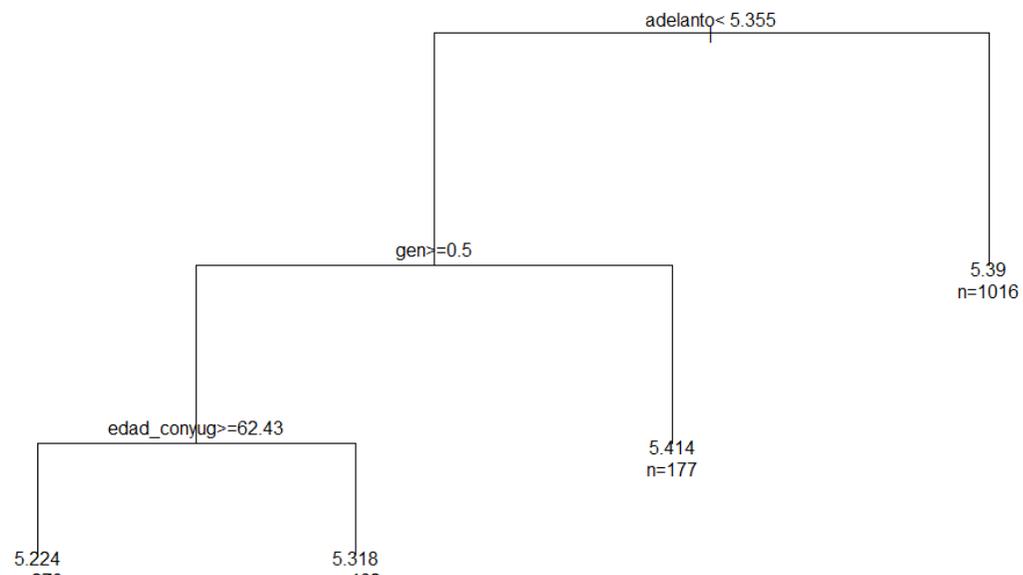


Meses utilizados para el CART: 15.

4 grupos generados:

- i. Adelanto < 5.952 años, hombre, RV conjunta;
- ii. Adelanto < 5.952 años, hombres, RV individual;
- iii. Adelanto < 5.952, mujeres;
- iv. Adelanto  $\geq$  5.952.

## b. GRUPO B



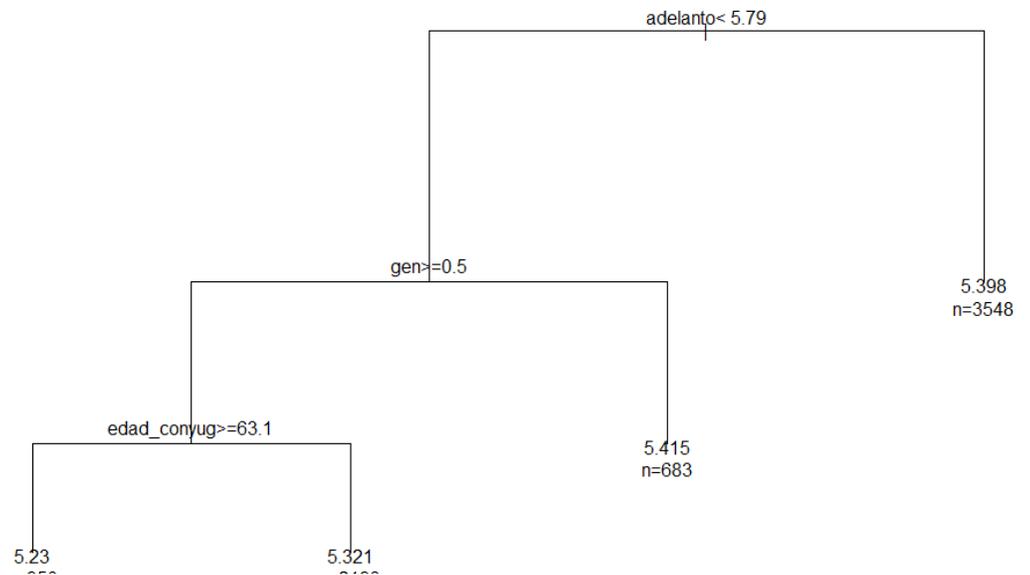
Meses utilizados para el CART: 5.

4 grupos generados:

- i. Adelanto < 5.355 años, hombre, RV conjunta;
- ii. Adelanto < 5.355 años, hombres, RV individual;
- iii. Adelanto < 5.355, mujeres;
- iv. Adelanto  $\geq$  5.355.

## c. GRUPO C

Dado que al utilizar sólo los meses correspondientes al grupo B se obtiene la estructura del grupo A, se corre un nuevo CART con los 20 meses de los grupos A y B. La estructura se mantiene.



Meses utilizados para el CART: 20.

4 grupos generados:

- i. Adelanto < 5.979 años, hombre, RV conjunta;
- ii. Adelanto < 5.979 años, hombres, RV individual;
- iii. Adelanto < 5.979, mujeres;
- iv. Adelanto  $\geq$  5.79.

Las agrupaciones definitivas para las pólizas normales es la del CART A y para las pólizas adelantadas es la del CART C