

ANALES DE ECONOMÍA APLICADA 2016

Número XXX

**DATOS, INFORMACIÓN Y CONOCIMIENTO
EN ECONOMÍA**

**DATA, INFORMATION AND
KNOWLEDGE IN ECONOMICS**

**Santiago Murgui
Jose Manuel Pavía
Alejandro Casino
Belén García-Cárceles
(Coordinadores)**



ANALES DE ECONOMÍA APLICADA

Año 2016 - Número XXX

©2016 ASEPELT

Reservados todos los derechos. El contenido de esta publicación, tanto de la obra escrita como electrónica, puede ser utilizado, de común acuerdo con ASEPELT, para usos exclusivamente particulares y/o profesionales y, en ningún caso, comerciales.

ISSN: 2174-3088

Indexada en:

BSB (Bavarian State Library)

WorldCat (OCLC Online Computer Library Center)

ZBW (German National Library of Economics)

COPAC (Catálogo colectivo Reino Unido)

SUDOC (Catálogo colectivo Francia)

ANÁLISIS DEL RIESGO DE CAÍDA DE CARTERA SEGÚN EL PERFIL DE RIESGO DEL ASEGURADO

JUAN ROCHER MOLIO

Departamento Actuarial
Divina Pastora Seguros

VICENTE COLL SERRANO

Facultad de Economía
Dpto. Economía Aplicada. Universidad de Valencia

e-mail: juan.rocher.molio@gmail.com - Teléfono: (+34) 630 947 100
e-mail: Vicente.Coll@uv.es - Teléfono: (+34) 963 828 640

Resumen

Actualmente el sector asegurador se encuentra ante un nuevo marco regulatorio, Solvencia II, cuyo objetivo principal es mejorar el control y la medición de los riesgos a los que están expuestos las aseguradoras, entre ellos se contempla el riesgo de caída de cartera. En este trabajo se estiman y analizan las tasas de caída de cartera de un seguro de vida-ahorro atendiendo al perfil del asegurado con la finalidad de ajustar las provisiones garantizando así la solvencia de la compañía frente a los asegurados. Para ello, se recurre al análisis de segmentación (algoritmo CHAID) para determinar los distintos perfiles de riesgo y posteriormente se realiza un ajuste analítico de las curvas de caída de los distintos segmentos obtenidos mediante la estimación de modelos lineales generalizados. Se obtienen un total de 9 perfiles de riesgo con sus correspondientes tasas de caída, lo que permitirá realizar una dotación de provisiones más adecuada al nivel de riesgo real de la cartera de la compañía aseguradora.

Palabras clave: Solvencia II, Riesgo de caída, Seguro vida-ahorro, Perfil del asegurado, Análisis de segmentación

Área Temática: Aula de Jóvenes Investigadores.

Abstract

Currently, the insurance sector is facing a new regulatory framework, Solvency II, whose aim is to improve the control and measurement of risks to which insurance companies are exposed to, including the portfolio's lapse risk. In order to adjust the reserves that guarantee the solvency of the insurer against the insured, in this paper we estimate and analyze lapse rates of life-savings insurance depending on the insured profile. For this purpose, we apply the segmentation analysis (CHAID algorithm) to determine the different risk profiles, and then we perform an analytical adjustment of the lapse rates of each segment by estimating generalized linear models. A total of 9 risk profiles with its corresponding lapse rates are identified. It will allow the insurance company to set a better provisioning for its actual risk level.

Key Words: Solvency II, Lapse Risk, life-savings insurance, Insured Profile, Segmentation Analysis.

1. INTRODUCCIÓN

En la actualidad, el sector asegurador se enfrenta a un importante reto que implica un gran cambio en la gestión y cuantificación de los riesgos expuestos impulsado en gran medida por la necesidad de adaptarse al nuevo marco regulatorio de los seguros, Solvencia II.

Según UNESPA (Asociación Empresarial del Seguro), *Solvencia II es una Directiva que cambia las normas europeas del Seguro para reforzar esta industria, que es plenamente solvente, y propiciar mejores productos aseguradores para los ciudadanos. El objetivo principal consiste en mejorar el control y medición de los riesgos (de mercado, operacionales, de crédito y de liquidez) a los que están expuestas las aseguradoras; y se estructura en tres pilares o principios: (i) Medida de activos, pasivos y capital; (ii) Proceso de Supervisión; (iii) Requerimientos de transparencia.*

Uno de los factores de riesgo más importantes a los que están expuestas las compañías aseguradoras es el posicionamiento en el mercado. Las entidades aseguradoras operan en un entorno cada vez más competitivo, donde existe gran cantidad de información que puede tener consecuencias en la fidelidad de los asegurados y, por tanto, en las caducidades de las carteras. Dicho riesgo se materializa en la caída de cartera que se registra cada año y la gran rotación existente de los asegurados que cambian de una entidad aseguradora a otra, alterando las composiciones y calidad de las carteras. (Ayuso-Gutiérrez et al., 2011)

Las tasas de caducidad nos dan la frecuencia con que determinado plan de seguro se cancela o se rescata antes de llegar a su vencimiento debido a diferentes motivos que llevan a los asegurados a tomar tal decisión. La cancelación tiene efectos directos sobre el valor de las provisiones matemáticas correspondientes a una cartera de planes de seguro; por ello, a efectos de proyectar el valor futuro de las provisiones matemáticas de una cartera de planes es necesario contar con tasas de caducidad (Aguilar-Beltrán, 2006).

En este trabajo se estiman y analizan las tasas de caída de cartera de un seguro de vida-ahorro atendiendo al perfil del asegurado con la finalidad de ajustar las provisiones garantizando así la solvencia de la compañía frente a los asegurados. Los distintos perfiles de riesgo se obtienen a partir de las siguientes variables: (i) antigüedad; (ii) edad; (iii) prima.

En la mayoría de los contratos de seguro de vida se prevé la posibilidad de rescate por parte del cliente, es decir, la resolución anticipada total o parcial del contrato con el fin de recuperar su ahorro. Cada vez que la entidad aseguradora vende un producto, modeliza, a través de su experiencia, una determinada tasa de rescate. Solvencia II calcula la carga del riesgo de rescate derivada de situaciones en las que los rescates no se ajustan de forma razonada a esas provisiones. El cálculo de la carga de capital, por lo tanto, tiene en cuenta varias hipótesis: (i) Un aumento permanente de las tasas de rescates en un 50%; (ii) una disminución permanente de las tasas de rescate en un 50% y (iii) un rescate masivo.

Una de las aportaciones principales de este trabajo es la elaboración de escenarios extremos en cada perfil de riesgo para obtener coeficientes de caída extremos (en adelante tasas estresadas) a partir de la experiencia interna de la entidad y determinar un nivel de riesgo acorde a situaciones de rescate extremas.

Para el desarrollo de este trabajo se ha aplicado el análisis de segmentación con el fin de determinar los distintos perfiles de riesgo, el bootstrap para la elaboración de escenarios y posteriormente se realiza un ajuste analítico de las curvas de caída de los distintos segmentos obtenidos mediante la estimación de modelos de regresión binomial.

2. DATOS Y VARIABLES

Los datos utilizados en este trabajo corresponden a una cartera simulada de asegurados de un seguro de vida ahorro y en ella podemos identificar las siguientes variables:

- Rescate: Variable dicotómica (variable dependiente) que toma el valor 1 si se produce un rescate y el valor 0 en caso contrario.
- Edad: Variable cuantitativa que hace referencia a la edad, en años, del asegurado.
- Antigüedad: Variable cuantitativa que indica la antigüedad, en años, que lleva el contrato de seguro en vigor.

- Prima: Variable cuantitativa que se refiere a la cuantía de la prima comercial mensual pagada al inicio del contrato de seguro.

Para poder aplicar el análisis de segmentación ha sido necesario la discretización de las variables cuantitativas en variables cualitativas ordinales. Para ello se ha procedido a su clusterización.

3. METODOLOGÍA

3.1. ANÁLISIS DE SEGMENTACIÓN

La segmentación, como técnica de análisis multivariante, tiene su lugar entre los métodos de dependencia, ya que se distinguen aquellas variables sobre las cuales se quiere explicar un comportamiento (variables dependientes) y aquellas que se utilizan para explicar las anteriores (variables explicativas o predictoras), que son de naturaleza independiente. Por el contrario, el análisis clúster (y otras técnicas), que también se utiliza para obtener grupos homogéneos agrupando elementos (objetos o individuos) a partir de un conjunto de variables consideradas de forma conjunta, se encuadra entre los métodos de interdependencia.

Dado que en este trabajo pretendemos explicar el rescate de un seguro de vida-ahorro (variable dependiente dicotómica), se efectúa un análisis de segmentación jerárquico.

La segmentación jerárquica es una técnica explicativa y descomposicional que utiliza un proceso de división secuencial, iterativo y descendente. Partiendo de la variable dependiente discreta (rescate), se forman grupos homogéneos mediante combinaciones de variables independientes explicativas (edad, antigüedad y prima, discretizadas) en los que se incluyen la totalidad de los casos recogidos en la muestra. Esta técnica también se conoce como “técnicas de árboles”, puesto que utiliza una estructura de representación gráfica similar a la de un diagrama de árbol ofreciendo una visión global y simple de los segmentos obtenidos en el análisis. La raíz del árbol es el conjunto de datos íntegro, los subconjuntos conforman las ramas del árbol. Un conjunto en el que se realiza una partición se denomina nodo.

El árbol de decisión se construye partiendo el conjunto de datos en dos o más subconjuntos de observaciones a partir de los valores que toman las variables explicativas; cada uno de estos subconjuntos vuelve a ser dividido utilizando el mismo algoritmo. Este proceso continúa hasta que no se encuentran diferencias significativas en la relación de las variables predictoras de uno de estos grupos hacia el valor de la variable dependiente. El proceso es secuencial e interactivo ya que recoge no sólo el efecto principal de las variables explicativas sobre la variable a explicar, sino también la influencia que cada variable independiente produce en la variable dependiente en función de los valores que adoptan el resto de variables dependientes contempladas en el análisis.

Uno de los primeros algoritmos empleados en la segmentación jerárquica es el AID (Automatic Interaction Detection) o Detección Automática de Interacciones, creado por Sonquist y Morgan de la Universidad de Michigan a principios de los años 60. Este algoritmo se utilizó extensamente en la década de los setenta y principios de los ochenta hasta que surgió el CHAID (Chi-Squared Automatic Interaction Detection) desarrollado por Kass en 1980 y completado por Magidson en 1993 (Román-González y Lévy-Mangin, 2003).

3.1.1. EL MÉTODO CHAID

El algoritmo que se utiliza en el presente trabajo para el análisis de segmentación es el CHAID. El algoritmo CHAID es considerado como un algoritmo general de segmentación y utiliza el test Ji-Cuadrado para contrastar independencia entre las distintas fases del proceso.

Entre las ventajas del método CHAID cabe destacar:

- Identifica aquellas clases o perfiles de las variables explicativas que no difieren desde el punto de vista estadístico con respecto a la variable dependiente agrupándolas en un mismo nodo.
- Mantiene todas las categorías que sean heterogéneas con respecto a la variable a explicar.
- Posibilita la supresión de variables no significativas de forma segura.
- Contempla las correlaciones entre las variables predictoras. Así, si se selecciona una variable altamente correlacionada con otras, el resto no las considera. Esto no

significa que éstas últimas carezcan de poder predictivo sino que tal valor ya ha sido recogido en la otra variable. Esto supone la unión de variables desde el punto de vista de su impacto explicativo.

- Permite conocer las variables que tienen una fuerte interacción entre ellas.

No obstante, el algoritmo CHAID no está exento de ciertos inconvenientes:

- Se requiere un tamaño muestral elevado ($n = 1.000$ suele ser habitual).
- La necesidad de muchas variables pronosticadoras.
- La carencia de un modelo explícito. Ninguno de los algoritmos anteriores proporciona una función que determine la relación existente entre la variable dependiente y las variables independientes, por lo que muchas veces resulta conveniente su combinación con otras técnicas de análisis multivariante (como por ejemplo técnicas de regresión).
- Las particiones resultantes dependen de la primera variable elegida, lo que condiciona sucesivas particiones.

3.2. REGRESIÓN BINOMIAL

La regresión binomial es un tipo de análisis de regresión utilizado para estimar el resultado de una variable cualitativa en función de las variables predictoras. Los modelos de regresión binomial se describen generalmente en términos del modelo lineal generalizado, un intento de generalizar diversos tipos modelos de regresión lineal. La regresión binomial analiza datos distribuidos binomial de la forma: $Y_i \sim B(p_i, n_i)$, para $i = 1, \dots, m$ donde los números de [ensayos Bernoulli](#) n_i son conocidos y las probabilidades de éxito p_i son desconocidas. El modelo presenta la siguiente formulación:

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i})}}$$

3.3. BOOTSTRAP

Para obtener las tasas de caída estresadas se aplicará un procedimiento de remuestreo conocido como bootstrap (Efron, 1979). Cabe distinguir entre el bootstrap paramétrico, en el que se conoce la función de distribución de la variable que se desea analizar, y el bootstrap no paramétrico, que se desconoce por completo (Álvaro-Jareño y Coll-Serrano, 2012). En nuestro trabajo empírico recurrimos este segundo tipo de bootstrapping.

El bootstrap no paramétrico permite estimar la distribución muestral de un estadístico sin hacer presunciones sobre la estructura de la población. De esta forma, se extrae una muestra de tamaño n de los elementos de $K = \{X_1, X_2, \dots, X_n\}$, muestreando con reemplazamiento. El resultado es la muestra bootstrap: $K_1 = \{X_{11}, X_{12}, \dots, X_{1n}\}$. Esto equivale a tratar la muestra k como una estimación de la población P , de manera que cada elemento X_i de K aparece en la muestra bootstrap con probabilidad $1/n$, tal y como se seleccionó inicialmente la muestra k de la población P . Se repite el procedimiento un número elevado de veces (en general con $b=1.000$ es suficiente), de forma que la muestra de orden b es: $K_b = \{X_{b1}, X_{b2}, \dots, X_{bn}\}$. (Hesterber y otros, 2013).

4. ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

En esta sección se presenta el análisis de los resultados obtenidos al aplicar las metodologías a las que nos hemos referido en el apartado anterior a la cartera simulada de asegurados de un seguro de vida ahorro.

4.1. RESULTADOS ANÁLISIS SEGMENTACIÓN. MÉTODO CHAID.

Tabla 1. Resultados de la discretización de variables.

Variable\Categoría	1	2	3	4
EdadD	14 – 36 años	37 – 46 años	47 – 55 años	56 – 65 años
AntigüedadD	0 – 2 años	3 – 4 años	5 – 6 años	7 – 9 años
PrimaD	45 € - 166€	167€ - 342€	343€ - 524€	525€ - 1000€

Para poder aplicar esta metodología se requiere que las variables predictoras o explicativas (edad, antigüedad y prima) sean cualitativas, por lo que se han discretizado con un análisis clúster. En la tabla 1 se muestran intervalos de clasificación.

Algunas de las categorías se han agrupado puesto que al realizar el análisis de segmentación se han obtenido que no existen diferencias significativas entre ellas. De esta forma, se han agrupado las categorías 1 y 2 de la variable AntigüedadD, las categorías 3 y 4 de la variable EdadD y las categorías 3 y 4 de la variable PrimaD.

En la Figura 1 se presenta el diagrama de árbol final que se ha obtenido al aplicar el algoritmo de CHAID. En este árbol se distinguen un total de 9 nodos de carácter final, nodos que se corresponderán con los 9 perfiles de riesgo obtenidos.

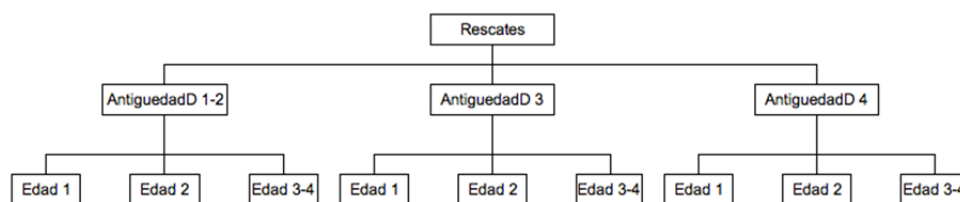


Figura 1. Diagrama de árbol final.

Cabe destacar que la variable con mayor relación con la variable rescate en el análisis (conocida como mejor pronosticador) ha sido la AntigüedadD y la segunda la EdadD. El hecho que la variable PrimaD no haya segmentado no significa que no tenga relación con la variable rescate (recuérdese que una de las ventajas del algoritmo CHAID es que se contemplan las correlaciones entre las variables predictoras), sino que tiene menor relación (menor estadístico Ji-Cuadrado) que la variable AntigüedadD y EdadD, y además que no es significativa en los nodos finales obtenidos (nivel de significación mayor que el establecido en la regla de decisión: 1%).

A continuación se describen las características que tiene un contrato de seguro en cada uno de los segmentos obtenidos:

- **Segmento 1:** Contratos de seguro con una antigüedad entre 0 y 4 años y cuya edad del asegurado está entre 14 y 36 años.
- **Segmento 2:** Contratos de seguro con una antigüedad entre 0 y 4 años y cuya edad del asegurado está entre 37 y 46 años.
- **Segmento 3:** Contratos de seguro con una antigüedad entre 0 y 4 años y cuya edad del asegurado está entre 47 y 65 años.
- **Segmento 4:** Contratos de seguro con una antigüedad entre 5 y 6 años y cuya edad del asegurado está entre 14 y 36 años.
- **Segmento 5:** Contratos de seguro con una antigüedad entre 5 y 6 años y cuya edad del asegurado está entre 37 y 46 años.
- **Segmento 6:** Contratos de seguro con una antigüedad entre 5 y 6 años y cuya edad del asegurado está entre 47 y 65 años.
- **Segmento 7:** Contratos de seguro con una antigüedad entre 7 y 9 años y cuya edad del asegurado está entre 14 y 36 años.
- **Segmento 8:** Contratos de seguro con una antigüedad entre 7 y 9 años y cuya edad del asegurado está entre 37 y 46 años.
- **Segmento 9:** Contratos de seguro con una antigüedad entre 7 y 9 años y cuya edad del asegurado está entre 47 y 65 años.

4.2. RESULTADOS AJUSTE REGRESIÓN BINOMIAL Y TASAS ESTRESADAS

Para poder ajustar analíticamente las tasas de rescate se ha aplicado los resultados del análisis de segmentación al censo mensual de la cartera de los últimos 3 años. Es decir, un total de 36 periodos de observación, que están segmentados con los 9 perfiles de riesgo obtenidos en el apartado anterior.

Por otra parte, para obtener las tasas estresadas, se ha aplicado la metodología del bootstrap donde se han simulado 1000 escenarios, los cuales han permitido obtener las tasas estresadas “up” simulando un shock de aumento en la tasa de rescate (percentil 99,975) y las tasas

estresadas “down” simulando un shock de disminución en la tasa de rescate (percentil 0,025). Dichas tasas han sido ajustadas analíticamente utilizando los modelos de regresión binomial, permitiendo estimar un intervalo de confianza para futuras tasas de rescate y prevenir cuál será la evolución de los rescates en cada uno de los segmentos (ver Figura 2).

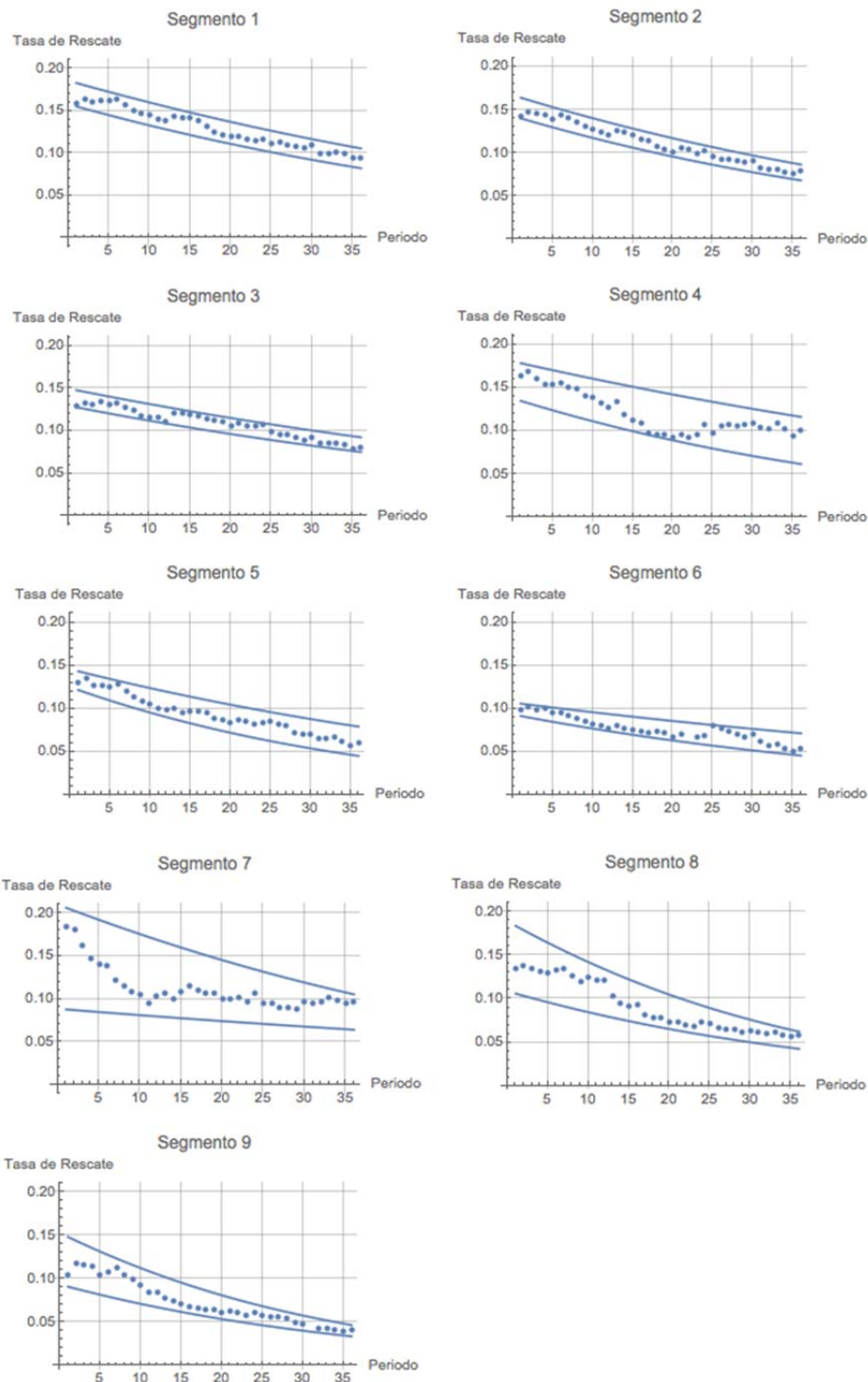


Figura 2. Ajuste Analítico Tasas Estresadas Segmentos: Tasas “up” y Tasas “down”

A partir de la Figura 2 se puede observar que existe una tendencia decreciente en la tasa de rescate para todos los perfiles de riesgo, además como es de esperar, los segmentos que mejor comportamiento tienen en relación con el rescate son los segmentos 9, 8 y 6 (una tasa de rescate actual alrededor del 5%), que coincide con aquellos segmentos que tienen mayor antigüedad (categorías 3 y 4) y edad (categorías 2 y 3-4). En contraposición, los segmentos

que presentan una mayor tasa de rescate son los segmentos 1, 2 y 4 (una tasa de rescate actual alrededor del 10%) coincidiendo con aquellos asegurados que tienen menor antigüedad de contrato (categoría 1-2) y son más jóvenes (categoría de edad 1).

Por otra parte, cabe destacar que para los periodos más recientes las tasas estresadas, tanto la tasa “up” como la tasa “down”, no se distancian mucho de la tasa real de rescate (siendo el shock más destacado la tasa down del segmento 7). De esta forma se garantiza que el shock del 50% que aplica la fórmula estándar de Solvencia II para la carga de capital por el riesgo de caída será más que suficiente (incluso excesivo), garantizando así la suficiencia de las provisiones de la compañía frente a sus obligaciones con la cartera de asegurados.

5. CONCLUSIONES

En este trabajo se ha realizado una clasificación de la cartera según distintos perfiles de riesgo con el objetivo de explicar el comportamiento de los asegurados frente al rescate del contrato. Para ello se ha utilizado el análisis de segmentación jerárquica. Como resultado se ha obtenido una segmentación constituida por un total de 9 segmentos o grupos de riesgo distintos. Sin embargo, resulta difícil explicar el rescate a partir de las variables conocidas (edad del asegurado, antigüedad del contrato, prima pagada, etc.), ya que es habitual que existan otras variables no medibles que tengan mayor relación con la variable rescate. Por ejemplo, situaciones económicas que provocan que el asegurado necesite de forma imprevista liquidez y se vea obligado a cancelar su inversión antes de la fecha de vencimiento establecida. En segundo lugar, también puede deberse a una mejor oferta por parte de la competencia ya que el sector de los seguros es un sector que cada día es más competitivo. Por último, la pérdida de interés por parte del asegurado de cubrir el riesgo puede ser otro motivo de cancelación antes de la fecha de vencimiento.

Respecto a las tasas estresadas obtenidas con el bootstrap, se ha observado que para todos los segmentos se obtienen para el periodo actual tasas estresadas inferiores a las que se aplicarían con el shock propuesto por Solvencia II en su fórmula estándar, permitiendo conocer el nivel de riesgo real de la cartera y aportando una garantía de solvencia por parte de la entidad hacia los asegurados de la cartera.

A pesar de las limitaciones que se ha hecho referencia, para una buena gestión del riesgo de caída de cartera es importante conocer cuáles son los distintos perfiles de riesgo de su cartera. A partir de estos perfiles la entidad aseguradora puede concentrar todos sus esfuerzos comerciales en aquellos segmentos que presentan unas tasas de rescate relativamente bajas (segmentos 9, 8 y 6) y tratar de mejorar el proceso de suscripción y fidelización de los asegurados con un perfil de riesgo de rescate más elevado (segmentos 1, 2 y 4).

Para concluir el trabajo y como futuras líneas de desarrollo, este análisis podría extenderse a otros riesgos contemplados en la nueva regulación de Solvencia II como por ejemplo el riesgo de mortalidad, longevidad, gastos, etc. ya que la proliferación de los modelos de predicción en los seguros de vida poco a poco va convirtiéndose en una realidad.

REFERENCIAS

AGUILAR BELTRÁN, P. (2006): Tasas de Caducidad: Guía de apoyo para la construcción y aplicación”. Asociación Mexicana de Instituciones de Seguros (AMIS).

ÁLVAREZ JAREÑO, J.A. y COLL SERRANO, V. (2012): Estimación de reservas en una compañía aseguradora. Una aplicación en Excel del método Chain-Ladder y Bootstrap. *Métodos Cuantitativos para la Economía y la Empresa*, 14, 124-136.

AYUSO GUTIÉRREZ, M.; GUILLÉN ESTANY, M. y PÉREZ-MARÍN, A. M. (2011): Metodología para el cálculo de escenarios de caída de cartera. *Anales del Instituto de Actuarios Españoles*, 17, 13 – 30.

AYUSO GUTIÉRREZ, M.; GUILLÉN ESTANY, M. y PÉREZ-MARÍN, A. M. (2011): Modelos internos en Solvencia II: Su aplicación al cálculo del coeficiente de caída de cartera. *Gerencia de Riesgos y Seguros*, 112, 38 – 48.

EFRON, B (1979): Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics*. 7(1), 1-26.

HESTERBER, T.; MONAGHAN, S.; MOORE, D.; CLIPSON, A. y EPSTEIN, R. (2013): Bootstrap methods and permutation tests. The practice of Business Statistics. W.H. Freeman and company, New York.

KASS, GORDON V. (1980): An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics*, 29(2), 119–127.

MAGIDSON, J. (1993): The use of the new ordinal algorithm in CHAID to target profitable segments. *The Journal of Database Marketing*, 1, 29–48.

COMISIÓN EUROPEA (2014): Reglamento Delegado (UE) 2015/35 de la Comisión de 10 de octubre 2014. Disponible en: <https://www.boe.es/doue/2015/012/L00001-00797.pdf>

ROMÁN GONZÁLEZ, M. V. y LÉVY MANGIN J. (2006). Análisis multivariable para las Ciencias Sociales. Pearson Educación, Madrid.

SONQUIST, J.A. y MORGAN, J.N. (1964): The Detection of Interaction Effects Survey Research Center. University of Michigan, Ann Arbor.