

**Predicción de fracaso empresarial en
organizaciones de Economía Solidaria: Un análisis
empírico**

Por

Nicolás Romero Castaño

Trabajo de grado para optar al título de

Economista

en la

UNIVERSIDAD DEL VALLE

Noviembre 2018

Autor.....

Departamento de Economía

Noviembre 22, 2018

Aprobada por.....

Inés María Ulloa

Profesora Titular

Tutora

Predicción de fracaso empresarial en organizaciones de Economía Solidaria: Un análisis empírico

por

Nicolás Romero Castaño

Abstract

The bankruptcy literature has focused mostly on for-profit organizations. However, non-profits organizations provide services in difficult access areas –Rural Credit Cooperatives– or special services –such as education or health– and provide loans to partially rationed people in the credit market; In general, they fulfill a complementary function to that of commercial banks. Using the financial statements reported by entities to the regulator between 2002 and 2017, I estimate multi-discriminant models and logit models to provide the more efficient predictive model taking into account the economic environment to identify business failure. I found that both models have similar results and robustness tests provide greater confidence in the results.

KEYWORDS: Corporate Finance, Bankruptcy, Liquidation, Nonprofit Institutions.

JEL CLASSIFICATION: G33, L31, M41.

Resumen

El estudio del fracaso empresarial se ha enfocado mayoritariamente en entidades con ánimo de lucro. Sin embargo, las entidades sin ánimo de lucro prestan servicios en zonas de difícil acceso (Cooperativas de Crédito Rural) o servicios especiales (como educación o salud) y otorgan créditos a personas parcialmente racionadas del mercado de crédito; en general cumplen una función complementaria a la de los bancos comerciales. Utilizando los estados financieros reportados ante la Superintendencia de Economía Solidaria entre 2002 y 2017, se estiman modelos de análisis discriminante múltiple y logit para encontrar el modelo predictivo más eficiente que tome en cuenta el entorno económico para identificar el fracaso empresarial. Se encuentra que ambos modelos tienen resultados similares que se mantienen con las pruebas de robustez.

PALABRAS CLAVE: Finanzas Corporativas, Bancarrota, Fracaso Empresarial, Economía Solidaria, Sin Ánimo de Lucro.

CLASIFICACIÓN JEL: G33, L31, M41.

Tutora: Inés María Ulloa

Título: Profesora Titular

Agradecimientos

Al ser supremo, por darnos la sabiduría para enfrentar cada día.

A mi familia –incluidas mis mascotas–, por ser el apoyo y el sustento de mis días.

*A mis padres, por su comprensión y dedicación. Gracias a ustedes por cada lección,
por cada mañana al despertar, y por enseñarme el valor de la educación.*

A ustedes, por su guía.

*A mi novia, Sofía, compañera del camino. A ti por apoyarme, por darme cariño y
ánimo, y por ser mi mano izquierda cuando necesité ayuda.*

El presente trabajo es un esfuerzo en el cual participaron muchas personas, de forma directa, leyendo, opinando o corrigiendo, y de manera indirecta, dándome ánimo, escuchando mis quejas y siendo un apoyo cuando no encontraba la luz al final.

A mis abuelos, mis tías y primos, por tener paciencia cuando pasaba estudiando durante las reuniones familiares. A mi familia en Medellín que celebra cada uno de mis logros con efusividad.

A mi mentora, Inés María, por creer en mí y en mis capacidades, y por darme guía en el momento que no encontraba el camino.

A la Federación Nacional de Estudiantes de Economía por permitirme participar en el concurso de ponencias “Jesús Antonio Bejarano” y a los evaluadores por sus aportes, que me permitieron mejorar este documento.

A la Universidad del Valle y el departamento de Economía por permitirme hacer parte de esta carrera que forja líderes críticos. Además, le agradezco a cada uno de mis profesores, porque me han forjado no solo académicamente sino también en valores y en carácter.

A Cristhian Larrahondo, Jhon Sebastián Villarreal, Geovany Uribe, Brayan Salgado y Luis Vélez, por cinco años de amistad –y los que vienen–, por su apoyo y por escuchar mis agobios.

A Jhonatan Galeano, Evelyn Orozco, Daniela Rojas, Santiago Mosquera y Camila Rebellón, por estar pendientes de mí y por su invaluable amistad.

Nicolás Romero Castaño

Índice general

1. Introducción	10
2. Contexto	13
3. Metodología	18
4. Datos, Muestra y Variables Relevantes	21
5. Resultados	28
6. Conclusiones	46
A. Estadísticas Descriptivas	51
B. Pruebas de Validación	53

Índice de figuras

1.	Evolución de las Entidades.	23
2.	Coefficiente de Correlación de Pearson Bivariante.	27
3.	Entidades por Tipo.	51
4.	Entidades por Tamaño y Situación.	51
5.	Entidades por Situación y Servicio de Ahorro y Crédito.	52
6.	Entidades por Nivel de Supervisión y Situación.	52

Índice de tablas

1.	Aspectos jurídicos de distintos tipos de organizaciones.	13
2.	Listado de Variables.	22
3.	Estadísticas descriptivas de la muestra.	24
4.	Contraste paramétricos y no paramétricos. Nivel de Significancia. . .	26
5.	Análisis Discriminante Múltiple Lineal.	30
6.	Comparación del grado de error entre modelos. MDA.	30
7.	Regresión Logística. Mejor estimación.	31
8.	Regresión Logística. Mejor predicción.	34
9.	Comparación del grado de error entre modelos. LOGIT.	36
10.	Test de Diferencia de Proporciones entre Modelos. MDA vs Logit. . .	38
11.	Comparación del acierto y error entre el modelo original para cada año y el modelo en t-1 usando la muestra de validación correspondiente a cada año. MDA.	40
12.	Test de Diferencia de Proporciones. Modelo t-1. MDA.	41
13.	Comparación del acierto y error entre el modelo original para cada año y el modelo en t-2 usando la muestra de validación correspondiente a cada año. MDA.	41
14.	Test de Diferencia de Proporciones. Modelo t-2. MDA.	41
15.	Comparación del acierto y error entre el modelo original para cada año y el modelo en t-3 usando la muestra de validación correspondiente a cada año. MDA.	42
16.	Test de Diferencia de Proporciones. Modelo t-3. MDA.	42
17.	Comparación del acierto y error entre el modelo original para cada año y el modelo en t-1 usando la muestra de validación correspondiente a cada año. Logit.	43
18.	Test de Diferencia de Proporciones. Modelo t-1. Logit.	43

19.	Comparación del acierto y error entre el modelo original para cada año y el modelo en t-2 usando la muestra de validación correspondiente a cada año. Logit.	44
20.	Test de Diferencia de Proporciones. Modelo t-2. Logit.	44
21.	Comparación del acierto y error entre el modelo original para cada año y el modelo en t-3 usando la muestra de validación correspondiente a cada año. Logit.	44
22.	Test de Diferencia de Proporciones. Modelo t-3. Logit.	45
23.	Análisis Discriminante Múltiple Lineal. Submuestra Aporte y Crédito.	53
24.	Comparación del grado de error entre modelos. MDA. Submuestra Aporte y Crédito.	53
25.	Regresión Logística con Submuestra Aporte y Crédito. Mejor estimación.	54
26.	Regresión Logística con Submuestra Aporte y Crédito. Mejor predicción.	55
27.	Comparación del grado de error entre modelos. Logit. Submuestra Aporte y Crédito.	56
28.	Test de Diferencia de Proporciones entre Modelos. MDA vs Logit. Submuestra Aporte y Crédito.	57
29.	Análisis Discriminante Múltiple Lineal. Submuestra 2002.	57
30.	Comparación del grado de error entre modelos. MDA. Submuestra 2002.	58
31.	Regresión Logística con Submuestra 2002. Mejor estimación.	58
32.	Regresión Logística con Submuestra 2002. Mejor predicción.	59
33.	Comparación del grado de error entre modelos. Logit. Submuestra 2002.	61
34.	Test de Diferencia de Proporciones entre Modelos. MDA vs Logit. Submuestra 2002.	61

1. Introducción

El análisis de fracaso empresarial ha tomado relevancia en las últimas décadas en la literatura debido a que los niveles de supervisión a las compañías, por parte de los entes de control y de los mismos accionistas, han aumentado. Entre los artículos seminales que estudiaron el riesgo de insolvencia se encuentran el de Altman (1968), quien realizó un análisis discriminante multivariado para encontrar cuáles compañías tenían mayores probabilidades de sufrir de insolvencia, y el de Beaver (1966), quien buscaba las razones financieras relevantes para determinar el fracaso empresarial. Con el devenir de los años los estudios y las metodologías aplicadas han aumentado de forma vertiginosa, buscando siempre la generalización y las variables que mejor predigan el comportamiento de una empresa.

La teoría respecto al fracaso empresarial no está muy extendida debido a los problemas para generalizar los hallazgos, tales como las variables relevantes, las características muestrales, entre otros, lo que en conjunto se ha denominado el paradigma clásico (Mateos, Marín, Marí & Seguí, 2011). De hecho, este problema se amplifica una vez se analizan los estudios sobre las empresas del sector solidario, dada la existencia de características particulares como la naturaleza de la firma (sin ánimo de lucro), la estructura organizacional y la estructura financiera.

Los estudios existentes para el sector solidario se centran en cooperativas, dejando por fuera otros tipos de entidades como los fondos de empleados, las asociaciones mutuales, etc. Con base en lo mencionado, Mateos *et al.* (2011) analizan los alcances y limitaciones que tendría un modelo logit aplicado a cooperativas agrarias. Asimismo, el estudio hecho por Dietrich, Arcelus y Srinivasan (2005) estima dos modelos de predicción preexistentes en la teoría para cooperativas agrarias en Canadá siguiendo la metodología de Análisis Discriminante Múltiple.

Por otro lado, las firmas del sector solidario actúan como intermediarios financieros donde los bancos no llegan –a los sectores rurales y a los sectores más pobres de la zona urbana. Una empresa fracasada genera fallos en el mercado, ya que las personas no podrán acceder al crédito, que, junto al riesgo de pérdida de sus ahorros, ocasionan

efectos negativos en la población (Tafur-Saiden, 2009). En particular, en Colombia ya se han registrado casos en los cuales las firmas solidarias fracasan; se pierden ahorros, se restringe el crédito y se pierde la confianza en el sector no bancario. Entre 1997 y 1999 se presentó la crisis del sector cooperativo, en especial las cooperativas financieras, período en el que hubo pérdida de ahorros por cerca de 650.000 millones de pesos, lo cual afectó, aproximadamente, a 887.000 ahorradores¹.

Por consiguiente, el objetivo de este trabajo es, por medio de la estimación de distintos modelos predictivos, encontrar potenciales variables predictivas que les permitan a los entes de control y a los mismos dueños de las empresas del sector tener un indicativo del riesgo, lo cual permitirá la toma de medidas preventivas para evitar dichos fallos que conduzcan a los efectos negativos mencionados. Además, proveer información respecto a cuál modelo tiene mejores niveles de pronóstico teniendo en cuenta el entorno económico colombiano.

La metodología empleada será la estimación de los dos principales tipos de modelos en la economía, el modelo de Análisis Discriminante Múltiple (MDA, por sus siglas en inglés) y el modelo Logit. Ambos son reconocidos en la literatura gracias a su practicidad en la estimación y a su capacidad predictiva; en particular, el modelo MDA es utilizado, a pesar del supuesto de normalidad de las variables, con resultados satisfactorios que lo mantienen vigente (Jiménez, 1996; Romero, 2013b; Tascón & Castaño, 2012; Zavgren, 1985). Por su parte, el logit permite una estimación más práctica en el sentido que no requiere la normalidad en las variables, ni proporcionalidad entre los grupos (Hair, Black, Babin & Anderson, 2010).

Los datos han sido obtenidos de los estados financieros de las empresas del sector solidario y disponibles en la Superintendencia de Economía Solidaria, para el periodo 2002 - 2017. Las variables explicativas de los modelos han sido tomadas de acuerdo a su popularidad en la literatura (Bellovary, Giacomino & Akers, 2007; Romero, 2013b), la significancia en estudios pasados y la disponibilidad de los datos.

Los resultados obtenidos indican que hay similitudes entre aplicar el modelo MDA o el modelo Logit, teniendo en cuenta también el bajo costo para la estimación.

¹EL TIEMPO. Bogotá D.C. 14, diciembre, 1998. Disponible en <http://archive.is/WSMal>

Se encuentra que el modelo Logit en el periodo $t-1$ presenta niveles de predicción ligeramente superiores, que son opacados por el costo en la predicción del error tipo II al utilizar muestras de periodos distintos. Por lo anterior, se considera que es mejor utilizar la especificación logit en $t-1$ cuando se requiere un modelo que reduzca al mínimo el error tipo I. Finalmente, se encuentra que, aunque los resultados son similares, la mejor opción siempre será obtener la mayor cantidad de información posible, con el mayor número de estimaciones posibles considerando las ventanas de tiempo propuestas a lo largo de la literatura, uno, dos y tres años antes de la quiebra.

El documento se encuentra dividido en seis secciones, siendo la primera esta introducción. La segunda sección trata el contexto relacionado con las empresas del sector solidario en Colombia, resaltando características relevantes que motiven su estudio, así como una breve revisión de literatura asociada a problemas intrínsecos a la determinación del fracaso empresarial. La tercera sección es la metodología. La cuarta sección está dedicada a los datos, la selección de variables y la selección muestral. La quinta sección está dedicada a los resultados de la estimación de los modelos de predicción de fracaso empresarial y a la validación de los modelos presentados. La última sección son las principales conclusiones de la estimación.

2. Contexto

El sector solidario es un sector particular, con características distintivas tanto en su constitución organizacional, como en sus relaciones con el público y sus asociados. En principio, la ley que rige todo el sector solidario es la Ley 454 de 1998, aunque en años anteriores se había legislado para organizaciones específicas – reseñado en la Tabla 1 – esta es la primera regulación de la economía solidaria como sector. En esta ley se dan los principales aspectos de gestión, control y otras disposiciones que regirán a todas las empresas que conforman el sector. Las entidades del sector solidario son empresas sin ánimo de lucro y tienen como objetivo el desarrollo socioeconómico de sus asociados y la comunidad en general.

También, en esta ley se señala que los excedentes obtenidos de la actividad durante el año deberán ser destinados “... a la prestación de servicios de carácter social, al crecimiento de sus reservas y fondos, y a reintegrar a sus asociados...” (Congreso de Colombia, 1998, p. 3).

Tabla 1: Aspectos jurídicos de distintos tipos de organizaciones.

TIPO DE ORGANIZACIÓN	CARACTERÍSTICAS
Cooperativas	Se rige por la Ley 78 de 1988, el Decreto 1482 de 1989 y el Decreto 4588 de 2006. Sociedad Limitada, cada asociado responde hasta el monto de sus aportes. Los excedentes se reparten según el monto de sus aportes. Cuando un socio se retira tiene derecho al reintegro de sus aportes de acuerdo a determinadas pautas legales.
Fondos de Empleados	Se rige por el Decreto 1481 de 1989 y la Ley 1392 de 2010. Están constituidos por trabajadores dependientes, trabajadores asociados o servidores públicos. La Responsabilidad con terceros se dará con la totalidad del patrimonio. Se diferencia de las Cooperativas de Trabajo Asociados en que los empleados deben pertenecer a la misma empresa.
Asociaciones Mutuales	Se rige por el Decreto 1480 de 1989. La Responsabilidad con terceros se limitará al monto de su patrimonio social. Los excedentes no se reparten. Los asociados realizan una contribución periódica llamada cuota social. Cuando un asociado renuncia no tiene derecho a recibir reintegros.

Fuente: Elaboración propia.

Junto con unas características organizacionales particulares respecto a las entidades con ánimo de lucro, las entidades de economía solidaria enfrentan retos en la prestación de sus servicios y la relación con sus asociados. Las entidades de economía solidaria prestan servicios a comunidades en zonas deprimidas o alejadas del casco urbano, realizando un esfuerzo por llevar oferta educativa, de salud, productiva y financiera a estos sectores.

Las restricciones de acceso al crédito formal, tales como el requerimiento de un nivel de activos, la presentación de colaterales y la distancia hasta los centros urbanos, justifican la existencia de entidades alternativas que prestan servicios financieros a esta población con restricciones.

Si bien no todos los procesos de estrés financiero acaban con la disolución de la empresa, el riesgo existente genera desequilibrios en el mercado, poniendo especial atención en las cooperativas con sección de ahorro y crédito, pues son las que llegan a los sectores populares y rurales, donde los bancos no llegan para colocar recursos en manos de la población de bajos ingresos².

Las cooperativas con sección de ahorro y crédito son las entidades con permiso para, no sólo colocar recursos al público, sino también captarlos – de allí la importancia de su correcto funcionamiento para el bienestar social en el mercado de crédito. Sin embargo, las empresas del sector solidario abarcan otras entidades como las cooperativas especializadas, las cooperativas integrales y las asociaciones mutuales, quienes tienen otras características más allá de colocar recursos al público. La disolución de estas organizaciones cierra las posibilidades de acceso a servicios a la población circunscrita.

Aunado a lo anterior, las entidades solidarias que realizan actividades financieras también se enfrentan al riesgo financiero del país, dado su carácter de intermediarias, en especial en países en desarrollo (Braga, Fully Bressan, Colosimo & Bressan, 2006, p. 85). La disolución de estos intermediarios – que atienden población vulnerable y PYMES – causaría efectos negativos en el mercado, debido a que los bancos no tienen incentivos para llegar a dichas poblaciones.

²EL ESPECTADOR. Bogotá D.C. 31, julio, 2017. Disponible en <http://archive.is/2xPjS>

De allí la importancia de estimar cuándo se podrían generar las distorsiones en el mercado y, de esta forma, actuar en consecuencia, evitando que las firmas caigan en procesos de liquidación. Esto para evitar que su población objetivo pierda un derecho adquirido, como es el acceso al crédito y todas aquellas actividades no financieras que realizan estas entidades solidarias.

Así pues, las cooperativas con sección de ahorro y crédito, y los fondos de empleados, son entidades con autorización para colocar y captar recursos del público. Además, la disolución de entidades que prestan otros servicios, más allá de la intermediación financiera, también afecta el bienestar social, en especial aquellas entidades con servicios educativos, de salud y hasta productivos.

Durante las últimas décadas el número de investigaciones empíricas respecto de modelos predictivos ha crecido de manera exponencial; prueba de esto es que Bellovary *et al.* (2007) realizan un análisis de 165 estudios de predicción del fracaso empresarial para mostrar cuáles han sido las variables relevantes usadas en la literatura. A pesar de todos los esfuerzos y las publicaciones, los alcances de dichos resultados se ven minados debido a que estudios posteriores los ponen en tela de juicio, por ende, no se ha extendido una teoría de largo alcance sobre el fracaso empresarial. Al respecto, Mateos *et al.* (2011) mencionan que:

Sin embargo, y a pesar del gran número de contribuciones, no se ha conseguido formular una teoría sobre el fracaso empresarial ni sobre sus factores determinantes. Los resultados de cada investigación, a pesar de estar basados en datos objetivos obtenidos de rigurosos procedimientos cuantitativos, son generalmente insuficientes para, de forma aislada, concluir sobre una cuestión específica relativa al fracaso empresarial, ya que suelen ser replicados o matizados a la luz de trabajos posteriores (Jiménez, 1996). (p.181).

No obstante, los estudios realizados en este campo siguen una especie de manual, tratando de encontrar soluciones a los problemas comunes a los que se enfrentan quienes tratan de modelar una teoría sobre el fracaso empresarial. Esto es lo que Mateos *et al.* (2011) llaman el “paradigma clásico”. Dicho paradigma, engloba cuatro problemas esenciales –a los cuales también hace referencia la literatura (Balcaen &

Ooghe, 2006) – a saber: i) El concepto de fracaso empresarial; ii) La relación inestable de los datos en el tiempo; iii) la selección de la muestra y; iv) Elección arbitraria del criterio de optimización aplicado al modelo para verificarlo, lo que puede ser determinante para dar como válidos modelos que realmente no lo son y viceversa.

El concepto de fracaso empresarial ha sido neurálgico en los resultados obtenidos. Watson y Everett (1999) muestran que la definición del concepto de fracaso empresarial es crítica en el momento de medir resultados en empresas pequeñas.

Romero (2013a) hace referencia a tres definiciones que han sido las más utilizadas en los estudios analizados. La primera definición hace referencia a un lapso en el cual la compañía no está en posición de pagar sus deudas, pero no implica su desaparición. La segunda definición está enmarcada en un auto-reconocimiento del problema de insolvencia, pero tampoco implica, necesariamente, la desaparición de la firma, sino que puede enmarcarse también en un proceso de restructuración. Por último, la tercera definición refleja la determinada “quiebra técnica” que se muestra con un patrimonio neto contable negativo, visión que se relaciona a un entorno económico.

El segundo problema, el de selección de variables es, quizás, uno de los temas más complejos a tratar, puesto que no hay consenso respecto a esto. Algunos estudios usan variables que son significativas en los estudios previos más relevantes, otros estudios optan por realizar pruebas estadísticas y quedarse con variables que reflejen mejor capacidad analítica y mejor capacidad de asociación con los datos; por último, algunos estudios utilizan las variables que tuvieron debido a la disponibilidad limitada de la información (Bellovary *et al.*, 2007; Romero, 2013a; R. . J. . Taffler, 1982; Tascón & Castaño, 2012).

La estabilidad de la relación entre los datos y las variables resulta de vital importancia para la creación de una teoría del fracaso empresarial duradera, sin embargo, los resultados de algunas variables dependen de los datos que hayan sido seleccionados. Ante esto, Mateos *et al.* (2011) aseguran que:

La no estacionalidad y la inestabilidad de los datos generan consecuencias severas en los modelos predictivos del fracaso empresarial (Balcaen y Ooghe, 2006).

En primer lugar, porque sus capacidades predictivas resultarán pobres cuando

se apliquen sobre potenciales muestras futuras. En segundo término, porque la robustez del modelo será deficiente cuando exista inestabilidad en los datos utilizados. (p. 195).

En términos de selección de la muestra, Enguítanos (1994) ha señalado limitantes a la hora de dicho proceso: i) La proporción de empresas sanas y fallidas, ii) La aleatoriedad de la muestra y iii) El horizonte temporal de la muestra. Esto debido a que debe evitarse la sobrerrepresentación de empresas de algún tipo; además, se hace relevante la aleatoriedad en la selección, puesto que, limitaciones en la obtención de las firmas pueden sesgar los resultados en favor de la habilidad predictiva (Pacey & Pham, 1990).

El caso de las entidades del sector solidario es particular puesto que tienen una estructura de gobierno corporativo y restricciones sobre el lucro, que son diferentes frente a las entidades mercantiles. Dietrich *et al.* (2005) consideran fracasadas aquellas entidades que se encuentren listadas como tal por la entidad reguladora o que no hayan presentado sus estados financieros durante los últimos tres años consecutivos; con esto realizan un modelo de predicción de bancarrota de cooperativas agrícolas en Canadá. Mientras que Braga *et al.* (2006) definen como fracasadas aquellas cooperativas que estén obligadas a mantener el 35 % de sus depósitos como reservas y realizan un modelo de predicción de fracaso en cooperativas brasileñas siguiendo una metodología no paramétricas.

Por último, en términos de la selección de la muestra, ambos obtienen una muestra que depende de la información disponible al momento de realizar dicho estudio, en particular el de Dietrich *et al.* (2005), debido a que se realiza para una región pequeña al interior de Canadá y, por ende, dependen de los datos disponibles.

Como se ha podido observar, el paradigma clásico presenta una dificultad a la investigación en este campo de la predicción de bancarrota, pero también da líneas de base sobre las preguntas que deben ser respondidas para tratar con dichos problemas. Esto debido a que es un problema transversal a toda la elaboración del modelo de predicción.

3. Metodología

Dada la importancia que tiene la predicción del fracaso empresarial y teniendo en cuenta las dificultades planteadas para definir una teoría sobre dicho tema, la literatura relacionada ha tenido su más grande expansión en la definición, crítica y mejoramiento de modelos predictivos.

A lo largo de la literatura han aparecido distintos modelos que han sido la frontera, mientras encuentran alguna restricción en su modelación o se define un modelo que tiene mejor respuesta empírica en términos predictivos (de Llano, Piñeiro & Rodríguez, 2016). Se pueden distinguir dos grandes vías de desarrollo de los modelos de fracaso empresarial: los modelos estadísticos clásicos o modelos paramétricos, y los modelos no paramétricos. Dentro de los modelos estadísticos algunos de los más conocidos son los modelos de análisis discriminante univariado y multivariado, y la regresión logística, mientras en los modelos no paramétricos, están los modelos heurísticos como las redes neuronales artificiales, las máquinas de soporte vectorial, entre otros (Romero, 2013a). Por último, los modelos DEA (Data Envelopment Analysis) han sido destacados en la literatura, debido a que son populares para otras evaluaciones en los negocios y, al ser un método no paramétrico, permite flexibilizar los supuestos de distribución y proporcionalidad para la estimación (Premachandra, Bhabra & Sueyoshi, 2009; Sueyoshi & Goto, 2009).

Jiménez (1996) señala que la superioridad predictiva de los modelos no paramétricos, en especial los que utilizan técnicas avanzadas están en duda, ya que las mejoras predictivas se ven opacadas por los costos de información y la complejidad técnica que revisten para su estimación. Algunos autores que han evaluado estas técnicas frente a técnicas estadísticas más sencillas como el Logit o el Análisis Discriminante Múltiple (MDA por sus siglas en inglés) han encontrado resultados similares, o si se encuentran mejoras predictivas, surge entonces la duda de si el costo de estimación es lo suficientemente alto como para dilucidar los beneficios (Altman, Marco & Varetto, 1994; de Llano *et al.*, 2016; Mateos *et al.*, 2011; Tascón & Castaño, 2012). Por lo anterior, en este documento se utilizarán solo los modelos MDA y Logit, con el fin de

comparar la efectividad en la predicción, teniendo como base un bajo costo técnico de estimación.

“The classic statistical failure prediction models involve a classification procedure to classify firms as failing or non-failing” (Balcaen & Ooghe, 2006, p. 65). Los modelos estadísticos se fundamentan en el desarrollo de un punto óptimo de clasificación y en la reducción de los errores de clasificación tipo I y tipo II. Los errores de clasificación tipo I son aquellos en los que una firma que está en riesgo de insolvencia se clasifica como sana, mientras el error tipo II clasifica a las firmas sanas como firmas quebradas (Hair *et al.*, 2010).

Los modelos estadísticos paramétricos han sido desarrollados desde el trabajo pionero de Beaver (1966), quien desarrolló un modelo discriminante univariado, en el cual se estima un punto de corte óptimo para cada ratio considerado, con el fin de considerar a las firmas en saludables o en quiebra; según dicho punto de corte y se determina la significancia de dicho ratio. El objetivo de este trabajo no era el hallazgo de un modelo predictivo óptimo, sino la demostración del potencial de los datos contables. Sin embargo, de acuerdo con Enguítanos (1994), esta metodología está restringida por la suposición de que todas las variables en el modelo tienen una relación lineal con la variable de fracaso empresarial.

Altman (1968) en contraposición, mejoró el análisis univariado y desarrolló un modelo multivariado o Z-score. El modelo de MDA, se fundamenta en una función discriminante multivariada con ratios relevantes y se busca hallar un punto de corte para clasificar a las firmas, a partir de la combinación lineal de dichos ratios (R. J. Taffler, 1983). La debilidad de este modelo radica en que i) atribuye una distribución normal a los variables, supuesto difícil de cumplir teniendo en cuenta que son datos de corte financiero (Deakin, 1976); ii) Asume el error de clasificación a priori (*prior probabilities*) lo que sesga el resultado (Edmister, 1972); y iii) la efectividad del modelo para predecir es mayor en validación *in-sample*, es decir, predecir sobre la muestra utilizada para estimar, y en un periodo corto en validación *out-sample*, es decir, predecir bancarrota en entidades que no se encontraban en la muestra de estimación (Altman, 1978).

Para enfrentar los problemas restrictivos del modelo MDA y dado el carácter dicotómico de la variable dependiente, se ha extendido en la literatura el uso de modelo de regresión logística (Ohlson, 1980). En esta, se asume que los datos siguen una distribución logística y, a diferencia de los MDA, no clasifica las firmas en fallidas o sanas, sino que estima una probabilidad de que el evento suceda –que una firma caiga en insolvencia– (de Llano *et al.*, 2016); aunado a que, en la estimación Logit, no se requiere una muestra proporcional de empresas fallidas y sanas (Hair *et al.*, 2010). El principal problema de estos modelos es la alta probabilidad de que la estimación tenga multicolinealidad, esto debido a la alta correlación entre las razones contables (Gómez, de la Torre & Román, 2008).

Sin embargo, los dos métodos han tenido gran desarrollo a lo largo de la literatura por su bajo costo técnico en la estimación. El modelo MDA ha tenido diversas variantes como el modelo MDA no estacionario (Mare, Moreira & Rossi, 2017) y mejoramientos en la interpretación de la probabilidad (Lepetit & Strobel, 2015); mientras que los modelos Logit han sido presentados con los logit anidados y logit multinomial, para el caso donde se hace explícita la zona de ignorancia y se utiliza un rango intermedio.

Lo anterior surge, en general, para la predicción de empresas del sector industrial o de servicios, es decir, con ánimo de lucro. Para la predicción de insolvencia de firmas del sector solidario, especialmente las cooperativas de crédito, el espectro es más reducido, a pesar de ello, se ha encontrado la estimación de un modelo de análisis discriminante múltiple (Dietrich *et al.*, 2005), un modelo *Cox Proportional Hazard* (Braga *et al.*, 2006) y un modelo DEA (Pille & Paradi, 2002).

4. Datos, Muestra y Variables Relevantes

Para la estimación de los modelos de predicción de fracaso empresarial se usarán los estados financieros de las firmas del sector solidario, provistos por la Superintendencia de Economía Solidaria con corte a diciembre de cada año. El horizonte temporal considerado será entre 2002 y 2017, de acuerdo a disponibilidad de los datos.

Con relación a la construcción de la variable dependiente categórica de fracaso empresarial, se sigue a Dietrich *et al.* (2005) quienes la definen como la declaración de liquidación –voluntaria o forzosa– por parte del ente regulador (en este caso la Supersolidaria) o no presentar estados financieros al ente regulador durante al menos tres años.

Para la construcción de las variables explicativas, siguiendo la revisión de Bellovary *et al.* (2007), Tascón y Castaño (2012) y Romero (2013b), además, de acuerdo a la disponibilidad de los datos obtenidos, se construyen 27 ratios contables y 2 variables de control (Tabla 2). Siguiendo los factores más relevantes en la literatura se construyen: i) ratios de Solvencia y Liquidez (SL), que permitan observar si la entidad puede responder a choques de corto plazo; ii) ratios de Endeudamiento (E) que permitan observar la carga de deuda que tiene la firma y, junto con iii) ratios de Cobertura y Apalancamiento (CA), sensibilizar el modelo si hay existencia de sobre costos en los Gastos y si la firma, al igual que con SL, puede responder a las obligaciones pactadas; iv) ratios de Rentabilidad (RE) que devalen si hay un incentivo similar al de las firmas con ánimo de lucro, además relacionar el ciclo de los negocios con la quiebra; y, v) ratios de Rotación (R), que señalan el flujo de generación de ingresos con respecto a los costos de mantener la operación. Finalmente, se describen las variables de control, un indicador por tamaño y un indicador de si la entidad presta servicio de aporte y/o crédito, también se utilizan los activos como control.

La limpieza de la base de datos parte, en principio, de la eliminación de las cooperativas y precooperativas de trabajo asociado, una vez que se presenta un choque exógeno por el artículo 63 de la ley 1429 de 2010 que prohíbe su uso para intermediación laboral, por lo que varias entidades –según se observa en la base– decidieron

Tabla 2: Listado de Variables.

Nombre	Fórmula
SL1	Activo Corriente/Activo
SL2	(Activo Corriente – Inventario)/Activo
SL3	(Activo Corriente – Inventario)/Activo Corriente
SL4	Disponible/Activo Corriente
SL5	Activo/Pasivo
SL6	Disponible/Activo
SL7	Disponible/Pasivo
SL8	Cuentas por Cobrar/(Activo – Pasivo)
E1	Patrimonio/Pasivo
E2	(Obligaciones Financieras + Cuentas por Pagar)/Activo
E3	Gastos no Operacionales/Pasivo
E4	Gastos no Operacionales/Activo
E5	Pasivo/(Patrimonio + Pasivo)
RE1	Utilidad Operacional/Activo
RE2	Utilidad Operacional/Ingresos Operacionales
RE3	Utilidad Neta/Activo
RE4	Utilidad Operacional/Patrimonio
RE5	EBIT/Activo
RE6	Utilidad Neta/Patrimonio
CA1	Pasivo/Activo
CA2	Pasivo/Patrimonio
CA3	Patrimonio/Activo
CA4	Capital Social/Patrimonio
CA5	Reservas/Patrimonio
CA6	(Obligaciones Financieras + Cuentas por Pagar)/Patrimonio
R1	Ingresos Operacionales/Pasivo
R2	Ingresos Operacionales/Activo
APORTE Y CRÉDITO	1 si la Firma presta servicios de Ahorro y/o Crédito. 0 en otro caso.
SIZE ^a	1: Microempresa si Activos < 500 SMMLV 2: Pequeña Empresa si $500\text{SMMLV} \leq \text{Activos} \leq 5000\text{SMMLV}$ 3: Mediana Empresa si $5000\text{SMMLV} \leq \text{Activos} \leq 30000\text{SMMLV}$ 4: Gran Empresa si Activos > 30000SMMLV

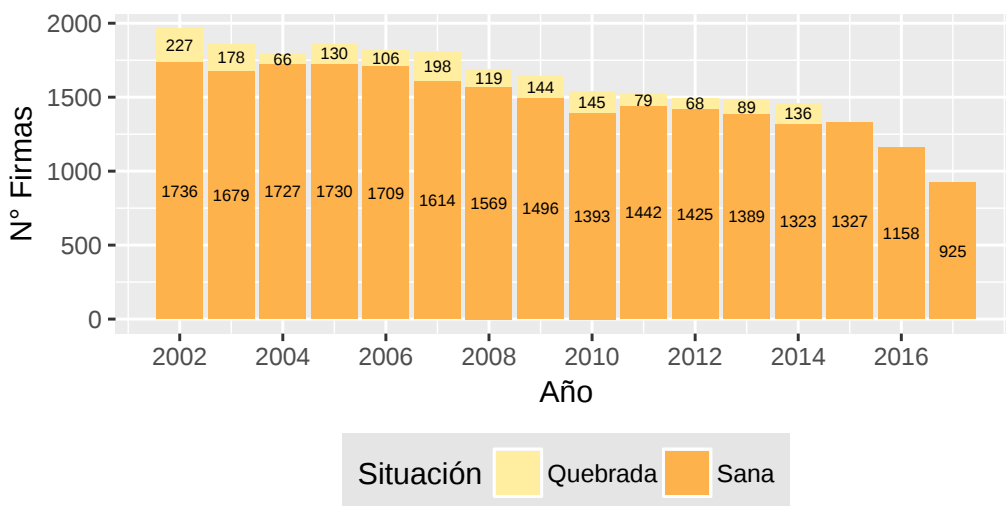
^a La variable de tamaño se construye siguiendo lo erigido por la ley 905 de agosto 2 de 2004, disponible en línea: http://www.secretariasenado.gov.co/senado/basedoc/ley_0905_2004.html

Fuente: Elaboración propia.

salir o quebraron debido a esto. Asimismo, se eliminan los organismos de representación, como las confederaciones de cooperativas, las asociaciones cooperativas, los organismos de carácter económico y las instituciones especializadas auxiliares, puesto que representan entidades de segundo piso, es decir, asociaciones de cooperativas, o

entidades con un manejo financiero común que no tiene mayor incidencia en el fracaso empresarial. Con respecto a los datos, se han eliminado las entidades que presenten al menos un año con activo, pasivo, patrimonio, ingresos, ingresos operacionales o gastos³ con ceros y unos, así como aquellas que presenten crecimientos de más del 1000 % entre un año y otro en alguna de las cuentas principales. También, se han eliminado aquellas firmas donde la igualdad contable Activo = Pasivo + Patrimonio no se mantiene. Finalmente, siguiendo a Romero (2013a), se eliminan las entidades con patrimonio negativo y que estén sanas, puesto que entorpece el algoritmo de predicción y no representa un caso común. Después de realizar esta limpieza, la base de datos cuenta con 25327 observaciones, 3088 entidades, de las cuales 1686 se reportan en quiebra (Figura 1).

Figura 1: Evolución de las Entidades.



Fuente: Elaboración propia.

Las estadísticas descriptivas de la muestra (Tabla 3) señalan que en el periodo hay más firmas quebradas que sanas. Sin embargo, hay que recordar que no todas las organizaciones quiebran en el mismo momento y que esto, al observarse en conjunto el gráfico anterior, señala la importancia de hacer una selección muestral para la estimación predictiva. También, se aprecia que las entidades pequeñas y con niveles de supervisión bajos son las que más quiebran, en particular debido a que no tiene la

³Cuentas principales de los estados financieros (Balance general y Estado de Resultados).

obligación de presentar estados financieros con la misma regularidad que firmas con altos niveles de supervisión. Además, en general, son estas entidades que presentan niveles bajo de cobertura ante las crisis. En el Anexo A se presentan otras estadísticas descriptivas de carácter general, que permiten observar cómo se encuentra distribuida la muestra según características relevantes, así como corroborar la información contenida en la Tabla 3.

Tabla 3: Estadísticas descriptivas de la muestra.

	Total	Sanas	Quebradas
Número de Firmas	3088	1402	1686
Tamaño			
Microempresa	1875	589	1286
Pequeña	1636	988	648
Mediana	431	356	75
Grande	90	72	18
Nivel de Supervisión			
Nivel 1 (Alto)	139	112	27
Nivel 2 (Medio)	315	280	35
Nivel 3 (Bajo)	2658	1096	1562
Aporte y Crédito			
Si	1554	908	646
No	1522	507	1015

Fuente: Elaboración propia.

Continuando con el análisis, el paso siguiente es determinar la muestra para estimación. Considerando que las firmas cuyos estados financieros estén reportados hasta el año 2014 aparecen como fracasadas, dividir la muestra en dos periodos provocaría sobrerrepresentación de firmas quebradas en la muestra de estimación y de firmas sanas en la muestra de validación. Por lo tanto, se optó por subdividir la totalidad de la base de datos en dos muestras siguiendo un proceso aleatorio. Esto a su vez, aumenta ligeramente el sesgo en favor de la predictibilidad del algoritmo, puesto que el entorno económico se mantendrá constante en ambas muestras, sin embargo, debido al costo que tiene la sobrerrepresentación para la predicción, es un riesgo que debe asumirse. También se opta por separar la muestra en t-1, t-2 y t-3 años antes de la quiebra, con el fin de obtener el mejor modelo predictivo.

Una vez se tiene lo anterior y teniendo en cuenta que no existe una teoría definida

respecto a especificar cuáles variables deben utilizarse en el análisis, el siguiente paso es determinar cuáles variables de la batería propuesta son relevantes. Para ello, se propone realizar el análisis de la varianza (ANOVA) con el fin de determinar cuáles variables presentan medias distintas según el grupo (Fracasadas o Sanas). Sin embargo, este test requiere que la variable siga una distribución normal, pero, dado el tamaño de la muestra, se puede proceder al análisis bajo el Teorema Central del Límite. Asimismo, se presentan estadísticos no paramétricos como la U de Mann-Whitney-Wilcoxon y el test Kolmogorov-Smirnov que le darán refuerzo al estadístico paramétrico.

Los resultados obtenidos en la Tabla 4 muestran que hay varias variables que no son significativas para el análisis. Con el fin de obtener un acercamiento a las variables con más capacidad predictiva, sin minar la especificación del modelo, se han eliminado de la estimación las variables que no presenten significancia en al menos dos de los tres estadísticos. En este caso, la variable Capital Social/Patrimonio no es significativamente diferente entre grupos en ningún periodo de análisis, así mismo, se señala que para t-1 las variables Disponible/Pasivo y Utilidad Neta/Activo tampoco son significativas, y en t-2 son Pasivo/Activo, Patrimonio/Activo y Pasivo/(Patrimonio + Pasivo) las que no son relevantes para el análisis.

Con el fin de obtener la especificación de los modelos con capacidad predictiva que no caigan en la sub-especificación ni el sobre-especificación y evitando la multicolinealidad, se ha propuesto el coeficiente de correlación de Pearson bivalente, con el fin de conocer cuál es el grado de asociación entre las variables. La figura 2 presenta los resultados de la estimación de dicho coeficiente, agrupando por clústeres de alta correlación positiva (en azul) y de alta correlación negativa (en rojo). Es importante señalar, que todas las correlaciones en los clústeres son significativas al 5 %.

Se observa que las correlaciones positivas se dan entre denominadores similares, es decir que los numeradores son menos volátiles que las tres cuentas principales (activo, pasivo y patrimonio), por lo que se generan estos clústeres de correlación positiva. En contraposición, la correlación negativa de la variable de utilidad neta/patrimonio se da con respecto al clúster de correlación positiva cuyo denominador es el patrimonio

Tabla 4: Contraste paramétricos y no paramétricos. Nivel de Significancia.

	ANOVA				MANN-WHITNEY- WILCOXON				KOLMOGOROV- SMIRNOV			
	Total	t-1	t-2	t-3	Total	t-1	t-2	t-3	Total	t-1	t-2	t-3
ACTIVO ^d ^a	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SL1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SL2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SL3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SL4	0.08	0.18	0.00	0.57	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SL5	0.00	0.19	0.09	0.06	0.00	0.06	0.16	0.05	0.00	0.00	0.03	0.04
SL6	0.00	0.00	0.00	0.00	0.00	0.73	0.04	0.26	0.00	0.00	0.00	0.03
SL7	0.02	0.23	0.15	0.14	0.00	0.68	0.02	0.08	0.00	0.01	0.00	0.01
SL8	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R1	0.01	0.16	0.13	0.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R2	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CA1	0.07	0.31	0.35	0.14	0.00	0.06	0.16	0.05	0.00	0.00	0.03	0.04
CA2	0.06	0.24	0.86	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CA3	0.07	0.31	0.35	0.14	0.00	0.06	0.16	0.05	0.00	0.00	0.03	0.04
CA4	0.13	0.93	0.93	0.67	0.48	0.48	0.93	0.54	0.00	0.00	0.00	0.00
CA5	0.11	0.36	0.65	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CA6	0.02	0.00	0.21	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
E1	0.00	0.19	0.09	0.06	0.00	0.06	0.16	0.05	0.00	0.00	0.03	0.04
E2	0.06	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
E3	0.02	0.20	0.19	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
E4	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
E5	0.07	0.31	0.35	0.14	0.00	0.06	0.16	0.05	0.00	0.00	0.03	0.04
RE1	0.18	0.00	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RE2	0.11	0.00	0.05	0.00	0.00	0.75	0.07	0.02	0.00	0.00	0.00	0.00
RE3	0.08	0.13	0.00	0.02	0.00	0.23	0.93	0.03	0.00	0.00	0.00	0.00
RE4	0.15	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RE5	0.16	0.00	0.00	0.00	0.31	0.35	0.14	0.00	0.00	0.00	0.00	0.00
RE6	0.10	0.08	0.44	0.56	0.12	0.63	0.07	0.00	0.00	0.00	0.00	0.00

^a Activo deflactado con IPC base 2008.

Variable de agrupación: Situación (Sana o Fracasada); $p \leq 0,05$.

Fuente: Elaboración propia.

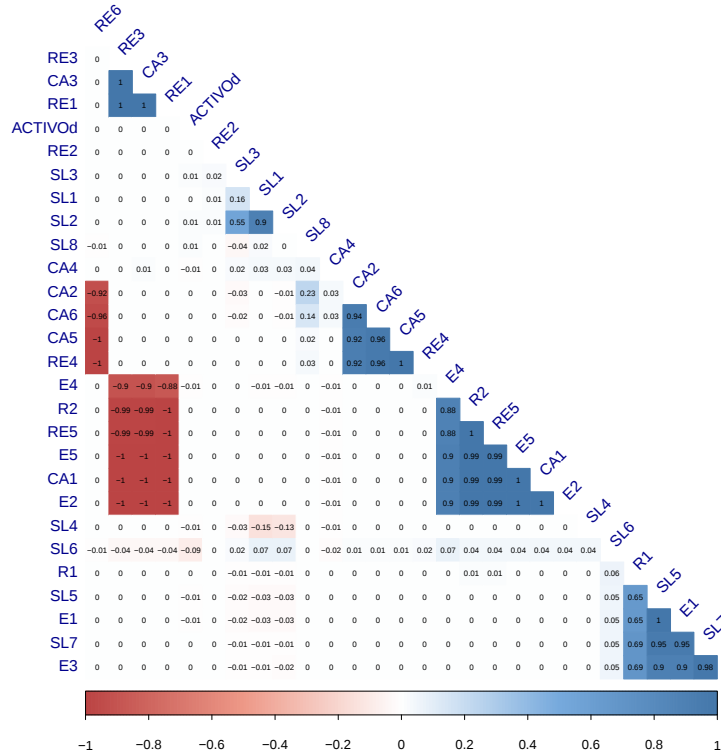
(Pasivo/Patrimonio, Reservas/Patrimonio, Utilidad Operacional/Patrimonio y (Obligaciones Financieras+CxP)/Patrimonio), lo que indica que las utilidades netas van en oposición a los numeradores del clúster de correlación positiva. Situación similar que se presenta con las variables cuyo denominador es el activo (Patrimonio/Activo, Utilidad Operacional/Activo y Utilidad Neta/Activo) frente a otro clúster de correlación positiva con el activo como denominador, lo que indica el movimiento opuesto

entre numeradores.

Para definir la especificación más apropiada, en cada modelo se decide si se deja la especificación completa, sólo con la significancia de la ANOVA o si se le añaden restricciones eliminando variables altamente correlacionadas, de acuerdo a su interés económico para la valoración predictiva. En particular porque lo más relevante es si la variable contribuye a la predicción, la cual siempre es el objetivo principal de este documento, por ende, no debe sobreestimarse el efecto de la multicolinealidad (Lev, 1978).

Finalmente, teniendo todas las variables relevantes en cada modelo, se procede a la estimación del MDA y el logit teniendo en cuenta las variaciones en t-1, t-2 y t-3 años antes de la quiebra.

Figura 2: Coeficiente de Correlación de Pearson Bivalente.



Fuente: Elaboración propia.

5. Resultados

Antes de iniciar con la presentación de los resultados debe hacerse hincapié en que, el objetivo principal de este documento es la estimación de modelos predictivos, por lo tanto, el esfuerzo principal no radica en la estimación del modelo más eficiente en términos de estimación, ni en la estimación de variables significativas, sino en el modelo que permita obtener el menor nivel de error en la clasificación; por lo tanto, algunas variables no son significativas en el análisis, sin embargo, su eliminación implica la reducción del poder predictivo del modelo (Lev, 1978).

Otro aspecto importante a tener en cuenta es que, la muestra elegida es más grande de lo que normalmente se trabaja en la literatura, un esfuerzo que se ha realizado para obtener lo más cercano a la realidad colombiana, es decir, que capture de forma implícita el entorno económico que rodea al sector. Este hecho, en general debería aumentar la precisión de las estimaciones, sin embargo, también dispersa los resultados puesto que hay más probabilidades de encontrar firmas que salieron por choques no observables en los datos, por lo que los resultados también se pueden sesgar.

La Tabla 5 presenta la función discriminante que ha sido estimada en cada uno de los periodos y sus respectivos coeficientes. Para esta estimación se ha utilizado un método de repeticiones que consiste en dividir la muestra en diez partes iguales, estimar con nueve y testear con la última división, luego se iteró diez veces con el fin de probar cada una de las divisiones, esto se repitió diez veces y con esto se obtienen los resultados promedio; la finalidad es reducir el error de clasificación. Asimismo, se observan las probabilidades pre-estimación que son obtenidas como proporción de los datos según la situación, lo que da a entender que, en los tres casos, hay más de un 50 % de probabilidades de obtener una firma sana, frente las 47.5 % en t-1 y el 38-39 % en t-2 y t-3. En los tres modelos se tiene que la mayoría de variables de liquidez tienen signo negativo, lo que indica que mayores niveles de estas disminuyen el riesgo de que la firma sea clasificada como fracasada, excepto en la variable Cuentas por Cobrar/(Activo – Pasivo) indica que las firmas requieren de suficiente efectivo para

sus transacciones, es decir, la liquidez es decisiva para las entidades. La variable Reservas/Patrimonio con signo negativo indica que las reservas son un indicador importante de la estabilidad de la entidad. La variable (Obligaciones Financieras + Cuentas por Pagar)/Activo con signo positivo indica que el endeudamiento desmedido aumenta el riesgo de clasificarse como entidad quebrada. Las variables de rentabilidad tienen signos más dispersos lo que se esperaba, puesto que estas entidades tienen otros objetivos y la rentabilidad es usada, por ejemplo, en reinversiones. Finalmente, con respecto a las variables de control, se obtiene que las firmas que prestan servicios de Aporte y Crédito tienen menor probabilidad de clasificarse como quebrada, lo que tiene relación con que estas entidades, por lo general, son supervisadas con mayor rigurosidad; y se observa que las firmas más pequeñas tienen mayor riesgo de ser clasificada como fracasadas frente a las firmas más grandes.

Ahora bien, con respecto a la capacidad de predicción de los modelos anteriores, se presenta la Tabla 6, donde se realizan dos predicciones: i) *in-sample* y ii) *out-sample*. Los resultados muestran que la efectividad es en promedio superior al 70 %, el coeficiente Kappa⁴ se encuentra entre un 41-46 % en la muestra de estimación y entre el 39-50 % en la muestra de validación, los errores tipo I (clasificar como sana una entidad que está quebrada) son del orden del 26-38 % en la muestra de estimación y del 27-35 % en la de validación; finalmente, los errores tipo II (clasificar como quebrada una entidad que está sana) son del orden del 21-26 % en la muestra de estimación y del 20-26 % en la muestra de validación. Es relevante señalar que los resultados son mejores entre más cerca esté la muestra del suceso (Gómez *et al.*, 2008).

En el caso de la estimación del modelo logit⁵, se han realizado dos estimaciones, la primera con el fin de obtener el mejor modelo en términos de estimación, buscando el modelo más parsimonioso posible, pero que mantenga una capacidad predictiva buena; y en segundo lugar, se ha estimado el mejor modelo predictivo, a pesar de que este tenga problemas de multicolinealidad, signos no esperados y coeficientes

⁴El coeficiente Kappa ajusta el porcentaje de efectividad añadiendo la probabilidad de que cada uno de los eventos que representa éxito ocurra. Valores entre 0.2 – 0.4 representan efectividad moderada, pero lejana; entre 0.4 – 0.6 efectividad moderada; 0.6 – 0.8 efectividad suficiente y mayores niveles, efectividad casi perfecta (Landis & Koch, 1977).

⁵Para obtener los *odds ratio* basta con hacer exponente base Euler del coeficiente.

Tabla 5: Análisis Discriminante Múltiple Lineal.

	Fracaso Empresarial = 1		
	t-1	t-2	t-3
Prior Probabilities			
Sana	0.525	0.530	0.611
Quebrada	0.475	0.381	0.389
Variables			
SL1	-1.582	-1.940	-1.783
SL3	-1.520	-1.406	-2.080
SL4	-0.001	-0.288	0.346
SL5	-0.001	-0.002	0.001
SL6	-0.334	0.328	-0.556
SL7		0.010	0.011
SL8	0.031	0.057	0.021
CA3	0.755		0.012
CA5	-0.065	-0.272	-0.762
CA6	-0.006	-0.016	-0.001
E2	1.064	1.558	1.612
RE2	-0.209	0.052	-0.316
RE5	-0.027	0.046	-0.011
RE6	-0.027	-0.011	-0.093
ACTIVO ^a	0.000	0.000	0.000
R1	0.000	0.002	-0.006
R2	0.028	-0.037	0.015
APORTE Y CRÉDITO	-0.154	-0.142	-0.093
SIZE: Mediana	0.185	-0.059	-0.511
SIZE: Pequeña	1.052	0.863	0.400
SIZE: Micro	2.500	2.272	1.867

^a Activo deflactado con IPC base 2008.

Fuente: Elaboración propia.

Tabla 6: Comparación del grado de error entre modelos. MDA.

Modelo	Muestra	Muestra de Estimación				Muestra de Validación			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	2089	73.48 %	46.90 %	26.38 %	26.64 %	75.48 %	50.77 %	27.02 %	22.26 %
t-2	1850	71.46 %	41.84 %	31.31 %	26.54 %	75.97 %	50.65 %	28.87 %	20.52 %
t-3	1715	72.36 %	41.34 %	37.93 %	21.09 %	70.73 %	39.00 %	34.94 %	25.67 %

El coeficiente Kappa se calcula como la probabilidad de éxito de cada escenario, ponderando por las probabilidades de ocurrencia.

Fuente: Elaboración propia.

muy grandes. Los coeficientes están inflados debido a un error de la muestra en la diferenciación de los grupos⁶, para lo cual se propuso la estimación de un logit bayesiano para corregir el sesgo. Se mantuvo la estimación logit en los casos donde el grado de acierto fuera mayor que por medio del logit bayesiano. Sin embargo, en ningún caso se consiguieron mejores resultados utilizando la estimación del logit bayesiano.

La Tabla 7 muestra el mejor modelo estimado, donde se encuentran variables significativas en Solvencia y Liquidez, lo que indica la importancia de poder responder a las obligaciones de corto plazo, así como el signo negativo y significativo de las variables de cobertura y apalancamiento que señalan la importancia de tener un nivel de patrimonio que responda al nivel de activos y evitar el sobre-endeudamiento. Las variables de rentabilidad con signo negativo muestran la importancia de que la firma se piense rentable para la reinversión y las variables de control muestran que las firmas con nivel de activos más grande tienen una menor probabilidad de caer en insolvencia -que se observa a través de las variables de tamaño y el coeficiente más grande entre menor sea el tamaño.

Tabla 7: Regresión Logística. Mejor estimación.

Variable	Fracaso Empresarial = 1		
	t-1	t-2	t-3
SL1	-2.552*** (-0.484)	-2.138*** (0.440)	-2.244*** (0.384)
SL3		-2.210*** (0.582)	-2.862*** (0.769)
SL4		-0.416 (0.857)	
SL5	-0.001 (0.001)	-0.001 (0.001)	
SL6	-0.832* (0.488)	0.687 (0.971)	-0.022 (0.372)
SL8	0.089*** (0.031)	0.108*** (0.032)	0.057* (0.034)
CA3	0.5195* (0.298)		
CA5	-1.106** (0.451)	-0.459** (0.227)	-1.011*** (0.391)

⁶Un error de diferenciación perfecta que no permite evaluar con eficacia, en algunos casos, la diferencia entre una entidad quebrada y una sana respecto a una variable particular. Sin embargo, esto no detiene la estimación.

Variable	t-1	t-2	t-3
CA6	-0.063*** (0.024)	-0.063*** (0.021)	-0.011 (0.009)
E2	3.559*** (0.525)	2.100*** (0.338)	1.723*** (0.330)
RE2	-0.853*** (0.269)		-0.230 (0.205)
RE5	-0.783*** (0.290)	0.010 (0.010)	-0.009 (0.012)
RE6	-0.341** (0.134)	-0.016 (0.017)	-0.065 (0.093)
R1	0.001 (0.001)		
R2	0.818*** (0.290)		
APORTE Y CRÉDITO	-0.083 (0.188)	-0.079 (0.131)	-0.072 (0.136)
SUPERV: Nivel 2	-0.694 (0.474)	-0.962** (0.416)	-1.094** (0.470)
SUPERV: Nivel 3	0.223 (0.551)	0.311 (0.423)	0.605 (0.504)
ASOCIADOS	0.000 (0.000)		
SIZE: Mediana	1.234* (0.713)	0.562 (0.512)	0.279 (0.562)
SIZE: Pequeña	2.138*** (0.808)	1.135** (0.557)	0.601 (0.643)
SIZE: Micro	4.291*** (0.819)	2.507*** (0.564)	1.983*** (0.653)
Constante	-0.401 (0.766)	1.748** (0.758)	2.986*** (0.870)
Observaciones	2089	1850	1715
Pseudo R2	0.54	0.22	0.21
Log likelihood	-667.081	-982.799	-907.859
AIC	1376.163	2001.6	1849.717
Desviación Residual	1334	1965	1816
Desviación NULO	2148	2493	2288
Método	Logit	Logit Bayesiano	Logit

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$. Errores estándar en paréntesis.

Fuente: Elaboración propia.

En el modelo predictivo se tiene un número mayor de variables explicativas, debido a que se han eliminado menos variables tomando en cuenta que, a pesar de la correlación, su presencia aumenta la capacidad predictiva del modelo. Es de recalcar que la mejora en la predicción -fin último de este documento- al utilizar este modelo, frente al más parsimonioso, es destacable y es la razón por la cual se sigue con este modelo en el resto de la estimación.

La Tabla 8 presenta el análisis de regresión, donde destacan las variables (Activo Corriente – Inventario)/Activo con coeficiente negativo y Cuentas por Cobrar/(Activo – Pasivo) con coeficiente positivo, que demuestran la importancia de la liquidez, también (Obligaciones Financieras + Cuentas por Pagar)/Patrimonio que corresponde a las obligaciones respecto al patrimonio y el signo negativo indica la importancia de un suficiente proceso de endeudamiento, con el fin de que la compañía no dependa sólo del aporte de sus asociados. Sorprende en la estimación en t-3, la significancia obtenida en las variables Pasivo/Activo y Patrimonio/Activo con signo negativo, contrario a lo esperado, dado que indicarían que mayores niveles de endeudamiento disminuyen la probabilidad de quiebra. En términos de rentabilidad, la variable Utilidad Operacional/Activo con signo negativo señala que, al igual que en entidades con ánimo de lucro, los ingresos son relevantes para mantener en funcionamiento; similar razonamiento encontrado con la variable EBIT/Activo en t-1 con signo negativo que señala que las entidades con mayores niveles de beneficios antes de intereses e impuestos respecto al activo, se encuentran más alejadas del fracaso empresarial. En contraposición, la variable Utilidad Neta/Activo con signo positivo en t-2 y t-3 resulta contradictoria a lo esperado.

Finalmente, las variables de control mejoran la estimación, dada su significancia. Los activos, aunque con valores pequeños de los coeficientes, resultan ser importantes para la clasificación de insolvencia; el nivel de supervisión 2 significativo y negativo indica que reduce la probabilidad de fracaso tener supervisión nivel medio a tener un alto nivel de supervisión, pero se debe recordar que las entidades con nivel 1 son entidades con actividades financieras, que se encuentran con procesos que los pueden llevar más rápido al fracaso. La cantidad de asociados se estiman en el modelo en t-1⁷ y resulta significativo con signo positivo y coeficiente pequeño, lo que indica que un mayor número de asociados aumenta la probabilidad de insolvencia. El tamaño de las firmas también es significativo, pero sus signos no son los esperados, ya que

⁷No fue posible utilizar la variable ASOCIADOS en t-2 y t-3 debido a que presentan una cantidad alta de valores nulos y la estimación se detiene, eliminar esos valores nulos implica reducir considerablemente la muestra de estimación y minar la capacidad predictiva del modelo, por lo que se opta por eliminar la variable.

indicarían que firmas más pequeñas tienen menos probabilidad de ser insolventes. El valor de la constante es grande debido a la presencia de los activos en niveles y no en desviaciones, el R^2 de McFadden indica que el modelo en t-1 explica el 59 % de la muestra frente al 24 % de los modelos en t-2 y t-3.

Tabla 8: Regresión Logística. Mejor predicción.

Variable	Fracaso Empresarial = 1		
	t-1	t-2	t-3
SL1	5.605 (3.60)	2.683 (2.249)	2.739 (2.569)
SL2	-9.696*** (3.744)	-5.152** (2.345)	-5.067* (2.706)
SL3	6.128* (3.132)	1.384 (1.909)	0.818 (2.246)
SL4	-0.006 (0.054)	-0.561 (0.944)	0.889 (1.356)
SL5	711100 (460000)	-438100 (441200)	722800 (541800)
SL6	-0.345 (0.545)	0.812 (1.062)	-1.263 (1.483)
SL7		0.009 (0.013)	0.031 (0.026)
SL8	0.172*** (0.048)	0.152*** (0.039)	0.085** (0.038)
CA1	-1246000 (3409000)		-6546000** (2764000)
CA2	0.049** (0.025)	-0.024* (0.013)	-0.003 (0.011)
CA3	-2096000 (3968000)		-6923000** (3166000)
CA5	-0.880 (0.536)	-0.054 (0.379)	0.093 (0.53)
CA6	-0.149*** (0.040)	-0.061** (0.025)	-0.018 (0.016)
E1	-711100 (460000)	438100 (441200)	-722800 (541800)
E2	4.514*** (0.626)	2.079*** (0.379)	1.848*** (0.418)
E3	-0.007 (0.015)	-0.012 (0.019)	-0.025 (0.021)
E4	-18.01*** (2.584)	1.279 (1.26)	1.924 (1.949)
E5	-849200 (4175000)		-377300 (3333000)

Variable	t-1	t-2	t-3
RE1	-4.32*** (1.113)	-3.01*** (0.999)	-2.949** (1.378)
RE2	-0.556 (0.347)	0.21 (0.144)	-0.18 (0.218)
RE3		1.63** (0.808)	2.761*** (1.069)
RE4	-0.079** (0.037)	-0.01 (0.022)	-0.032 (0.032)
RE5	-23.32*** (2.726)	-1.712 (1.437)	-1.324 (2.104)
RE6	-0.073 (0.163)	0.011 (0.03)	0.011 (0.093)
ACTIVO ^a	-0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
R1	0.003 (0.011)	0.014 (0.016)	-0.002 (0.013)
R2	23.35*** (2.726)	1.712 (1.437)	1.316 (2.104)
APORTE Y CRÉDITO	0.223 (0.208)	-0.069 (0.139)	0.024 (0.166)
SUPERV: Nivel 2	-1.102** (0.520)	-1.673*** (0.463)	-1.998*** (0.506)
SUPERV: Nivel 3	0.368 (0.612)	-0.581 (0.481)	-0.464 (0.533)
ASOCIADOS	0.000** (0.000)		
SIZE: Mediana	-2.495** (1.075)	-3.135*** (1.032)	-4.29*** (1.127)
SIZE: Pequeña	-2.803** (1.353)	-3.461*** (1.25)	-5.057*** (1.375)
SIZE: Micro	-0.815 (1.390)	-2.276* (1.292)	-3.862*** (1.421)
Constante	1384000 (4041000)	438100 (441200)	6200000 (3268000)
Observaciones	1559	1832	1710
Pseudo R^2	0.594	0.24	0.235
Log likelihood	-587.288	-956.343	-876.996
AIC	1240.576	1974.686	1821.992
Desviación Residual	1175	1923	1754
Desviación NULO	2148	2493	2288
Método	Logit	Logit	Logit

^a Activo deflactado con IPC base 2008.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$. Errores estándar en paréntesis.

Fuente: Elaboración propia.

Para evaluar la capacidad predictiva de los modelos Logit se presenta la Tabla 9, en esta se consignan los resultados de la predicción del modelo; es importante recalcar que, debido al alto número de observaciones, la demarcación del valor límite entre firmas quebradas y sanas se vuelve difuso, lo que, aunado con la representación mayoritaria de entidades sanas, hace más difícil el control sobre la probabilidad de fracaso; por consiguiente, se opta por reducir el punto de inflexión en la probabilidad a 0.45, es decir que, aquellas entidades cuya probabilidad de caer en insolvencia sea mayor al 45 %, se reportan como quebradas. Lo anterior permitirá controlar de mejor manera la sobrerrepresentación y, dado que se han comparado distintos puntos de quiebre, se tiene certeza de que provee mejores predicciones y de forma más equilibradas, ya que moverse más hacia el cero aumenta el error tipo II mientras el error tipo I disminuye, y moverse hacia el uno aumenta el error tipo I mientras el error tipo II disminuye, ambas con tradeoffs no lineales.

Allí se observa que la efectividad es, en promedio, superior al 70 %, con un destacado 84.5 % en la muestra de estimación y 80 % en la muestra de validación. El coeficiente Kappa se encuentra entre un 44-69 % en la muestra de estimación y entre el 39-60 % en la muestra de validación. Los errores tipo I son del orden del 13-33 % en la muestra de estimación y del 17-37 % en la muestra de validación. Finalmente, los errores tipo II rondan entre 19-33 % en la muestra de estimación y del 23-24 % en la muestra de validación. Al igual que con el modelo MDA, los resultados son mejores entre más cerca esté la muestra del suceso, asimismo, los resultados son muy similares en ambos modelos.

Tabla 9: Comparación del grado de error entre modelos. LOGIT.

Modelo	Muestra	Éxito	Muestra de Estimación			Muestra de Validación			
			Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	1560	84.49 %	68.58 %	12.51 %	19.15 %	80.05 %	59.89 %	17.39 %	22.83 %
t-2	1833	72.94 %	45.19 %	27.31 %	26.88 %	75.93 %	51.16 %	24.48 %	23.77 %
t-3	1720	73.47 %	44.43 %	32.98 %	22.41 %	71.36 %	39.79 %	36.75 %	23.46 %

El coeficiente Kappa se calcula como la probabilidad de éxito de cada escenario, ponderando por las probabilidades de ocurrencia.

Fuente: Elaboración propia.

Para iniciar, el lector se habrá dado cuenta de que, aunque se trata de la misma muestra, el número de observaciones en el logit es menor en casi todos los modelos, esto se debe a que el modelo elimina las observaciones con valores perdidos (*missings*), sin embargo, la fuerza de las predicciones no se ve afectada. Al comparar los modelos MDA y Logit, se encuentra que la predicción es más eficiente en todos los periodos en el modelo Logit frente al MDA, excepto con la muestra de validación en t-2 donde sus porcentajes de acierto son prácticamente iguales. Los porcentajes de error en los modelos son similares, con cierta ventaja para el modelo Logit que presenta, en los tres periodos antes del fracaso, niveles más bajos de error, especialmente el error tipo I, el cual se busca minimizar.

Al optar por reducir el error tipo I se aumentan, de forma no lineal, los niveles de error tipo II –en un *tradeoff*– lo que advierten todos los modelos estimados, sin embargo, estos son similares al error tipo I, lo que supondría que el modelo se encuentra equilibrado. Estos niveles de error tipo II son cercanos entre los modelos tipo Logit y los MDA, sin embargo, son los primeros los que siguen teniendo los menores niveles.

En el Anexo B se presentan los resultados de la estimación de los mismos modelos MDA y Logit utilizando submuestras, con el fin de comprobar la robustez de las estimaciones presentadas anteriormente. La Tabla 23, Tabla 24, Tabla ?? y Tabla 27 presentan las estimaciones realizadas utilizando como Submuestra aquellas entidades que prestan servicios de Aporte y Crédito, donde se conjugan el hecho de ser entidades sin ánimo de lucro con el servicio de ahorro y crédito; por su parte, las Tablas Tabla 29, Tabla 30, Tabla ?? y Tabla 33 presentan las estimaciones realizadas utilizando como Submuestra aquellas entidades que están activas desde el año 2002, esto con el fin de mantener constante el escenario inicial sobre el cual las entidades decidieron entrar al mercado. En estas estimaciones, se puede constatar que las predicciones mejoran ligeramente frente a la muestra total, lo que va en concordancia con el hecho de que la muestra más grande genera mayor distorsión al modelo aplicado. En ese orden de ideas, las tasas de éxito se encuentran superiores al 74 % y las tasas de error tipo I se mantienen ligeramente más bajas o similares en la Submuestra 2002, mientras que, para la Submuestra de entidades con Aporte y Crédito son similares, pero con rápido

crecimiento a medida que se aleja del periodo del suceso de quiebra. Finalmente, se destaca el hecho de que los errores tipo II son menores en estas estimaciones frente a la muestra total, con tasas entre el 10 % y el 20 %.

Continuando con las pruebas de robustez a los modelos estimados, se realiza el análisis de sensibilidad propuesto por Gómez *et al.* (2008) para comparar la efectividad de los modelos. La primera prueba consistente en la realización del test de diferencia de proporciones para comparar los niveles de los distintos tipos de error entre los modelos estimado. Dado que el artículo de Gómez *et al.* (2008) estima solo modelos logit en distintos periodos, en este documento se estimará dicho test, pero comparando, para el mismo periodo, entre modelos MDA y logit. La Tabla 10 presenta los resultados de dicho test realizado a un nivel de significancia del 95 % a doble cola. En el periodo anterior a la quiebra de las entidades, se observa que las proporciones son significativamente distintas entre el MDA y el logit en la muestra de estimación, pero sólo hay diferencias significativas en el error tipo I en la muestra de validación. En el caso de los periodos t-2 y t-3, no se encuentran diferencias significativas en los errores estimados entre los modelos MDA y Logit.

Tabla 10: Test de Diferencia de Proporciones entre Modelos. MDA vs Logit.

Modelo	Muestra de Estimación			Muestra de Validación		
	Error Tipo I	Error Tipo II	Error Global	Error Tipo I	Error Tipo II	Error Global
t-1	0.0000	0.0003	0.0000	0.0198	0.9781	0.1198
t-2	0.0942	0.8961	0.3339	0.3902	0.4239	1.0000
t-3	0.0670	0.4952	0.4910	0.8190	0.6285	0.8973

Ho: Proporciones iguales, Ha: Proporciones distintas, $p \leq 0,05$.

Fuente: Elaboración propia.

Lo anterior permite dar más evidencia a lo que se ha venido conjeturando respecto a la similitud de las estimaciones entre las especificaciones Logit y MDA. A su vez, se permite considerar el hecho de que las estimaciones Logit son ligeramente superiores, esto debido a que, al menos en el periodo t-1 para la muestra de estimación y para el error tipo I en la muestra de validación, se obtienen diferencias significativas que dan ventaja a la especificación Logit.

Al igual que con las estimaciones y la comparación entre modelos, en el Anexo B se presentan los resultados del test de diferencia de proporciones, aplicados a los modelos usando las submuestras. La Tabla 28 muestra los resultados para la submuestra de Aporte y Crédito, mientras la Tabla 34 hace lo correspondiente con la submuestra con las entidades que inician labores en 2002. Los resultados muestran que las diferencias se mantienen, exceptuando el error tipo I en el periodo t-2 de ambas submuestras y en el error tipo II en el periodo t-3 de la submuestra de Aporte y Crédito en las cuales el test rechaza la hipótesis nula de proporciones iguales.

Por lo anterior, se puede considerar el hecho de que las estimaciones, aun considerando pruebas de robustez, mantienen el hecho de que las estimaciones son similares y le da una ligera ventaja al modelo Logit, dada la diferencia significativa en t-1 en la muestra de estimación.

La siguiente prueba de sensibilidad de los modelos, utilizada también en Gómez *et al.* (2008), considera el hecho de que, en el mundo real, no es posible saber en qué periodo antes de la quiebra se encuentra la entidad, por lo que utilizar la especificación del periodo t-1 cuando una firma está a dos o tres años de la quiebra, podría resultar en la alteración de los resultados. Por ende, los autores proponen utilizar los modelos en cada periodo para predecir las muestras de otros periodos, por ejemplo, utilizando la especificación MDA del periodo en t-1, para predecir la muestra de validación del periodo t-2 y t-3. Con esto, se espera encontrar el modelo que se adecúe de forma más generalizada a la realidad y tenga capacidad de adaptación.

Dado que las muestras entre periodos son iguales entre especificación, es decir, la muestra en t-2 es la misma para el modelo Logit como el MDA, se compararán entre periodos solamente, manteniendo la diferenciación Logit y MDA. Y, finalmente, se utilizará de nuevo el test de diferencia de proporciones que permita evaluar diferencias significativas entre los errores obtenidos entre el modelo original, es decir, utilizando la especificación inicial para el periodo, y el modelo en consideración. A manera de ejemplo, considere que la prueba está buscando diferencias significativas entre usar la especificación en t-1 para estimar la muestra de entidades que están a dos años de la quiebra, y utilizar la especificación en t-2 como originalmente se planteó.

Las Tablas 10 a 15 presentan los resultados de la estimación para los modelos MDA. Para el modelo en t-1 (Tabla 11), predecir con la muestra en t-2 aumenta ligeramente los niveles de error y disminuye la efectividad, especialmente el coeficiente Kappa que controla por la probabilidad de ocurrencia, sin embargo, el test de diferencia de proporciones señala que no hay diferencias significativas (Tabla 12), por lo cual, es plausible considerar que el modelo en t-1 podría ser útil para estimar en cualquier caso. En el caso del modelo en t-2 (Tabla 13), se observan sutiles mejoras en la predicción frente al modelo en t-1 excepto en el error tipo II, donde se presenta una diferencia favorable al modelo en t-1 de 0.73 % y en el caso del modelo en t-3 el error es el mismo, pero, al considerar el test (Tabla 14), no se encuentran diferencias significativas. Finalmente, el modelo en t-3 (Tabla 15) presenta leves mejoras en los niveles de éxito, que se ven reflejados en la disminución del error tipo II, pero se aumenta el error tipo I –el más grave–, el test de diferencia de proporciones aplicado (Tabla 16) no muestra diferencias significativas entre los modelos.

De lo anterior, se puede observar que los modelos tienen una capacidad relativamente igual de predicción, y que podrían ser intercambiables a la hora de predecir. Sin embargo, en algunos casos esto tiene unos costos en el *tradeoff* entre el error tipo I y tipo II.

Tabla 11: Comparación del acierto y error entre el modelo original para cada año y el modelo en t-1 usando la muestra de validación correspondiente a cada año. MDA.

Modelo- Muestra	Obs.	Muestra de Original				Muestra de Validación (Modelo t-1)			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-2	462	75.97 %	50.65 %	28.87 %	20.52 %	73.59 %	45.64 %	32.47 %	22.02 %
t-3	427	70.73 %	39.00 %	34.94 %	25.67 %	70.73 %	39.14 %	34.34 %	26.05 %

Las columnas 3-6 utilizan el modelo mencionado en la columna 1 sobre la muestra de validación mencionada en la misma columna 1 (es decir, es la réplica de la segunda parte de la tabla 6). Por su parte, las columnas 7-10 utilizan el modelo estimado en t-1 sobre la muestra en el año de la columna 1.

Fuente: Elaboración propia.

Consecutivamente, las Tablas 16 a 21 presentan los resultados para la estimación Logit. El modelo en t-1 utilizado para predecir las muestras en t-2 y t-3 (Tabla 17)

Tabla 12: Test de Diferencia de Proporciones. Modelo t-1. MDA.

Modelo	Modelo t-1		
	Error Tipo I	Error Tipo II	Error Global
t-2	0.5089	0.7515	0.4487
t-3	1.0000	1.0000	1.0000

Ho: Proporciones iguales, Ha: Proporciones distintas, $p \leq 0,05$.

Fuente: Elaboración propia.

Tabla 13: Comparación del acierto y error entre el modelo original para cada año y el modelo en t-2 usando la muestra de validación correspondiente a cada año. MDA.

Modelo-Muestra	Obs.	Muestra de Original				Muestra de Validación (Modelo t-2)			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	522	75.48 %	50.77 %	27.02 %	22.26 %	75.86 %	51.60 %	25.40 %	22.99 %
t-3	427	70.73 %	39.00 %	34.94 %	25.67 %	70.96 %	39.56 %	34.34 %	25.67 %

Las columnas 3-6 utilizan el modelo mencionado en la columna 1 sobre la muestra de validación mencionada en la misma columna 1 (es decir, es la réplica de la segunda parte de la tabla 6). Por su parte, las columnas 7-10 utilizan el modelo estimado en t-1 sobre la muestra en el año de la columna 1.

Fuente: Elaboración propia.

Tabla 14: Test de Diferencia de Proporciones. Modelo t-2. MDA.

Modelo	Modelo t-2		
	Error Tipo I	Error Tipo II	Error Global
t-1	0.7594	0.9187	0.9425
t-3	1.0000	1.0000	1.0000

Ho: Proporciones iguales, Ha: Proporciones distintas, $p \leq 0,05$.

Fuente: Elaboración propia.

es superior a los originales, al considerar las tasas de éxito y los niveles considerablemente más bajos de error tipo I, sin embargo, al controlar por las probabilidades de ocurrencia, se observa que el coeficiente Kappa es muy bajo y los niveles de error tipo II son elevados en comparación al modelo original; esto también lo demuestra el test de diferencia de proporciones (Tabla 18). Debido a lo anterior, este modelo sería considerado “alarmista” al utilizar otras muestras, puesto que reduce el error tipo I, aumentando el error tipo II. El modelo en t-2 presenta resultados similares a los

Tabla 15: Comparación del acierto y error entre el modelo original para cada año y el modelo en t-3 usando la muestra de validación correspondiente a cada año. MDA.

Modelo-Muestra	Obs.	Muestra de Original				Muestra de Validación (Modelo t-3)			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	522	75.48 %	50.77 %	27.02 %	22.26 %	75.86 %	51.47 %	28.23 %	20.44 %
t-2	462	75.97 %	50.65 %	28.87 %	20.52 %	76.19 %	50.64 %	31.96 %	17.91 %

Las columnas 3-6 utilizan el modelo mencionado en la columna 1 sobre la muestra de validación mencionada en la misma columna 1 (es decir, es la réplica de la segunda parte de la tabla 6). Por su parte, las columnas 7-10 utilizan el modelo estimado en t-1 sobre la muestra en el año de la columna 1.

Fuente: Elaboración propia.

Tabla 16: Test de Diferencia de Proporciones. Modelo t-3. MDA.

Modelo	Modelo t-3		
	Error Tipo I	Error Tipo II	Error Global
t-1	0.8408	0.6767	0.9425
t-3	0.5811	0.5107	1.0000

Ho: Proporciones iguales, Ha: Proporciones distintas, $p \leq 0,05$.

Fuente: Elaboración propia.

de los modelos originales (Tabla 19), se observa que el modelo en t-1 tiene mayores niveles de éxito y menos porcentaje de error de clasificación, mientras que el modelo en t-3 tiene un error tipo I más grande frente a utilizar la especificación en t-2 para la muestra de validación en t-3, lo que es compensado con un menor error tipo II; la prueba de diferencia de proporciones (Tabla 20) muestra que no hay diferencias significativas entre los modelos.

Para concluir esta prueba, se realiza la estimación considerando la especificación en t-3 (Tabla 21), donde se observa que la única ventaja que trae esta especificación sobre las demás muestras es la reducción del error tipo II, puesto que aumenta el error tipo I y reduce el nivel de acierto, lo que a su vez no es respaldado por el test de diferencia de proporciones (Tabla 22), donde sólo el error tipo I para la especificación en t-1 tiene distintas proporciones el considerar el modelo original en t-1 y el modelo en t-3.

Los resultados obtenidos en esta prueba señalan que no existirían diferencias marcadamente significativas entre aplicar el modelo en t-2 y en t-3 a las muestras en otros años, exceptuando el hecho de aumentar el error tipo I de forma significativa en t-3 respecto a la muestra en t-1. Por su parte, el modelo en t-1 presenta fallas considerables en los rangos de clasificación, por lo que no se recomendaría usar en el caso que se prefieran predicciones equilibradas, puesto que este modelo es “alarmista” y aumentaría los casos de entidades que estando sanas aparezcan como quebradas, lo que generaría una supervisión innecesaria y una toma de acciones de forma preventiva por parte de las entidades que se encuentran en buen estado.

Tabla 17: Comparación del acierto y error entre el modelo original para cada año y el modelo en t-1 usando la muestra de validación correspondiente a cada año. Logit.

Modelo-Muestra	Obs.	Muestra de Original				Muestra de Validación (Modelo t-1)			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-2	457	75.93 %	51.16 %	24.48 %	23.77 %	83.33 %	2.55 %	15.29 %	75.00 %
t-3	426	71.36 %	39.79 %	36.75 %	23.46 %	76.02 %	3.89 %	22.89 %	60.00 %

Las columnas 3-6 utilizan el modelo mencionado en la columna 1 sobre la muestra de validación mencionada en la misma columna 1 (es decir, es la réplica de la segunda parte de la tabla 9). Por su parte, las columnas 7-10 utilizan el modelo estimado en t-1 sobre la muestra en el año de la columna 1.

Fuente: Elaboración propia.

Tabla 18: Test de Diferencia de Proporciones. Modelo t-1. Logit.

Modelo	Modelo t-1		
	Error Tipo I	Error Tipo II	Error Global
t-2	0.0411	0.0754	0.0577
t-3	0.0083	0.1728	0.2918

Ho: Proporciones iguales, Ha: Proporciones distintas, $p \leq 0,05$.

Fuente: Elaboración propia.

Considerando toda la información obtenida en los resultados, que se podría resumir en el hecho de que se encuentran que las estimaciones son similares entre aplicar MDA o Logit. Asimismo, se encuentra que, exceptuando por la estimación del modelo logit en el periodo t-1 para pronosticar las muestras en los periodos t-2 y t-3, la aplicación

Tabla 19: Comparación del acierto y error entre el modelo original para cada año y el modelo en t-2 usando la muestra de validación correspondiente a cada año. Logit.

Modelo-Muestra	Obs.	Muestra de Original				Muestra de Validación (Modelo t-2)			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	391	80.05 %	59.89 %	17.39 %	22.83 %	76.27 %	52.51 %	22.92 %	24.44 %
t-3	426	71.36 %	39.79 %	36.75 %	23.46 %	71.13 %	40.15 %	33.13 %	26.15 %

Las columnas 3-6 utilizan el modelo mencionado en la columna 1 sobre la muestra de validación mencionada en la misma columna 1 (es decir, es la réplica de la segunda parte de la tabla 9). Por su parte, las columnas 7-10 utilizan el modelo estimado en t-1 sobre la muestra en el año de la columna 1.

Fuente: Elaboración propia.

Tabla 20: Test de Diferencia de Proporciones. Modelo t-2. Logit.

Modelo	Modelo t-2		
	Error Tipo I	Error Tipo II	Error Global
t-1	0.1839	0.7754	0.2029
t-3	0.5649	0.5424	1.0000

Ho: Proporciones iguales, Ha: Proporciones distintas, $p \leq 0,05$.

Fuente: Elaboración propia.

Tabla 21: Comparación del acierto y error entre el modelo original para cada año y el modelo en t-3 usando la muestra de validación correspondiente a cada año. Logit.

Modelo-Muestra	Obs.	Muestra de Original				Muestra de Validación (Modelo t-3)			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	391	80.05 %	59.89 %	17.39 %	22.83 %	75.88 %	51.52 %	27.08 %	21.48 %
t-2	457	75.93 %	51.16 %	24.48 %	23.77 %	74.62 %	47.60 %	32.29 %	20.38 %

Las columnas 3-6 utilizan el modelo mencionado en la columna 1 sobre la muestra de validación mencionada en la misma columna 1 (es decir, es la réplica de la segunda parte de la tabla 9). Por su parte, las columnas 7-10 utilizan el modelo estimado en t-1 sobre la muestra en el año de la columna 1.

Fuente: Elaboración propia.

de los modelos en distintos periodos no tiene diferencias significativas, por lo que todos los modelos son válidos para ser aplicados en los casos donde no se conozca el momento en que una firma caerá en insolvencia.

A pesar de que el modelo Logit en t-1 presenta predicciones ligeramente supe-

Tabla 22: Test de Diferencia de Proporciones. Modelo t-3. Logit.

Modelo	Modelo t-3		
	Error Tipo I	Error Tipo II	Error Global
t-1	0.0198	0.8227	0.1588
t-3	0.1131	0.4021	0.7015

Ho: Proporciones iguales, Ha: Proporciones distintas, $p \leq 0,05$.

Fuente: Elaboración propia.

riores a los demás modelos, el hecho de que falle notablemente en la predicción de muestras en otros periodos lo hace difícilmente aplicable a otras muestras cuando se requieran modelos equilibrados –aunque resulta conveniente en casos donde un modelo “alarmista” sea considerado una mejor opción–. Por ende, los modelos MDA, a pesar de su relativa simplicidad, presentan niveles de pronóstico satisfactorios frente a los modelos Logit. Mientras que los modelos Logit, presentan niveles de error más bajos que los modelos MDA, pero son más sensibles a las variables utilizadas y al entorno económico implícito.

6. Conclusiones

El fracaso empresarial ha sido ampliamente estudiado a lo largo de la literatura, donde se han aplicado decenas de modelos distintos que van desde estadísticos sencillos, hasta modelos no paramétricos de redes neuronales. La dificultad para generalizar los hallazgos y proferir un modelo insignia, así como la imposibilidad de hallar un grupo único de variables que sean relevantes, son los hechos que han permitido su amplio desarrollo, utilizando técnicas de frontera para enfrentar el problema. Sin embargo, existen modelos que han sido utilizados repetidamente gracias a su capacidad para proferir predicciones ajustadas al contexto, entre los que se destacan el modelo MDA y el modelo Logit. En este trabajo se han utilizado estas técnicas, teniendo en cuenta las limitaciones que pueden tener las estimaciones y el limitado campo de acción de los resultados, para encontrar una especificación que sea útil para el sector solidario en Colombia.

El conocimiento de la futura insolvencia de entidades del sector solidario, les permitiría a los directivos y al ente regulador tomar medidas preventivas que eviten el fallo de mercado, el cual afecta principalmente a los asociados y a los consumidores de los servicios que prestan dichas entidades. El ejemplo particular que motiva este documento es la cooperativa de ahorro y crédito rural, que provee servicios en zonas donde la banca comercial no llega, la quiebra de estas entidades condena a la marginalización del mercado financiero a los campesinos que utilizan estos servicios para sobrellevar la falta de acceso a créditos formales de la banca comercial ubicada, generalmente, en las zonas más urbanizadas.

La estimación del modelo MDA y el modelo Logit en estas entidades, ha permitido encontrar la relevancia que tiene la Liquidez dentro de la firma para responder a las obligaciones de corto plazo, así como tener un margen suficiente de Cobertura, pero también de Apalancamiento, no depender exclusivamente de los aportes de los socios, pero tampoco excederse con cupos de endeudamiento. Finalmente, se ha destacado la importancia de la Rentabilidad en estas entidades, más allá de no tengan ánimo de lucro, puesto que esta rentabilidad se convierte en reinversiones que les permiten

crecer. Los controles señalan que las firmas pequeñas tienen una mayor probabilidad de quiebra, así como aquellas que tienen niveles de supervisión bajos.

Respecto a la capacidad predictiva, es importante recalcar en que la estimación se encuentra limitada debido a la dispersión de los datos por el aumento de la muestra –en pro de la mayor exactitud del entorno colombiano– y debido a los problemas muestrales comunes como la presentación incorrecta de los estados financieros. Sin embargo, los resultados señalan que los modelos son efectivos en predecir, incluso en el caso donde no se tenga conocimiento del tiempo en que se encuentra una entidad de la quiebra, exceptuando el caso del modelo Logit en el periodo $t-1$ que es un modelo “alarmista” y sólo debería ser utilizado en los casos donde sea conveniente reducir al máximo las posibles quiebras. Asimismo, se encuentra que los modelos MDA presentan resultados satisfactorios, similares al modelo Logit y no son tan sensibles a la especificación.

Finalmente, los resultados muestran que los modelos Logit y MDA son efectivos para predecir el fracaso empresarial y se debe propender por utilizar toda la información posible para los análisis de insolvencia, esto debido a que el entorno económico es dinámico. Es imperante aplicar con rigurosidad los análisis de insolvencia con los modelos disponibles, donde Logit y MDA presentan resultados satisfactorios con bajo costo de estimación, para evitar los fallos de mercado sobre entidades tan relevantes como las sin ánimo de lucro.

Así pues, este documento presentó la primera estimación de modelos de fracaso empresarial en entidades de economía solidaria en Colombia, siguiendo recomendaciones de artículos estimados para entidades con ánimo de lucro que ya han sido estimados en este país. Este acercamiento permitió encontrar algunas variables relevantes en el análisis y que, a diferencia de lo que se podría considerar a priori, son similares a los comportamientos observados en las firmas con ánimo de lucro, como la importancia de la liquidez, el endeudamiento positivo pero controlado y la rentabilidad. También, reconoció que los modelos aplicados en la literatura, tienen validez en Colombia y que, a pesar de la dispersión provocada por el aumento de la muestra, obtienen resultados consistentes con la literatura.

Referencias

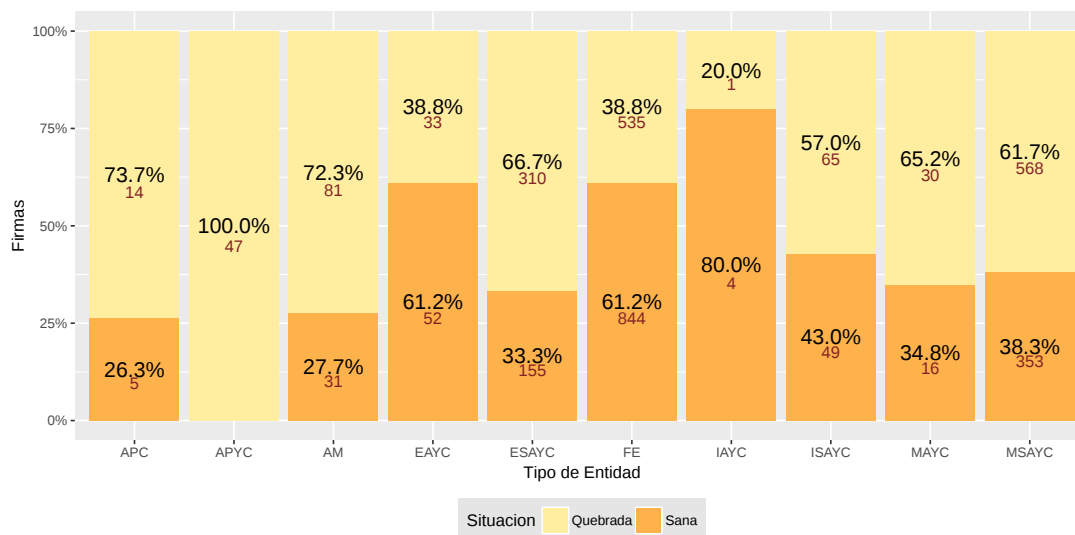
- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4), 589. doi:10.2307/2978933
- Altman, E. I. (1978). Examining Moyer's Re-examination of Forecasting Financial Failure. *Financial Management*, 7(4), 76-79. doi:10.2307/3665090
- Altman, E. I., Marco, G. & Varetto, F. (1994). Corporate distress diagnosis : Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance*, 18, 505-529.
- Balcaen, S. & Ooghe, H. (2006). 35 years of studies on business failure: An overview of the classic statistical methodologies and their related problems. *British Accounting Review*, 38(1), 63-93. doi:10.1016/j.bar.2005.09.001
- Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 4(1966), 71. doi:10.2307/2490171
- Bellovary, J. L., Giacchino, D. E. & Akers, M. D. (2007). A Review of Bankruptcy Prediction Studies: 1930 to Present. *Journal of Financial Education*, 33(Winter 2007), 1-42. Recuperado desde <http://www.jstor.org/stable/41948574%7B%5C%%7D5Cn>
- Braga, M. J., Fully Bressan, V. G., Colosimo, E. A. & Bressan, A. A. (2006). Investigating the solvency of brazilian credit unions using a proportional hazard model. *Annals of Public and Cooperative Economics*, 77(1), 83-106. doi:10.1111/j.1370-4788.2006.00298.x
- Congreso de Colombia. (1998). Ley 454 de 1998. Recuperado desde <http://www.orgsolidarias.gov.co/sites/default/files/pagina-basica/pdf/LEY%20454%20DE%201998.pdf>
- de Llano, P., Piñeiro, C. & Rodríguez, M. (2016). Predicción del fracaso empresarial: Una contribución a la síntesis de una teoría mediante el análisis comparativo de distintas técnicas de predicción. *Estudios de economía*, 43(2), 163-198. doi:10.4067/S0718-52862016000200001
- Deakin, E. B. (1976). Distributions of Financial Accounting Ratios : Some Empirical Evidence. *American Accounting Association*, 51(1), 90-96.
- Dietrich, J., Arcelus, F. J. & Srinivasan, G. (2005). Predicting financial failure: some evidence from new brunswick agricultural co-ops. *Annals of Public and Cooperative Economics*, 76(2), 179-194. doi:10.1111/j.1370-4788.2005.00275.x
- Edmister, R. O. (1972). An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. *The Journal of Financial and Quantitative Analysis*, 7(2), 1477-1493. doi:10.2307/2329929
- Enguános, A. (1994). Limitaciones Metodológicas de los trabajos empíricos sobre la predicción del Fracaso Empresarial. *Revista Española De Financiación Y Contabilidad*, 23(80), 709-732. Recuperado desde <http://www.jstor.org/stable/42782129>
- Gómez, M. E., de la Torre, J. M. & Román, I. (2008). Análisis de sensibilidad temporal en los modelos de predicción de insolvencia: una aplicación a las PYMES

- industriales. *Revista Española de Financiación y Contabilidad*, 37(137), 85-111. doi:10.1080/02102412.2008.10779640
- Hair, J. F., Black, W. C., Babin, B. J. & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th). arXiv: arXiv:1011.1669v3
- Jiménez, S. M. (1996). Una evaluación crítica de la investigación empírica desarrollada en torno a la solvencia empresarial. *Revista Española De Financiación Y Contabilidad*, 25(87), 459-479.
- Landis, J. R. & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174. doi:10.2307/2529310
- Lepetit, L. & Strobel, F. (2015). Bank insolvency risk and Z-score measures: A refinement. *Finance Research Letters*, 13, 214-224. doi:10.1016/j.frl.2015.01.001
- Lev, B. (1978). *Análisis de estados financieros: Un nuevo enfoque* (Ediciones). Madrid.
- Mare, D. S., Moreira, F. & Rossi, R. (2017). Nonstationary Z-Score measures. *European Journal of Operational Research*, 260(1), 348-358. doi:10.1016/j.ejor.2016.12.001
- Mateos, A., Marín, M. d. M., Marí, S. & Seguí, E. (2011). Los modelos de predicción del fracaso empresarial y su aplicabilidad en cooperativas agrarias Los modelos de predicción del fracaso empresarial y su aplicabilidad en cooperativas agrarias. *ciriec, Revista de Economía pública, social y cooperativa*, (70), 179-208.
- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109-131. doi:10.2307/2490395
- Pacey, J. W. & Pham, T. M. (1990). The predictive of Bankruptcy models: methodological problems and evidence. *Australian Journal of Management*, 15(2), 315-337.
- Pille, P. & Paradi, J. C. (2002). Financial performance analysis of Ontario (Canada) Credit Unions: An application of DEA in the regulatory environment. *European Journal of Operational Research*, 139(2), 339-350. doi:10.1016/S0377-2217(01)00359-9
- Premachandra, I. M., Bhabra, G. S. & Sueyoshi, T. (2009). DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique. *European Journal of Operational Research*, 193(2), 412-424. doi:10.1016/j.ejor.2007.11.036
- Romero, F. (2013a). Alcances y limitaciones de los modelos de capacidad predictiva en el análisis del fracaso empresarial. *Ad-minister*, (23), 45-70. Recuperado desde <http://search.ebscohost.com/login.aspx?direct=true%7B%5C%7Ddb=bth%7B%5C%7DAN=92743171%7B%5C%7Dsite=bsi-live>
- Romero, F. (2013b). Variables financieras determinantes del fracaso empresarial para la pequeña y mediana empresa en Colombia: análisis bajo modelo Logit. *Pensamiento & Gestión*, (34), 235-277. Recuperado desde <http://search.ebscohost.com/login.aspx?direct=true%7B%5C%7Ddb=bth%7B%5C%7DAN=91100297%7B%5C%7Dlang=es%7B%5C%7Dsite=ehost-live>
- Sueyoshi, T. & Goto, M. (2009). Methodological comparison between DEA (data envelopment analysis) and DEA-DA (discriminant analysis) from the perspective

- of bankruptcy assessment. *European Journal of Operational Research*, 199(2), 561-575. doi:10.1016/j.ejor.2008.11.030
- Taffler, R. . J. . (1982). Forecasting Company Failure in the UK Using Discriminant Analysis and Financial Ratio Data. *Journal of the Royal Statistical Society*, 145(3), 342-358. Recuperado desde <http://www.jstor.org/stable/2981867>
- Taffler, R. J. (1983). The Assessment of Company Solvency and Performance Using a Statistical Model. *Accounting and Business Research*, 13(52), 295-308. doi:10.1080/00014788.1983.9729767
- Tafur-Saiden, C. (2009). Bancarización: una aproximación al caso Colombiano a la luz de América Latina. *Estudios Gerenciales*, 25(110), 13-37. doi:10.1016/S0123-5923(09)70060-9
- Tascón, M. T. & Castaño, F. J. (2012). Variables y Modelos Para La Identificación y Predicción Del Fracaso Empresarial: Revisión de La Investigación Empírica Reciente. *Revista de Contabilidad*, 15(1), 7-58. doi:10.1016/S1138-4891(12)70037-7
- Watson, J. & Everett, J. (1999). Small Business Failure Rates: Choice of Definition and Industry Effects. *International Small Business Journal*, 17(2), 31-47. doi:10.1177/0266242699172002. arXiv: 0803973233
- Zavgren, C. V. (1985). Assessing the Vulnerability to Failure of American Industrial Firms: A logistics Analysis. *Journal of Business Finance and Accounting*, 12(1), 19-45. doi:10.1111/j.1468-5957.1985.tb00077.x

A. Estadísticas Descriptivas

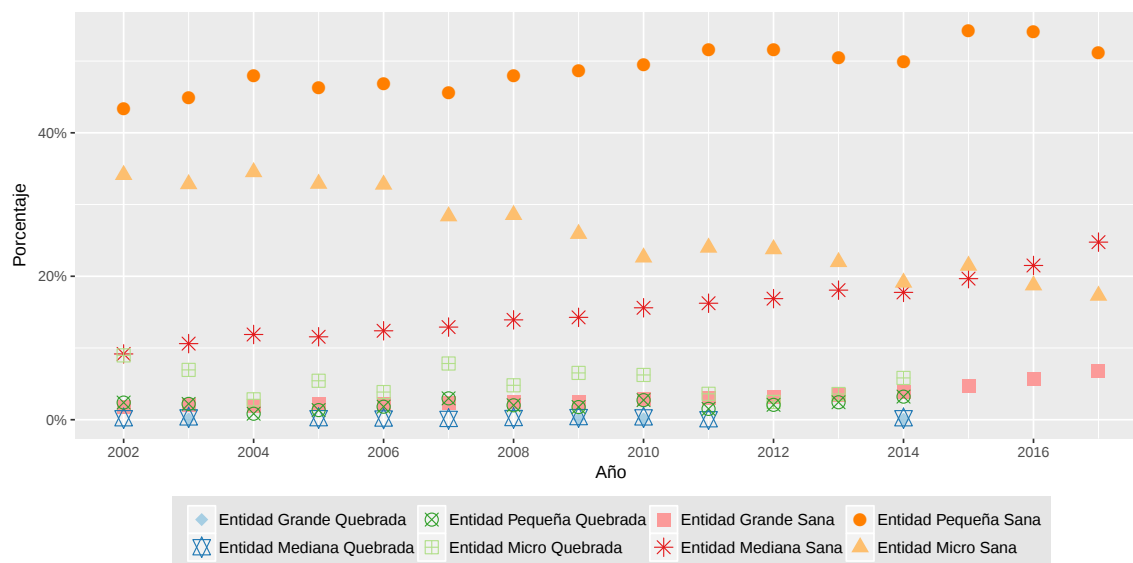
Figura 3: Entidades por Tipo.



Nota: APC: Administraciones Públicas Cooperativas, APYC: Cooperativa de Aporte y Crédito, AM: Asociación Mutua, EAYC: Cooperativa Especializada con Sección de Ahorro y Crédito, ESAYC: Cooperativa Especializada sin Sección de Ahorro y Crédito, FE: Fondos de Empleados, IAYC: Cooperativa Integral con Sección de Ahorro y Crédito, ISAYC: Cooperativa Integral sin Sección de Ahorro y Crédito, MAYC: Cooperativa Multiactiva con Sección de Ahorro y Crédito, MSAYC: Cooperativa Multiactiva sin Sección de Ahorro y Crédito.

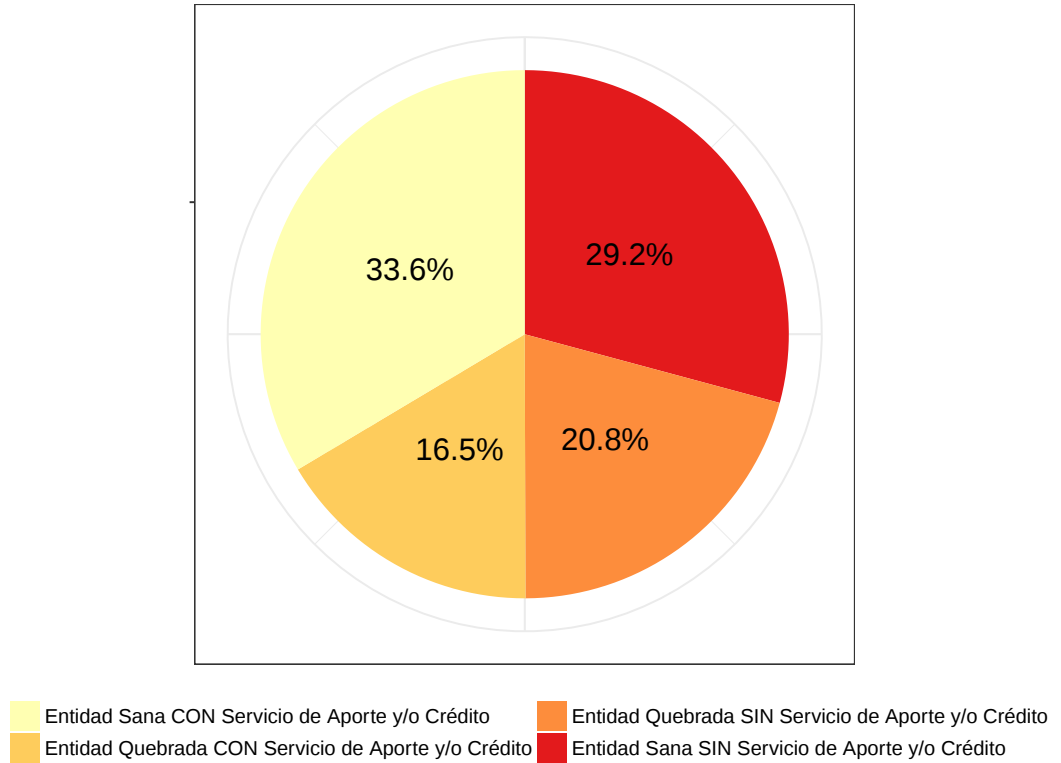
Fuente: Elaboración propia.

Figura 4: Entidades por Tamaño y Situación.



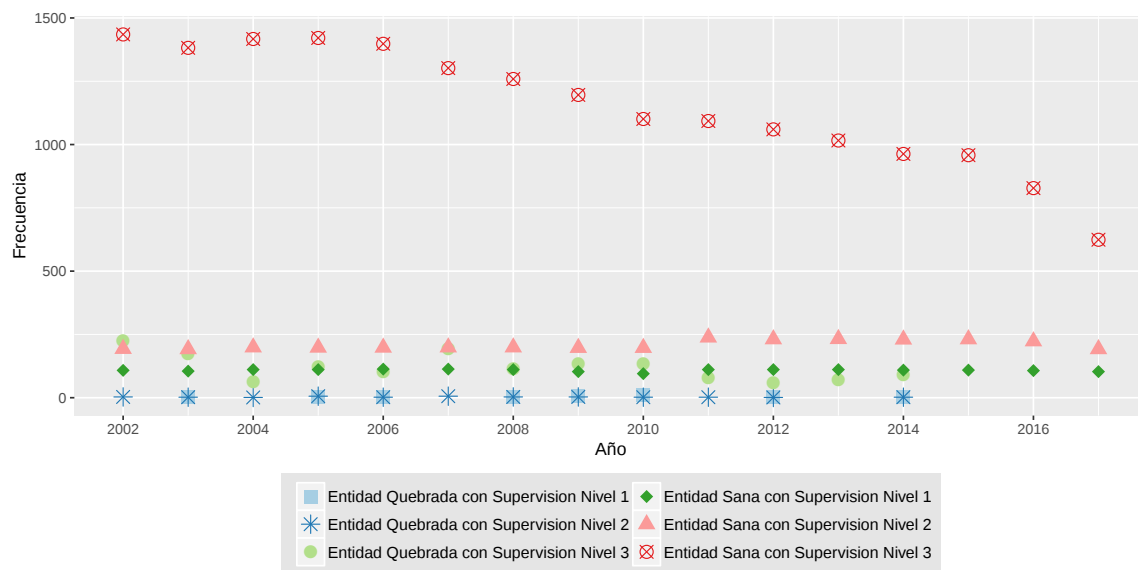
Fuente: Elaboración propia.

Figura 5: Entidades por Situación y Servicio de Ahorro y Crédito.



Fuente: Elaboración propia.

Figura 6: Entidades por Nivel de Supervisión y Situación.



Fuente: Elaboración propia.

B. Pruebas de Validación

Tabla 23: Análisis Discriminante Múltiple Lineal. Submuestra Aporte y Crédito.

Fracaso Empresarial = 1			
	t-1	t-2	t-3
Prior Probabilities			
Sana	0,641	0,680	0,705
Quebrada	0,359	0,320	0,295
Variables			
SL1	-2,163	-1,910	-4,715
SL3	-2,725	-3,078	-0,392
SL4	-0,029	0,328	-16,292
SL5	-0,001	0,008	0,003
SL6	-0,404	-0,636	16,401
SL7		-0,027	0,073
SL8	0,168	0,190	0,420
CA3	0,729		0,281
CA5	-0,832	-0,626	-0,233
CA6	0,107	-0,028	-0,008
E2	0,506	3,541	3,045
RE2	-0,231	0,149	-0,516
RE5	-3,039	-3,245	-5,272
RE6	-0,215	0,084	-0,077
ACTIVO ^b	0,000	0,000	0,000
R1	0,056	0,043	0,009
R2	3,093	3,489	5,360
SIZE: Mediana	0,145	0,173	-0,228
SIZE: Pequeña	1,026	1,114	0,746
SIZE: Micro	2,421	2,540	2,182

^a Se incluyen aquellas firmas que prestan servicios de Ahorro y/o Crédito.

^b Activo deflactado con IPC base 2008.

Fuente: Elaboración propia.

Tabla 24: Comparación del grado de error entre modelos. MDA. Submuestra Aporte y Crédito.

Modelo	Muestra	Muestra de Estimación				Muestra de Validación			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	1117	74.49 %	42.97 %	41.90 %	16.34 %	80.29 %	57.04 %	28.00 %	15.08 %
t-2	1036	75.77 %	41.73 %	46.39 %	13.78 %	74.42 %	35.68 %	56.63 %	10.86 %
t-3	982	75.66 %	36.08 %	55.86 %	11.13 %	73.36 %	33.00 %	52.78 %	15.70 %

El coeficiente Kappa se calcula como la probabilidad de éxito de cada escenario, ponderando por las probabilidades de ocurrencia.

Fuente: Elaboración propia.

Tabla 25: Regresión Logística con Submuestra Aporte y Crédito. Mejor estimación.

Variable	Fracaso Empresarial = 1		
	t-1	t-2	t-3
SL1	-5.079*** (1.183)	-3.217*** (1.147)	
SL3	-4.797* (2.527)		
SL6	-0.451 (0.544)	-0.518 (0.604)	-0.996 (0.695)
SL7			0.187** (0.095)
SL8	0.297** (0.125)	0.160* (0.091)	0.373** (0.161)
CA3	0.036 (0.094)		
CA5	-3.151*** (0.683)	-1.191* (0.666)	-0.748 (0.770)
CA6	0.485*** (0.168)	-0.129 (0.103)	-0.197 (0.138)
E2	0.148 (0.193)	5.906*** (1.079)	5.682*** (1.247)
RE2	0.228 (0.349)		-0.389 (0.333)
RE3	-3.630*** (1.124)	-0.548 (1.175)	
RE5			-0.186 (1.528)
RE6	-0.641 (0.446)	-0.134 (0.270)	-0.707 (0.519)
SUPERV: Nivel 2	-1.246* (0.672)	-0.997 (0.765)	-2.031** (0.882)
SUPERV: Nivel 3	0.852 (0.738)	1.261* (0.745)	0.672 (0.683)
SIZE: Mediana	1.456* (0.821)	1.051 (0.934)	0.487 (0.906)
SIZE: Pequeña	1.850** (0.934)	1.624 (0.988)	1.159 (0.996)
SIZE: Micro	3.394*** (0.944)	3.187*** (0.996)	2.574** (1.010)
Constante	5.915 (2.559)	-1.056 (1.385)	-2.988*** (0.781)
Observaciones	1117	1036	982
Pseudo R2	0.26	0.24	0.2
Log likelihood	-538.223	-496.945	-478.164
AIC	1110.4	1021.9	986.33
Desviación Residual	1076.4	993.89	956.33
Desviación NULO	1450.2	1295.72	1191.14
Método	Logit Bayesiano	Logit Bayesiano	Logit

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$. Errores estándar en paréntesis.

Fuente: Elaboración propia.

Tabla 26: Regresión Logística con Submuestra Aporte y Crédito. Mejor predicción.

Variable	Fracaso Empresarial = 1		
	t-1	t-2	t-3
SL1	-5.000 (6.49)	-69.86 (80.69)	-8.995 (8.979)
SL2	-4.7 (6.494)	67.02 (80.85)	3.998 (9.014)
SL3	-18.26 (13.7)	-67.18 (75.85)	-6.944 (7.592)
SL4	-0.045 (0.632)	0.124 (4.373)	-3.656 (5.557)
SL5	0.034 (0.045)	-1730000* (1026000)	-0.014 (0.056)
SL6	0.238 (1.075)	-0.719 (4.466)	3.538 (5.607)
SL7		0.025 (0.069)	0.129 (0.13)
SL8	0.514*** (0.174)	0.25** (0.111)	0.489*** (0.189)
CA1	0.018 (0.084)		0.013 (3.498)
CA2	-0.178** (0.081)	0.018 (0.055)	-0.124 (0.079)
CA3	-0.018 (0.084)		-0.013 (3.498)
CA5	-1.383 (0.912)	-0.063 (0.885)	-0.586 (1.313)
CA6	0.887*** (0.286)	-0.26 (0.185)	-0.119 (0.191)
E1	0.034 (0.045)	1730000 (1026000)	-0.014 (0.056)
E2	0.025 (0.119)	7.402 (1.532)	4.246*** (1.372)
E3	-0.191 (0.208)***	0.091 (0.123)	0.054 (0.294)
E4	-44.41 (6.636)	5.996 (5.281)	-1.248 (4.295)
E5	0.018 (0.084)		0.013 (3.498)
RE1	-4.704 (3.554)	-1.926 (3.273)	2.509 (4.282)
RE2	-1.196 (0.796)	0.388* (0.218)	-0.841* (0.462)
RE3		3.119 (2.228)	1.083 (2.538)
RE4	0.567 (0.643)	-0.476 (0.433)	0.885 (0.64)
RE5	-57.39*** (6.974)	0.633 (4.855)	-2.243 (3.119)
RE6	-0.058 (0.593)	-0.128 (0.393)	-0.778 (0.639)

Variable	t-1	t-2	t-3
ACTIVO ^a	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
R1	0.171 (0.261)	0.185 (0.175)	0.275 (0.26)
R2	55.75*** (7.077)	-0.737 (4.949)	1.8 (3.212)
SUPERV: Nivel 2	-0.949 (0.769)	-2.405** (1.003)	-2.103** (0.874)
SUPERV: Nivel 3	0.888 (0.838)	-0.794 (0.837)	0.258 (0.619)
ASOCIADOS	0.000** (0.000)		
SIZE: Mediana	-0.419 (1.179)	-2.353 (567.3)	-3.013** (1.267)
SIZE: Pequeña	-1.677 (1.208)	-4.086 (567.3)	-3.225** (1.409)
SIZE: Micro	0.239 (1.238)	-4.785 (567.3)	-2.041 (1.447)
Constante	28.56** (13.73)	1730000* (1026000)	13.79 (8.476)
Observaciones	849	1030	980
Pseudo R2	0.639	0.303	0.253
Log likelihood	-262.99	-453.05	-445.064
AIC	589.9	966.096	956.1
Desviación Residual	526	906.1	890.1
Desviación NULO	1142	1296	1191
Método	Logit Bayesiano	Logit	Logit Bayesiano

^a Activo deflactado con IPC base 2008.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$. Errores estándar en paréntesis.

Fuente: Elaboración propia.

Tabla 27: Comparación del grado de error entre modelos. Logit. Submuestra Aporte y Crédito.

Modelo	Muestra	Muestra de Estimación				Muestra de Validación			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	850	87.76 %	74.69 %	12.72 %	11.91 %	84.91 %	68.81 %	17.44 %	13.49 %
t-2	1031	76.82 %	46.87 %	36.14 %	17.02 %	78.99 %	50.71 %	37.35 %	13.22 %
t-3	981	74.72 %	37.42 %	47.93 %	15.77 %	72.13 %	34.59 %	43.06 %	21.51 %

El coeficiente Kappa se calcula como la probabilidad de éxito de cada escenario, ponderando por las probabilidades de ocurrencia.

Fuente: Elaboración propia.

Tabla 28: Test de Diferencia de Proporciones entre Modelos. MDA vs Logit. Submuestra Aporte y Crédito.

Modelo	Muestra de Estimación			Muestra de Validación		
	Error Tipo I	Error Tipo II	Error Global	Error Tipo I	Error Tipo II	Error Global
t-1	0.0000	0.0366	0.0000	0.1264	0.8227	0.2269
t-2	0.0093	0.1073	0.6117	0.0197	0.6077	0.2617
t-3	0.0675	0.0141	0.6664	0.3169	0.2124	0.8389

Ho: Proporciones iguales, Ha: Proporciones distintas, $p \leq 0,05$.

Fuente: Elaboración propia.

Tabla 29: Análisis Discriminante Múltiple Lineal. Submuestra 2002.

	Fracaso Empresarial = 1		
	t-1	t-2	t-3
Prior Probabilities			
Sana	0,537	0,589	0,611
Quebrada	0,463	0,411	0,389
Variables			
SL1	-1,013	-1,119	-1,176
SL3	-0,632	-1,151	-1,470
SL4	0,000	-0,245	-0,014
SL5	0,000	-0,002	-0,001
SL6	-0,648	-0,671	-0,399
SL7		0,008	0,025
SL8	0,028	0,129	0,142
CA3	0,291		0,031
CA5	-0,056	-0,457	-0,742
CA6	-0,001	-0,023	-0,008
E2	2,283	2,043	2,328
RE2	-0,456	-0,306	-0,344
RE5	-0,047	-0,054	-0,195
RE6	-0,061	-0,191	-0,007
ACTIVO ^b	0,000	0,000	0,000
R1	-0,002	0,001	-0,007
R2	0,068	0,054	0,219
APORTE Y CRÉDITO	-0,022	-0,074	-0,051
SIZE: Mediana	0,248	-0,057	-0,380
SIZE: Pequeña	1,189	0,955	0,715
SIZE: Micro	2,807	2,612	2,179

^a La Submuestra 2002 son aquellas firmas cuyo año de entrada es 2002, es decir que, se encuentra a lo largo de todo el periodo de la muestra.

^b Activo deflactado con IPC base 2008.

Fuente: Elaboración propia.

Tabla 30: Comparación del grado de error entre modelos. MDA. Submuestra 2002.

Modelo	Muestra	Muestra de Estimación				Muestra de Validación			
		Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	1389	78.26 %	56.12 %	25.97 %	18.10 %	80.64 %	60.83 %	25.00 %	14.52 %
t-2	1266	77.25 %	52.45 %	31.54 %	16.62 %	79.05 %	55.51 %	34.11 %	11.83 %
t-3	1221	74.61 %	45.63 %	37.47 %	17.69 %	80.59 %	58.69 %	27.97 %	13.98 %

El coeficiente Kappa se calcula como la probabilidad de éxito de cada escenario, ponderando por las probabilidades de ocurrencia.

Fuente: Elaboración propia.

Tabla 31: Regresión Logística con Submuestra 2002. Mejor estimación.

Variable	Fracaso Empresarial = 1		
	t-1	t-2	t-3
SL1			-1.495*** (0.441)
SL3	-3.425*** (1.218)	-2.914*** (0.882)	-2.565*** (0.854)
SL4	0.009 (0.036)		-0.292 (0.489)
SL5	0.003 (0.003)		
SL6		-1.231** (0.535)	
SL7			0.028 (0.028)
SL8	0.104 (0.086)	0.246*** (0.091)	
CA3	0.524 (0.402)		-0.401 (0.344)
CA4	-0.202 (0.164)		
CA5	-2.138*** (0.678)	-0.914* (0.473)	-1.365*** (0.498)
CA6		-0.042** (0.020)	
E2	5.178*** (0.745)	4.129*** (0.536)	3.298*** (0.556)
RE2		-0.252 (0.291)	-0.359 (0.258)
RE5	-1.658*** (0.579)	-0.011 (0.038)	
RE6	-0.909** (0.428)	-0.312 (0.227)	-0.052 (0.114)
R1	-0.012 (0.008)		
R2	1.757***		

Variable	t-1	t-2	t-3
	(0.577)		
APORTE Y CRÉDITO	0.117 (0.231)	-0.156 (0.155)	-0.148 (0.193)
SUPERV: Nivel 2	-0.804 (0.588)	-1.170** (0.552)	-1.179** (0.499)
SUPERV: Nivel 3	0.975 (0.786)	0.189 (0.581)	0.446 (0.540)
ASOCIADOS	0.000 (0.000)		
SIZE: Mediana	1.135 (0.897)	0.535 (0.709)	0.165 (0.590)
SIZE: Pequeña	1.585 (1.103)	1.627** (0.797)	1.032 (0.678)
SIZE: Micro	3.905*** (1.109)	3.585*** (0.812)	2.585*** (0.697)
Constante	-0.501 (1.449)	0.301 (0.963)	2.197** (0.988)
Observaciones	1389	1266	1221
Pseudo R2	0.54	0.3	0.25
Log likelihood	-442.678	-603.974	-611.624
AIC	925.36	1239.9	1255.2
Desviación Residual	885.36	1207.9	1223.2
Desviación NULO	1558.75	1714.5	1632.0
Método	Logit	Logit	Logit

^a Activo deflactado con IPC base 2008.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$. Errores estándar en paréntesis.

Fuente: Elaboración propia.

Tabla 32: Regresión Logística con Submuestra 2002. Mejor predicción.

Variable	Fracaso Empresarial = 1		
	t-1	t-2	t-3
SL1	9.894** (4.743)	0.74 (3.491)	0.677 (2.461)
SL2	-13.64*** (4.929)	-2.822 (3.644)	-2.436 (2.532)
SL3	8.363** (3.531)	-0.277 (2.948)	-0.98 (2.151)
SL4	-0.006 (0.055)	-0.644 (1.137)	-0.027 (0.049)
SL5	219700 (358800)	-733300 (587500)	0.000 (0.002)
SL6	-0.889 (0.822)	-0.626 (1.354)	-0.494 (0.581)
SL7		0.021 (0.024)	0.011 (0.014)
SL8	0.329**	0.559***	0.277***

Variable	t-1	t-2	t-3
	(0.13)	(0.116)	(0.096)
CA1	2647000		-0.067
	(4781000)		(2.973)
CA2	-0.069	-0.102***	-0.017
	(0.068)	(0.036)	(0.025)
CA3	356100		0.067
	(4977000)		(2.973)
CA5	-1.671	-1.534	-1.163
	(0.761)	(0.952)	(0.751)
CA6	-0.004	-0.008	-0.004
	(0.063)	(0.034)	(0.039)
E1	-219700	733300	0.000
	(358800)	(587500)	(0.002)
E2	5.88***	4.517***	3.328***
	(0.965)	(0.602)	(0.637)
E3	-0.029	0.004	0.004
	(0.026)	(0.028)	(0.02)
E4	-16.7***	-3.999	0.153
	(3.198)	(3.4)	(1.134)
E5	-2291000		-0.067
	(5587000)		(2.973)
RE1	-4.926***	0.939	-0.334
	(1.531)	(1.629)	(1.125)
RE2	-1.072**	-0.304	-0.338
	(0.454)	(0.317)	(0.272)
RE3		-0.496	0.794
		(1.34)	(1.007)
RE4	0.282	0.308*	0.315
	(0.318)	(0.185)	(0.211)
RE5	-22.42***	-2.935	0.002
	(3.365)	(3.171)	(0.193)
RE6	0.083	-0.346	0.068
	(0.472)	(0.39)	(0.265)
ACTIVO ^a	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
R1	0.003	-0.003	-0.002
	(0.016)	(0.024)	(0.006)
R2	22.44***	2.904	-0.006
	(3.362)	(3.171)	(0.193)
APOORTE Y CRÉDITO	0.482*	0.268	0.038
	(0.274)	(0.188)	(0.202)
SUPERV: Nivel 2	-0.731	-1.548***	-1.754
	(0.667)	(0.567)	(0.53)
SUPERV: Nivel 3	1.575*	-0.295	-0.42
	(0.929)	(0.596)	(0.541)
ASOCIADOS	0.000**		
	(0.000)		
SIZE: Mediana	-2.559	-3.987***	-6.932***
	(1.206)	(1.109)	(1.349)
SIZE: Pequeña	-3.346**	-4.047***	-7.81***
	(1.582)	(1.31)	(1.609)
SIZE: Micro	-0.834	-2.314*	-6.624***

Variable	t-1	t-2	t-3
	(1.612)	(1.348)	(1.67)
Constante	-575900	733300	10.57***
	(5011000)	(587500)	(4.042)
Observaciones	1125	1265	1220
Pseudo R2	0.608	0.333	0.296
Log likelihood	-375.284	-572.05	-574.372
AIC	816.6	1206.1	1216.7
Desviación Residual	750.6	1144	1149
Desviación NULO	1559	1714	1632
Método	Logit	Logit	Logit Bayesiano

^a Activo deflactado con IPC base 2008.

*** $p < 0,01$; ** $p < 0,05$; * $p < 0,1$. Errores estándar en paréntesis.

Fuente: Elaboración propia.

Tabla 33: Comparación del grado de error entre modelos. Logit. Submuestra 2002.

Modelo	Muestra de Estimación					Muestra de Validación			
	Muestra	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II	Éxito	Coef. Kappa	Error Tipo I	Error Tipo II
t-1	1126	85.88 %	71.73 %	14.46 %	13.75 %	86.12 %	72.14 %	11.56 %	16.42 %
t-2	1266	78.36 %	55.44 %	25.19 %	19.17 %	77.78 %	54.05 %	27.13 %	18.82 %
t-3	1221	76.17 %	49.69 %	31.58 %	18.90 %	78.95 %	55.81 %	26.27 %	17.74 %

El coeficiente Kappa se calcula como la probabilidad de éxito de cada escenario, ponderando por las probabilidades de ocurrencia.

Fuente: Elaboración propia.

Tabla 34: Test de Diferencia de Proporciones entre Modelos. MDA vs Logit. Submuestra 2002.

Modelo	Muestra de Estimación			Muestra de Validación		
	Error Tipo I	Error Tipo II	Error Global	Error Tipo I	Error Tipo II	Error Global
t-1	0.0000	0.0452	0.0000	0.0040	0.7575	0.0863
t-2	0.0277	0.2241	0.5341	0.2799	0.0841	0.7714
t-3	0.0654	0.5922	0.3978	0.8836	0.3944	0.6863

H₀: Proporciones iguales, H_a: Proporciones distintas, $p \leq 0,05$.

Fuente: Elaboración propia.