



UNIVERSIDADE DA CORUÑA

Facultade de Economía e Empresa

Trabajo de
fin de máster

Aplicación de
técnicas
estadísticas
multivariantes en R.

Determinantes de la
insolvencia en el sector
textil español.

Alba Avión Garrido

Tutores: David Peón Pose,
Xosé Manuel Martínez Filgueira

Máster Universitario en Banca y Finanzas.
Año 2021

Trabajo de Fin de Máster presentado en la Facultad de Economía y Empresa de la Universidade da Coruña para la obtención del Máster Universitario en Banca y Finanzas (MUBF)

Resumen

Este Trabajo de Fin de Máster (TFM) pretende encontrar un modelo de predicción del fracaso empresarial para PYMES del sector del comercio textil. Para ello, hace uso de la herramienta RStudio y las cuentas de balance de 535 empresas procedentes de la base de datos SABI en un periodo de 5 años (2014-2018).

De entre las numerosas técnicas de análisis multivariante revisadas, se decide contrastar dos de ellas, principalmente por su popularidad y sencillez: El análisis discriminante y el análisis de regresión logística. De su aplicación a la muestra se obtienen las variables clave a la hora de observar el balance de una empresa comercial textil y, al mismo tiempo, comprobar si este tipo de técnicas pueden presentar resultados satisfactorios para sectores no industriales.

Palabras clave:

Análisis discriminante, logit, modelos predictivos, análisis financiero, comercio textil.

Abstract.

This Final Master's Project aims to find a predictions model of business failure for SMES in the commercial textile sector. For it, it makes use of the RStudio tool and the balance sheet of 535 companies from the SABI database for a period of 5 years (2014-2018).

Among the multiple techniques studied in the analysis multivariate field, it was decided to contrast two of them, because of their popularity and simplicity. Discriminant analysis and logistic regression analysis. Thus, finding out which variables are critical when it is observed the balance sheet of a commercial textile company and, at the same time, checking whether this type of technique can present satisfactory results for non-industrial sectors.

Key words:

Discriminant analysis, logit, predictive models, financial analysis and textile trade.

Índice

Introducción.....	6
1. El sector de comercio textil.....	8
1.1 Evolución del comercio textil a nivel internacional y nacional	9
1.1.1 Situación del comercio textil a nivel mundial.....	9
1.1.2 Historia del sector textil en España.....	11
1.1.3 Situación actual del comercio textil en España.....	12
1.2 Principales estrategias en el sector de la distribución textil en España.....	17
2. Marco teórico.....	20
2.1 Análisis económico financiero empresarial (AEF).....	20
2.1.1 Técnicas de análisis económico-financiero.....	21
2.1.2 Revisión de la literatura de predicción de insolvencia	24
2.2 Modelos utilizados.....	26
2.2.1 Metodología de análisis discriminante	26
2.2.2 Metodología de regresión logit.....	29
2.3 Definición de insolvencia	32
2.3.1 El concurso de acreedores en España	34
3. Análisis empírico	37
3.1 Criterios de elección de la muestra	37
3.1.1 Subactividad y tamaño.....	37
3.1.2 Base de datos	38
3.1.3 Elección de las variables.....	39
3.2 Análisis descriptivo de la muestra	41
3.3 Factores del fracaso empresarial: modelo AD	46
3.3.1 Hipótesis requeridas en el AD.....	46
3.3.2 Resultados del análisis discriminante	52
3.4 Factores del fracaso empresarial: modelo logit.....	57
3.5 Análisis de los resultados obtenidos.....	61
Conclusiones.....	64
Anexo	70

Índice de gráficos

Gráfico 1: Evolución de las exportaciones a nivel mundial del textil y ropa en millones de dólares para el periodo de 1998 a 2018	9
Gráfico 2 : Participación de la demanda en el mercado de la ropa en todo el mundo desde 2005 hasta 2020	10
Gráfico 3: Evolución de la balanza comercial del sector textil (2000-2016).....	13
Gráfico 4: Evolución de la facturación del comercio textil en España (en millones de euros).....	14
Gráfico 5: Evolución del número de trabajadores y número de puntos de venta en el sector de la distribución	14
Gráfico 6: Origen de las importaciones de ropa de los países miembros de la Unión Europea en 2019	15
Gráfico 7: Las tres empresas con mayor volumen de negocio (Inditex, H&M y Primark) en el sector textil en España para el periodo 2010-2017. Datos relativos sobre el total del comercio de moda.....	16
Gráfico 8: Porcentaje del número de empresas según su número de empleados.	17
Gráfico 9: Evolución de la facturación por formato comercial (excluido canal online) ..	17
Gráfico 10: Ingresos de comercio electrónico por ventas online de ropa y confección en España del 2013 al 2019 (En millones de euros).....	18
Gráfico 11: Documentos subidos a la base Scopus por año que contienen "bankruptcy prediction models"	24
Gráfico 12: Punto de corte entre los grupos de un análisis binomial	28
Gráfico 13: Recta del MRLM ajustada a los puntos	31
Gráfico 14: Curva sigmoide.....	31
Gráfico 15: Número de Concursos en España por trimestre (2014-2018).....	35
Gráfico 16: Composición de la muestra en el año 2018 según la media por grupos de diferentes magnitudes (Tamaño, activo, ingresos y empleados)	41
Gráfico 17: Muestra dividida en función de empresas sanas o en quiebra para las variables ingresos y ROA en 2018	42
Gráfico 18: Perfil de la muestra según su activo, ingresos y empleo en el año 2018...	42
Gráfico 19: Muestra dividida en función de quiebra y su tamaño en el año 2018	43
Gráfico 20: Número de ingresos según tamaño de empresa para el periodo 2013-2018	45
Gráfico 21: Evolución del ROA entre empresas en quiebra y sanas para el periodo 2013-2018.....	45

Gráfico 22: Diagramas de cajas por niveles de factor de cada variable tras eliminar valores atípicos.....	48
Gráfico 23: Matriz de correlaciones	49
Gráfico 24: Histogramas de cada variable tipificada	51
Gráfico 25: Dispersión entre las variables endeudamiento, current y ROA.....	54
Gráfico 26: Resultados del modelo discriminante	54
Gráfico 27: Valor de los centroides por grupo.....	55
Gráfico 28: Matriz de confusión para el modelo discriminante	56
Gráfico 29: Resultados modelo logit utilizando la función best subsets	58
Gráfico 30: Valores de los Odd ratio para cada variable	58
Gráfico 31: Matriz de confusión para el modelo logit	59
Gráfico 32: Curva Roc.....	60
Gráfico 33: Función de coste asimétrica.....	61
Gráfico 34: Matriz de confusión con un punto de corte de 0,16.....	61
Gráfico 35: Evolución del Índice de Producción Industrial (IPI) industrial textil	70
Gráfico 36: Test de homocedasticidad.....	70
Gráfico 37: Resultados del contraste Manova para el modelo en conjunto.....	70
Gráfico 38: Poder de separación de las variables endeudamiento, current y ROA.....	71
Gráfico 39: Correlación de las variables con la función discriminante.....	71

Índice de tablas

Tabla 1: Clasificación por trabajo de las diferentes visiones acerca del fracaso empresarial	34
Tabla 2: Clasificación de la Comisión Europea de Mediana, Pequeña y Micro-empresa.	37
Tabla 3: Códigos CNAE relativos al sector textil.....	38
Tabla 4: Criterios establecidos para la selección de la muestra	39
Tabla 5: Parámetros característicos de cada variable	47
Tabla 6: Resultados para cada variable del contraste de normalidad	50
Tabla 7: Resultados del contraste Manova para cada variable	52
Tabla 8: Ratios utilizadas	71

Introducción

La industria de la moda en España es un sector dinámico y creciente, en el que los actores de mayor tamaño vienen ganando ventaja competitiva año tras año frente a las numerosas PYMES que lo componen. En un contexto caracterizado por la globalización de los mercados, la reducción del ciclo de vida de los productos, cambios cada vez más rápidos en las preferencias de los clientes, la incertidumbre y la creciente competencia, resulta conveniente anticiparse a posibles deterioros de la solvencia empresarial. Se pretende, por tanto, encontrar mecanismos que ayuden a prever como evolucionarán las distintas variables en el balance de una compañía, con el objetivo de encontrar patrones que señalen el proceso de deterioro de su situación económico-financiera.

El objetivo de este trabajo será realizar una aplicación práctica de dos técnicas estadísticas de análisis multivariante – análisis discriminante y regresión logística – a una muestra de PYMEs españolas del sector textil, con el propósito de identificar las variables clave que habrían anticipado el deterioro económico-financiero de las empresas que resultaron insolventes al final del período analizado.

Para ello, partiremos de un primer apartado en el que se tratarán aspectos teóricos acerca de la situación del sector y sus principales tendencias. Dada la complejidad de la cadena que integra el sector de la moda, este trabajo se dirigirá específicamente a la actividad comercial. Asimismo, nos centraremos en describir la evolución de las empresas que mayores índices de fracaso empresarial presentan, las pequeñas y medianas empresas, objeto del presente trabajo.

El segundo apartado del trabajo abordará la cuestión metodológica. Para hallar las variables que representan la diferencia entre el buen funcionamiento de una empresa y su fracaso se utilizará una metodología de análisis financiero multivariante que nos brinda la posibilidad de hacer un tratamiento cuantitativo de los datos más complejo que el de las técnicas de corte más tradicional. En este apartado, de revisión del marco teórico, examinaremos las distintas técnicas existentes, con especial atención a las aplicadas en este estudio – análisis discriminante y regresión logística. Además, dedicaremos un epígrafe específico a acotar la definición de fracaso empresarial, variable cualitativa clave en estos modelos.

El estudio empírico será el objeto del tercer apartado. En él, analizaremos las premisas necesarias para cada modelo y como la falta de estas afecta a los resultados. Posteriormente llevaremos a cabo la formación de los modelos a través del programa Rstudio y, finalmente, contrastaremos el poder predictivo de ambas técnicas a través de los resultados alcanzados por los modelos utilizando datos del año 2018.

Por último, realizaremos un breve repaso de los aspectos reseñables a lo largo del trabajo, comentando las conclusiones alcanzadas, las limitaciones existentes y las aportaciones personales a los conceptos estudiados en el máster.

1. El sector de comercio textil

La actividad del comercio ocupa el cuarto puesto entre los sectores que más aportan al Producto Interior Bruto (PIB) español, siendo ésta una de las razones por la que es considerado un sector estratégico dentro de la economía española¹ (Ministerio de Industria, s.f.). Aparte de la aportación monetaria que genera, es el sector que mayor empleo concentra, ya que según datos proporcionados por el Instituto Nacional de Estadística (INE), daba empleo a 3.153.498 de personas en 2018, suponiendo esta cifra un 17% del total de afiliados en la Seguridad Social. Asimismo, el perfil de contratos en el sector es uno de los más estables, por encima de la media de la economía (un 77,5% de contratos estables en el sector, frente a 63%² del conjunto de la economía). Por otra parte, genera un gran efecto arrastre sobre el resto de los sectores, produciendo un efecto multiplicador (Conferación Española de Organizaciones Empresariales, 2019).

En los últimos años, el sector está sufriendo una transformación propiciada por las grandes tendencias mundiales, como son la globalización, la digitalización y el cambio en los hábitos de consumo. Estos cambios traen consigo ventajas y desventajas a los distintos sectores, y no iba a ser menos para el sector que nos ocupa, el comercio textil. Para poder aprovechar las oportunidades que las nuevas tendencias proporcionan, se requiere invertir en aspectos clave en la adaptación y transformación del negocio. Esta inversión puede resultar más fácil de asumir para empresas de gran tamaño, pero no tanto para pequeños y medianos negocios. Sin una adaptación efectiva a estos nuevos paradigmas, el fracaso está asegurado, hecho que podemos condensar en el refrán “renovarse o morir”.

Se procede a mostrar una visión del sector del comercio de la moda y textiles en nuestro país en tres epígrafes. En primer lugar, un breve repaso del contexto internacional del textil. Segundo, la historia del textil en España y su contexto más reciente. Por último, comentamos los principales datos de mortalidad empresarial en el

¹ Según el Ministerio de Industria y Comercio. Información disponible en el siguiente link: <https://comercio.gob.es/ComercioInterior/Paginas/Index.aspx>. Último acceso: diciembre de 2020

² Información disponible en www.ccoo.es/51ba857456034a4944008f09a49b9512000001.pdf. Último acceso: diciembre de 2020

sector de la distribución textil en España. Con todo ello, pretendemos realizar una primera aproximación a las causas del fracaso empresarial de las PYME de comercio textil en España, poniendo de relieve la necesidad de creación de un modelo que ayude su diagnóstico prematuro.

1.1 Evolución del comercio textil a nivel internacional y nacional

1.1.1 Situación del comercio textil a nivel mundial

La industria de la moda tiene un peso relevante en el comercio global; concretamente, según información de la Organización Mundial del Comercio (OMC), en 2018 las exportaciones de “prendas de vestir”³ presentaban un valor absoluto de 494.076 millones de dólares, lo cual suponía un 3,76% del total de manufacturas, mientras que los “artículos textiles” suponían 312.644 millones de dólares (un 2,38% sobre el total de manufacturas). Por tanto, la suma de ambas categorías representaba más de un 6% sobre el total de exportaciones de manufacturas a nivel mundial. Además, el crecimiento medio interanual – en términos reales – de las exportaciones, tanto de prendas de vestir como de productos textiles, para el periodo 1998-2018 ha sido de un nada despreciable 5,02% – véase Gráfico 1.

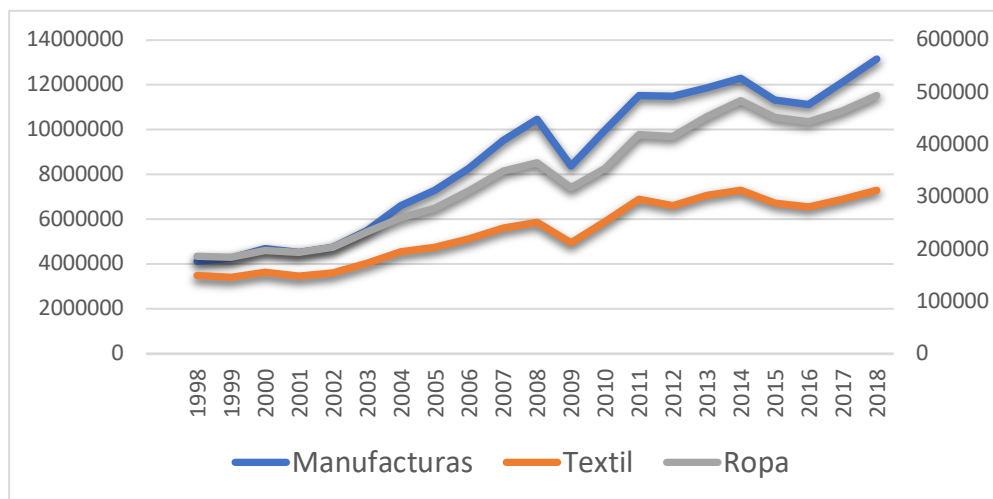


Gráfico 14: Evolución de las exportaciones a nivel mundial del textil y ropa en millones de dólares para el periodo de 1998 a 2018

Fuente: Elaboración propia a partir de datos de la OMC

³ En las estadísticas de la OMC se distinguen entre textiles y prendas de vestir. El primer término hace referencia a productos en cuya composición intervengan productos textiles, ya sea en su totalidad o solo en parte, mientras que el término “prendas de vestir” son productos textiles excluidos de la definición del anterior, destinados a servir como vestimenta.

⁴ La primera escala (eje izquierdo) hace referencia al volumen de Manufacturas, mientras que la segunda escala (eje derecho) indica el volumen del Textil y Ropa.

Después de los efectos negativos que sufrió el sector de la moda a nivel internacional, pudiéndose observar en 2009 una caída de las exportaciones (en el Gráfico 1 para todas las categorías), en los últimos años empezamos a ver un repunte con distintas intensidades dependiendo de la zona del mundo a la que hagamos referencia. Así, a escala mundial el gasto en prendas de vestir se ha incrementado un 72,8% entre 2009 y 2015 (EL COMERCIO JUSTO EN ESPAÑA, 2014). Los países que han presentado un mayor gasto en textil en 2018 fueron Estados Unidos y China, 374.595 millones de euros y 334.789 millones de euros, respectivamente.⁵

Como se puede observar en el Gráfico 2, según los datos proporcionados por Statista, la composición por regiones en la demanda de ropa se ha transformado en estos últimos años. Corroborado por el artículo de Mckinsey (2018), en 2017 el paradigma cambia por completo para Occidente: éste ya no se sitúa como líder de ventas de moda, debido a que más de la mitad de las demandas se originan fuera de Europa y América del Norte. Los focos de crecimiento se situarán en regiones como Asia Pacífico y América Latina, entre otras. Este hecho ocasiona en parte, que los grandes mercados mundiales reduzcan su tamaño, produciéndose así caídas de actividad en Alemania, Francia, Italia, etc. Mientras tanto, los países con mayor crecimiento en el sector textil en 2017, según EAE Business School, fueron: Indonesia, Irán, Vietnam, Corea del Sur, Uzbekistán, Japón e India.

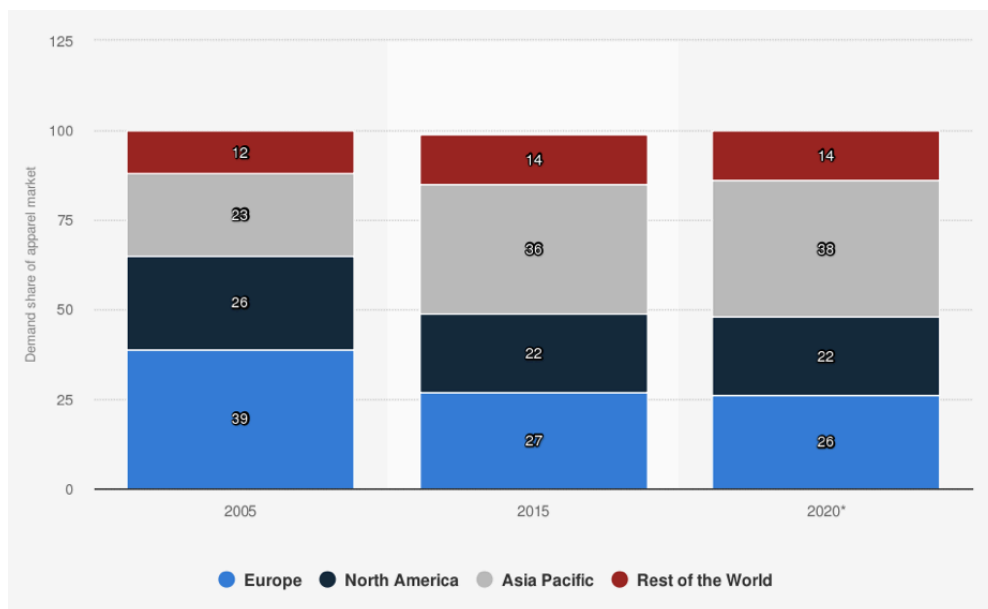


Gráfico 2 : Participación de la demanda en el mercado de la ropa en todo el mundo desde 2005 hasta 2020
Fuente: STATISTA

⁵ Información disponible en www.eae.es/actualidad/noticias/el-mercado-textil-espanol-alcanza-en-2015-los-28489-millones-de-facturacion. Último acceso: febrero de 2021.

1.1.2 Historia del sector textil en España

El sector textil en España fue un sector altamente proteccionista hasta finales de los años 60 y formado principalmente por PYMES. El crecimiento del sector era débil debido, entre otras cosas, a una orientación al mercado interno, con mínima presencia de las exportaciones y un bajo nivel salarial. La baja competencia producida por un mercado proteccionista contribuía a reducir las inversiones en innovación, mientras que el tejido empresarial formado fundamentalmente por pequeñas y medianas empresas⁶ provocaba una baja tasa de éxito empresarial (Costa Néstor Duch, 2004). Estas particularidades tuvieron un impacto importante en la economía en posteriores aperturas del mercado, además de producir una serie de problemas crónicos que llegan hasta nuestros días, como el bajo nivel tecnológico alcanzado y la baja tasa de inversión en I+D, provocando una dependencia tecnológica evidente (Sojo Calvo, 2012).

A partir de las dificultades económicas de los años 60 el sector sufrió una fuerte crisis, siendo el germen que da lugar a su posterior transformación. En la década de los 70, las empresas comenzaron a cambiar su estrategia, pasando de enfocarse en la maximización del volumen de producción a una estrategia más orientada a la venta y comercialización. En 1974 se llevaron a cabo los Acuerdos Multifibras (AMF) los cuales, según la Organización Mundial del Comercio, establecían un sistema de cuotas que limitaba en términos cuantitativos a los países importadores, para evitar así que el exceso de determinados productos perjudicase a la estabilidad de la balanza comercial. Este tipo de acuerdo pretendía proporcionar un clima de estabilidad en un mundo en plena globalización. Mientras tanto, existían algunos detractores que abogaban por la utilización de los acuerdos del Acuerdo General sobre Aranceles Aduaneros (GATT).⁷

A partir de entonces, el sector textil español ha sufrido numerosas reformas y reestructuraciones, con dos hitos en su historia más reciente: la entrada en la Comunidad Económica Europea (CEE) en 1986 y la liberalización mundial de productos textiles en 2005.

La entrada de España al proceso de integración económica en 1986 supuso en este sector, y en muchos otros, la apertura a un contexto más competitivo. Según el estudio llevado a cabo por Costa (2004), este hecho junto con una política monetaria restrictiva produjo una apreciación de la moneda que supuso una dificultad extra frente a los nuevos competidores. El mercado se inundó de productos más baratos pertenecientes

⁶ Estas no disponían de acceso crédito a largo plazo para poder prosperar como si los tenían otras empresas de mayor tamaño. Este hecho nos les permitía crecer e innovarse y por tanto no podían beneficiarse de economías a escala ni sinergias.

⁷ Véase https://www.wto.org/spanish/docs_s/legal_s/gatt47.pdf

a países en vías de desarrollo y de productos más sofisticados de otras economías de la Comunidad Europea, produciendo una caída interna de la producción.

Otro de los efectos de la entrada de España en la CEE fue el aumento de la renta per cápita y el nivel de vida de los españoles, permitiendo a las familias el acceso a una mayor cantidad de productos. Como era de esperar, este aumento salarial era beneficioso para las familias, pero no para las empresas autóctonas, ya que estaban obligadas a reducir su margen de beneficios para poder aportar un producto más barato y a la vez, proporcionar salarios congruentes con el contexto. Esta bajada del margen de beneficios junto con la entrada de nuevos productos foráneos más asequibles provocó el fracaso de muchas pequeñas y medianas empresas, aumentando así las cifras de desempleo (Costa Néstor Duch, 2004).

El segundo hito es la liberalización mundial del mercado de productos textiles. En 1995 se sustituyó el acuerdo AMF por el Acuerdo sobre los textiles y el Vestido (ATV)⁸, siendo este aplicable en un periodo de transición de 10 años, para finalmente, pasar a los acuerdos generales (GATT) en año 2005. Este hecho desencadenó grandes cambios en el sector textil, a los cuales se hizo alusión en el apartado anterior, con los países emergentes como China, Pakistán, Bangladesh, entre otros, afianzándose como líderes exportadores textiles, inundando los mercados occidentales de sus productos y haciendo de la deslocalización de la cadena productiva la única salida para las empresas autóctonas (Sojo Calvo, 2012). Como consecuencia de estas transformaciones, se produjo una caída de la actividad como había ocurrido en la entrada de la CEE – véase **¡Error! No se encuentra el origen de la referencia.** del Anexo

1.1.3 Situación actual del comercio textil en España

En España, el sector se considera estratégico debido a su elevado peso dentro del PIB (2,9%), y aportando un 13,2% al sector del comercio según datos de la Compañía Española de Seguros de Crédito a la Exportación (CESCE). Se posiciona como el cuarto sector que más exporta a nivel nacional y a su vez ocupa el quinto lugar entre los países de la Unión Europea, aunque, históricamente siempre ha presentado un saldo comercial deficitario (CESCE, 2019). Como se puede observar en el Gráfico 3, el volumen de transacciones ha ido elevándose a lo largo del tiempo. Las importaciones aumentaron de forma más acusada en los primeros años de la serie, mientras que, las exportaciones lo hicieron a un menor ritmo. El año posterior al estallido de la crisis y en el año 2013

⁸ Este acuerdo constaba de tres etapas con duraciones de: 3, 4, y 3 años, hasta finalizar los 10 años con la integración de todos los productos reflejados en el Anexo de las normas GATT.

podemos apreciar una disminución del déficit de la balanza comercial provocado por el estancamiento del crecimiento de las importaciones.

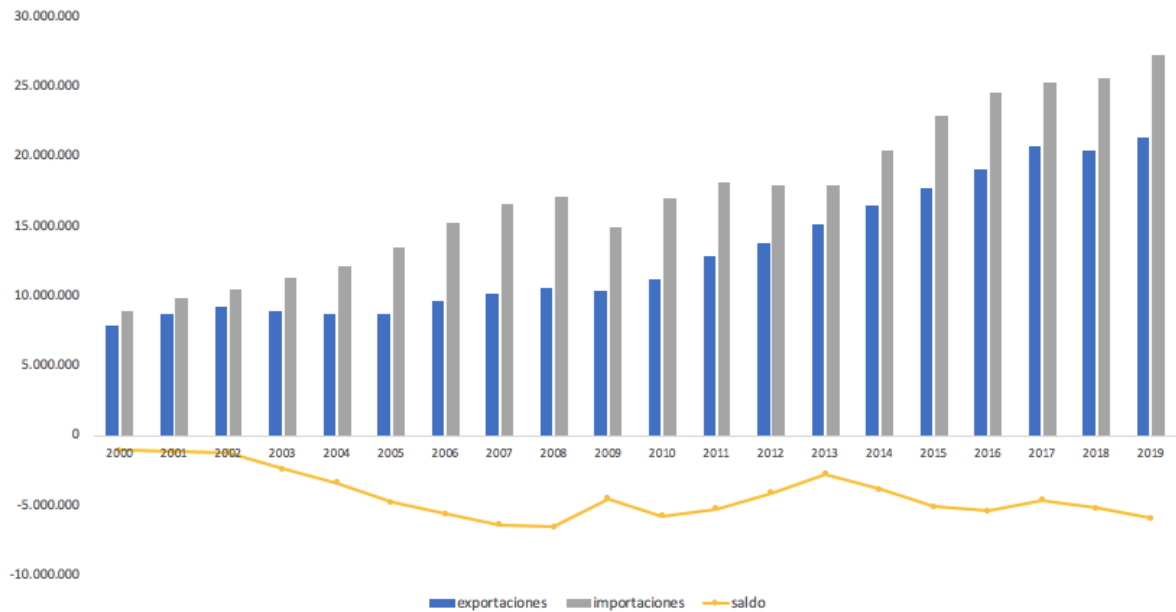


Gráfico 3: Evolución de la balanza comercial del sector textil (2000-2016)
Fuente: Elaboración propia a partir de datos de IDEPA

Actualmente el sector del comercio textil es un sector diverso, que integra múltiples niveles, desde grandes mayoristas hasta un amplio número de autónomos y microempresas. Además de esta variedad de tamaños, también nos encontramos con una gran diversidad en cuanto al público objetivo al que se dirige, desde la moda *low cost*, para todos los públicos, hasta la alta costura, orientado a un estrato de la población con una renta media-alta. Un análisis completo del sector supondría un mayor esfuerzo, debido a la multitud de categorías de producto que engloba y la complejidad de su cadena de valor, en la cual algunas empresas se especializan en una o varias fases. Por ello, como ya se expuso en la introducción, este trabajo se centrará en el último eslabón de la cadena: la comercialización del producto.

La parte del sector textil dirigida a la distribución ha experimentado también cambios a una gran velocidad. Nos situamos en un entorno global, marcado por una competencia creciente y una tendencia a la diversificación debido a la dinámica de las demandas de los consumidores, cada vez más exigentes. La venta del artículo se ha convertido en preocupación de primer nivel dentro de la cadena. Las empresas dedicadas a la actividad de la distribución presentan, en ocasiones, una mayor repercusión en la venta del producto que el propio producto en sí, ya que llevan a cabo una gran cantidad de acciones en las que mantienen contacto directo con el público. En este contexto, España es considerada la cuna del *fast fashion*, modelo que se consigue a través de una excelencia tecnológica y operacional. Pretende ofrecer prendas con una buena relación

calidad-precio, producción en series cortas y adaptándose a los gustos de los diferentes consumidores. Zara se posiciona como el *benchmark* en esta operativa.

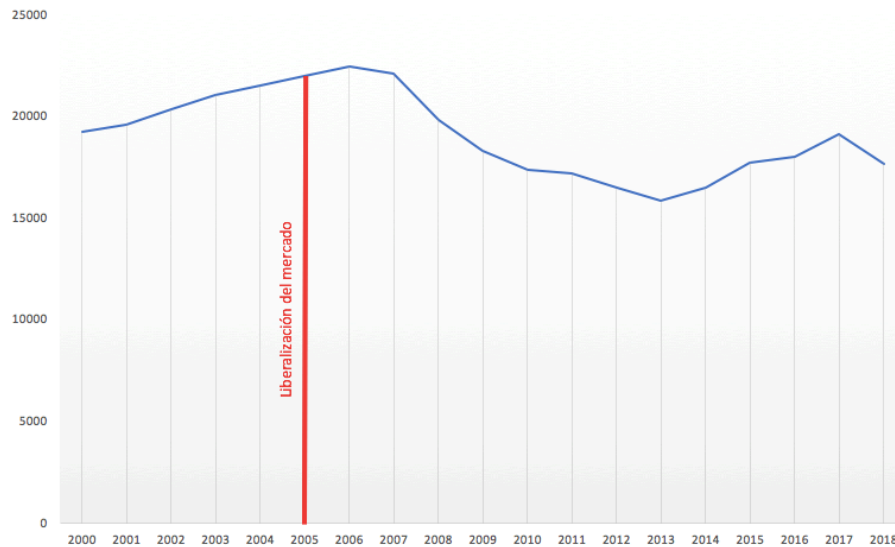


Gráfico 4: Evolución de la facturación del comercio textil en España (en millones de euros)
Fuente: Elaboración propia a partir de datos recogidos en el Informe anual de ACOTEX (2018)

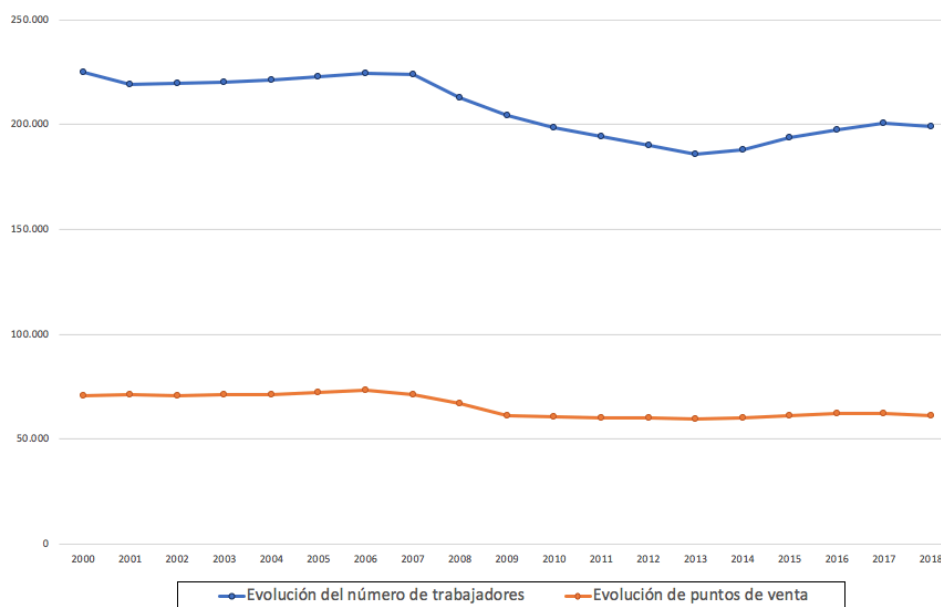


Gráfico 5: Evolución del número de trabajadores y número de puntos de venta en el sector de la distribución
Fuente: Elaboración propia a partir de datos recogidos en el Informe anual de ACOTEX (2018)

Como se puede observar en el Gráfico 4, la facturación del comercio textil ha sufrido una desaceleración de su ritmo de crecimiento a partir del 2006, cuando se alcanza un máximo de 22.450 millones de euros. Este frenazo en la facturación viene explicado, en parte, por la liberalización del mercado en 2005, con la entrada en vigor de los acuerdos GATT. La tendencia decreciente continuará hasta 2013, cuando se alcanza una facturación mínima, entorno a los 150 millones de euros (una caída acumulada del 29,43%) Siguiendo la misma dinámica, también se puede apreciar una disminución en

el número de puntos de venta y el número de empleados (véase Gráfico 5), alcanzándose un mínimo de 186 mil empleados en el sector en 2013 en 60 mil puntos de venta.

La apertura del mercado a un contexto mundial ha provocado un incremento en las importaciones de productos textiles, en su mayoría de países asiáticos (Gráfico 6). Los mercados experimentaron una saturación masiva con una ventaja en costes por parte de estos países, que contribuyeron al incremento del fracaso del pequeño comercio textil. Por eso, la solución de muchas empresas resultó ser la deslocalización de su proceso productivo hacia otros países donde la regulación en ciertos aspectos fuese más permisiva, como en el caso de la retribución de la mano de obra, permitiéndoles así, ofrecer sus productos a un menor precio.

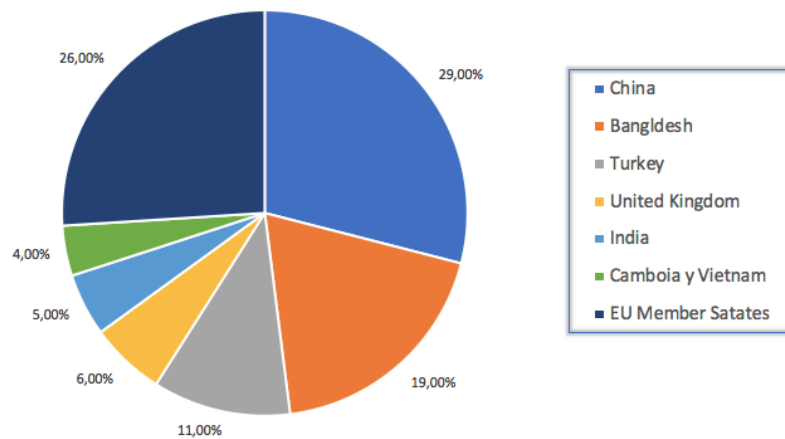


Gráfico 6: Origen de las importaciones de ropa de los países miembros de la Unión Europea en 2019
Fuente: Elaboración propia a partir de datos de Eurostat⁹

Otra fórmula utilizada para competir en esta coyuntura fue la integración de empresas, apostando por un sistema de integración vertical en el cual todos los procesos que debe llevar a cabo una prenda, desde el diseño hasta la venta, están controlados por una única empresa. Esta estrategia permitía crear una ventaja competitiva, centrando todos los esfuerzos en mejorar los resultados de forma global y no en cada proceso de manera independiente. Esta tendencia al control de todos los elementos de la cadena de suministros, desde el proveedor hasta el consumidor final, hizo que las grandes marcas fueran ganando terreno respecto al pequeño comercio. En España algunos de los precursores fueron El Corte Inglés y Cortefiel, a los que más adelante se le uniría Inditex (Goñi, 2015).

⁹ Se están teniendo en cuenta los 27 miembros de la Unión Europea a pesar de que Reino Unido se separó el 31 de enero de 2020. Pueden consultarse los datos en la web de Eurostat, en la siguiente dirección: <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20200424-1>. Último acceso en febrero de 2021.

Atendiendo a su origen nacional, las empresas que encabezan el mercado son 7 compañías, consideradas los mayores grupos de distribución españoles, destacando Inditex, Mango, Cortefiel y Desigual. En el Gráfico 7 se hace referencia, en particular, a los 3 mayores distribuidores de moda en España – Inditex, Primark y H&M, por orden de volumen de negocio. Obviamente, estas empresas no son PYME, por lo que no pertenecerán al segmento objetivo de este trabajo; sin embargo, es importante mencionarlas ya que mayoría de las empresas de la muestra que analizaremos están afectadas, en menor o mayor medida, por el devenir de estas compañías. Con lo cual, se puede observar una gran fragmentación en el sector liderado por unas pocas empresas de gran tamaño las cuales incrementan su volumen de negocio año tras año.

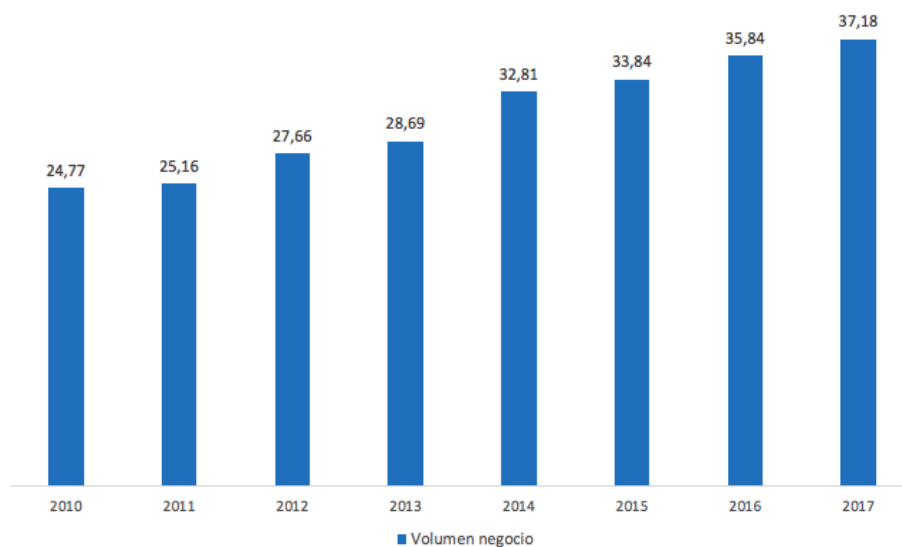


Gráfico 7: Las tres empresas con mayor volumen de negocio (Inditex, H&M y Primark) en el sector textil en España para el periodo 2010-2017. Datos relativos sobre el total del comercio de moda. Fuente: Elaboración propia a partir de datos de modaes.es

A pesar de que gran parte de la cuota de mercado está tomada por las grandes cadenas especializadas, el tejido empresarial español en el sector del comercio textil está integrado principalmente por pymes y micropymes. Este hecho se puede ver en las cifras de empleo en el Gráfico 8, con datos de IDEPA,¹⁰ ya que casi la mitad de las empresas del sector (45,2%) no emplea a ningún empleado, un 41,1% emplea entre 1 y 9 empleados, y solo el 13,7% emplean a más de 10 empleados (CESCE, 2019).

¹⁰ Disponible en www.idepa.es/detalle-oportunidad/-/asset_publisher/pZrNYOpxJB8w/content/sector-textil-el-sector-en-espana-informacion-general Último acceso: abril de 2021.

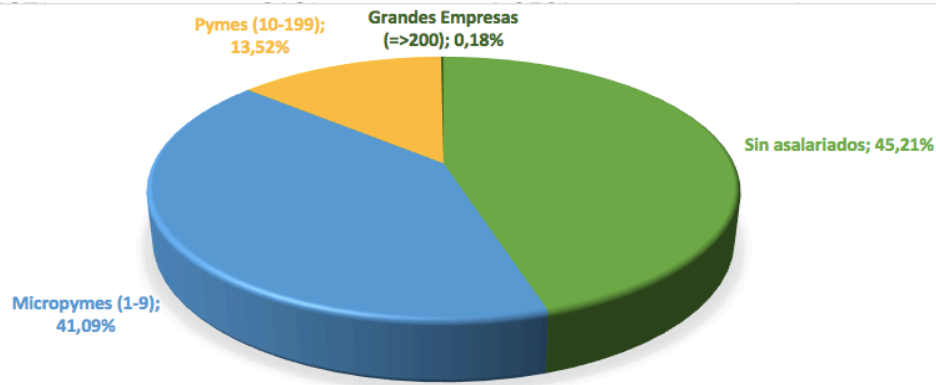


Gráfico 8: Porcentaje del número de empresas según su número de empleados.
Fuente: Elaboración propia a partir de datos recogidos en IDEPA

1.2 Principales estrategias en el sector de la distribución textil en España

Las grandes cadenas compiten con fórmulas que desplazan al pequeño comercio como lo es capacidad de mantener costes operativos bajos, debido entre otras cosas a un gran poder de negociación frente a los proveedores, y una gran rotación de productos que logra adaptarse a las necesidades de los consumidores. Su cuota de mercado ha ido incrementándose. Esta evolución es, entre otras cosas, debida a la estrategia llevada a cabo por estas grandes distribuidoras, que consiste en la apertura de nuevas tiendas, ampliando así su red comercial (Subdirección General de Estudios y Modernización del Comercio, 2007).

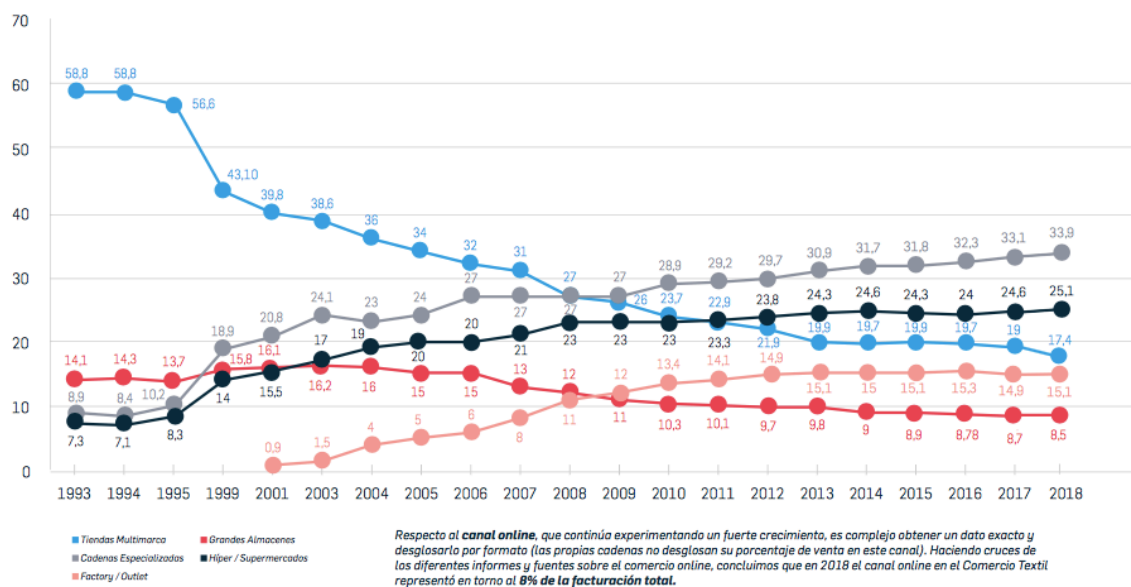


Gráfico 9: Evolución de la facturación por formato comercial (excluido canal online)
Fuente: ACOTEX (2018).

El Gráfico 9 muestra, con datos de la Asociación Nacional del Comercio Textil, Complementos y Piel (ACOTEX), como la facturación en los formatos de tiendas

multimarca y grandes superficies ha sufrido un descenso, mientras que outlet, cadenas especializadas y supermercados han experimentado un incremento de su facturación. Lo cierto es que las tiendas multimarca actualmente representan la parte más débil de este sector, viendo una clara tendencia de los hábitos del consumidor hacia puntos de venta con una única marca o centros comerciales. Este formato se convierte en el más vulnerable debido a una serie de problemas, entre los cuales destacan:

- Falta de personal cualificado, debido a unas condiciones de trabajo menos flexibles de las que pueden ofrecer otro tipo de grandes grupos textiles, en relación con horarios y salario.
- Incremento del precio de los alquileres, relegando este formato a establecerse en calles secundarias, no como las principales marcas, con poder adquisitivo suficiente como para situarse en los puntos estratégicos de las ciudades.
- Una capacidad de negociación frente a proveedores muy inferior a sus principales competidores. Limitados por su baja capacidad de compra, se abastecen con lo necesario para la temporada, corriendo el riesgo de una rotura de stock si el mercado repunta.

Ante estas dificultades, las principales estrategias del pequeño comercio son la especialización y adecuación en el mercado local (Subdirección General de Estudios y Modernización del Comercio, 2007).

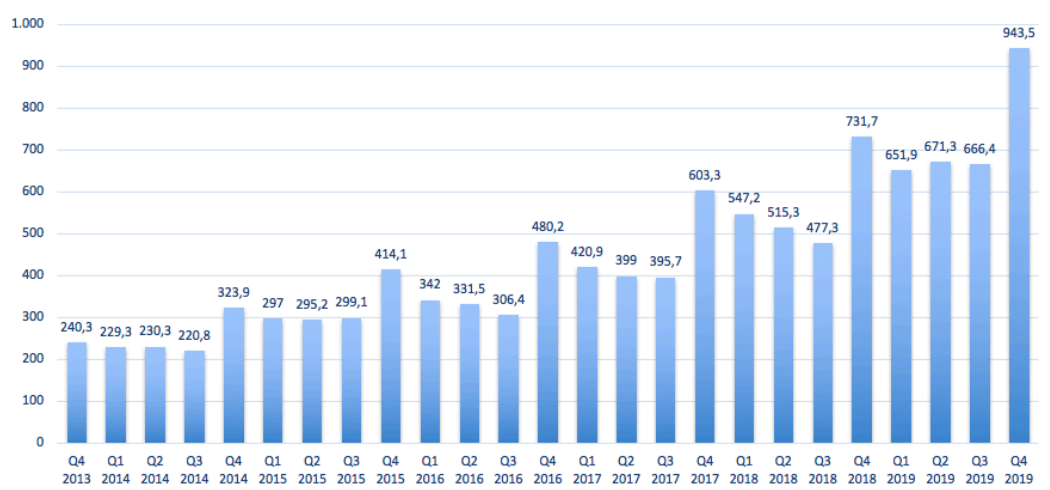


Gráfico 10: Ingresos de comercio electrónico por ventas online de ropa y confección en España del 2013 al 2019 (En millones de euros)

Fuente: Elaboración propia a partir de datos STATISTA.¹¹

Por último, es necesario hacer mención del canal online (Gráfico 10), ya que, a pesar de no estar en el Gráfico 9, es uno de los canales que está ganando notoriedad

¹¹ Disponible en: www.statista.com/statistics/444235/clothing-and-apparel-quarterly-e-commerce-revenue-in-spain/

los últimos años.¹² El e-commerce se presenta como una de las fórmulas que ayudaron a las empresas de este sector a recuperarse de la crisis. Uno de los factores que ha impulsado esta fórmula es el aumento en el uso de terminales móviles. Mientras tanto las páginas webs fueron adaptándose a los nuevos soportes, *smartphones* y *tablets* (Asociación Empresarial de Comercio Textil, Complementos y Piel, 2018).

La utilización de canales digitales está presente antes, durante o después de la compra. Gracias a la digitalización, los consumidores se han vuelto más exigentes en sus demandas, ya que pueden obtener gran cantidad de información acerca de ellas. Las nuevas tecnologías son una gran oportunidad para las empresas *low cost* y online, surgiendo nuevas fórmulas de promoción, como las figuras de los denominados *instagramers*. Estos efectos en los distintos formatos de venta comercial vienen impulsados por una serie de tendencias que se dan en el conjunto de la población:

- Un crecimiento exponencial de la población y un estrechamiento de las clases sociales concentrándose en mayor medida en la clase media-baja.
- El precio es un factor clave de competitividad debido, en parte, a la convergencia de la clase media mencionada en el punto anterior.
- Un incremento en el cambio del comportamiento del consumidor provocando así una época marcada por el dinamismo y la complejidad.
- Una mayor concienciación en temas relacionados con el medioambiente y los problemas sociales (Arribas Barreras, Josa, Bravo Durán, García Hiljding, & San Miguel Arregui, 2016).
- Un auge en el uso de las nuevas tecnologías para realizar las compras.

Como conclusión, podemos resumir tres efectos del proceso de transformación que se está produciendo en el sector de distribución comercial en España: un estancamiento de las ventas de grandes superficies; un fuerte crecimiento de las grandes cadenas y superficies especializadas; una pérdida de cuota y desplazamiento hacia el fracaso de las pymes que operan en este sector, sobre todo en las consideradas micro y el aumento en uso de los canales como internet y *e-commerce*.

¹² Debido a las dificultades que se producen a la hora de obtener los datos desglosados del canal online ACOTEX concluye que en 2018 la cifra del comercio online rondó el 8%.

2. Marco teórico

Los grandes cambios que están sucediendo en el sector dificultan la permanencia de algunos negocios. Para poder adelantarse a una quiebra o simplemente aprovechar las posibles oportunidades de negocio, la dirección de la empresa debe estar bien informada, con el fin de gestionar los recursos empresariales adecuadamente. Más aún, otros usuarios de la información financiera, como acreedores bancarios, accionistas, potenciales inversores o instituciones, estarán igualmente interesados en conocer la situación de la empresa.

Una de las herramientas más utilizadas para visualizar y analizar el estado de una empresa es el análisis económico financiero. En particular, el análisis financiero multivariante utiliza técnicas estadístico-económicas para condensar y procesar información de manera más eficiente que un análisis financiero de corte más tradicional. En este trabajo, trataremos este tipo de técnicas dentro del campo de estudio sobre predicción de insolvencia empresarial. Para ello, abordaremos primero cuestiones imprescindibles como acotar la definición del fracaso empresarial a utilizar, o la elección de variables que deben conformar la muestra. El objetivo de este apartado es analizar estas cuestiones para, posteriormente, proceder a analizar los resultados empíricos.

2.1 Análisis económico financiero empresarial (AEF)

La dirección de la empresa debe conocer en todo momento el estado económico-financiero actual de la compañía, su evolución, y contar con herramientas que ayuden a pronosticar la situación futura, con el fin de tomar decisiones más adecuadas y detectar potenciales contratiempos. Para llevar a cabo esta tarea, las empresas recurren al denominado análisis económico-financiero,¹³ entendido como “la aplicación de técnicas y herramientas analíticas en los estados financieros de propósito general y datos relacionados para obtener estimados e inferencias útiles en el análisis de negocios” (Wild, Subramanyam y Halsey (2007), p.4).

¹³ También denominado simplemente análisis financiero, diagnóstico económico, análisis de estados financieros, análisis de balances o análisis contable (Nogueira, y otros, 2017).

El proceso de análisis consta de 3 fases (García y Jiménez, 1996): primero, llevar a cabo una recolección e integración de información financiera, seguido de una formulación de hipótesis, y terminar con la obtención de unos resultados que sean útiles para la toma de decisiones. En síntesis, el análisis económico-financiero es un proceso de investigación y análisis cuantitativo aplicado con el fin de evaluar el desempeño económico y financiero en el momento actual o a lo largo de sucesivos periodos, valiéndose de información contable, la cual es a su vez transformada e interpretada mediante instrumentos analíticos.

Habitualmente se divide el objeto de análisis en sus dos partes; análisis económico y análisis financiero. El primero hace referencia a la medición de la capacidad de la empresa para generar beneficios y mide el rendimiento de los activos, mientras que el segundo hace referencia a la capacidad de hacer frente a las deudas en sus tres dimensiones – liquidez, solvencia y endeudamiento (Nogueira, y otros, 2017). En ambos casos, los estados más utilizados para recolectar la información tenemos el balance general y el estado de pérdidas y ganancias (P&G) o estado de resultados.

En cuanto a los métodos aplicados para la transformación e interpretación de la información, se encuentran: razones financieras, porcentajes integrales, tendencias monetarias mediante diferencia, tendencias de razones y porcentajes (análisis horizontales y verticales). En particular, el análisis financiero mediante cocientes fue una de las primeras herramientas utilizadas y una de las más habituales actualmente entre analistas contables y financieros. El cálculo de cocientes, también conocidos como ratios, razones o indicadores financieros, consiste en crear relaciones entre dos elementos de los estados contables. Entre los más destacados, encontramos los indicadores de liquidez, y solvencia, de eficiencia y actividad, de endeudamiento y rentabilidad (Nava Rosillón, 2009), indicadores todos ellos que emplearemos – y definiremos – en la parte empírica de este trabajo.

2.1.1 Técnicas de análisis económico-financiero

Existen diferentes métodos para llevar a cabo el análisis financiero. Mencionamos a continuación los más habituales, según Nava Rosillón (2009), haciendo hincapié en la última, debido a ser uno de los ejes principales del presente trabajo. Las dos primeras técnicas están enmarcadas en el análisis financiero tradicional. En todos ellos, las ratios financieras están presentes, pero mientras en las técnicas tradicionales representan el “output” del modelo, en el análisis financiero a través de técnicas estadísticas representan la materia prima, es decir, sus “inputs” (Ibarra Mares, 2009).

Análisis comparativo. Se pueden mencionar dos tipos: “el análisis seccional”, en el cual el *benchmark* o comparación utilizada puede ser un competidor de características similares o la empresa líder en el sector, y el “análisis de series de tiempo”, en el cual se compara contra los resultados del año anterior para comprobar el desempeño de la empresa a través de las tendencias que presente. También existe la posibilidad de combinar ambos análisis en el llamado “análisis combinado” (Gitman & Zutter, 2012).

Análisis Dupont. Es una técnica que permite encontrar las áreas clave del desempeño financiero de la empresa. Condensa el análisis de PyG y balance general en dos medidas de rentabilidad: *Return On Assets* (ROA) y el *Return on Equity* (ROE).

$$ROA = \text{Margen de utilidad neta} \times \text{Rotación total de activos}$$

$$= \left(\frac{BN}{Ventas} \right) \times \left(\frac{Ventas}{AT} \right)$$

$$ROE = ROA \times MAF$$

La primera expresión relaciona el Margen de utilidad neta, que mide la rentabilidad de las ventas y está compuesto por los Beneficios Netos entre el volumen de ventas, con la Rotación de Activos, la cual representa la eficacia de la empresa a la hora de utilizar su activo. Esta relación da lugar a la conocida medida de rentabilidad sobre activos ROA. La segunda fórmula relaciona el ROA con la rentabilidad para el accionista, el ROE, a través del multiplicador de apalancamiento financiero (MAF). La utilidad de este método radica en la posibilidad de descomponer la rentabilidad de una empresa en tres pilares: margen, eficiencia y apalancamiento.

Modelos estadísticos. Las metodologías anteriores proporcionan una interpretación, principalmente de carácter descriptivo, para un número limitado de empresas o periodos. Los modelos estadísticos, en cambio, permiten además realizar predicciones acerca de la condición futura de la empresa.

En el ámbito de la predicción del fracaso empresarial, en particular, los primeros estudios son de carácter univariante y su principal representante Beaver (1966), quien muestra que “las ratios financieras eran buenos predictores de las crisis empresariales, especialmente el cociente entre los Recursos generados por las operaciones y el Pasivo exigible” (p. 101). El propio Beaver admite que las ratios no son los únicos predictores de la quiebra de una empresa, pero presentan gran eficacia a la hora de anticiparse a una futura bancarrota – aspecto corroborado por Altman (1968), quien destaca su gran poder predictivo, disponibilidad y estandarización. En otros trabajos, además de las propias ratios, se han utilizado también sus desviaciones y valores relativos, y en ocasiones se han completado con datos macroeconómicos.

Existen diversos métodos a la hora de clasificar las técnicas estadísticas utilizadas. Una de las más populares es la clasificación de Stevens (1946), separando las variables según su tipología: variables métricas (de intervalo y razón) y no métricas (nominales y ordinales). No obstante, la clasificación que seguiremos versa en función del objetivo del análisis, ya sea “describir” o “predecir” (Fernández, 1986).

Las técnicas descriptivas (exploratorias) llevan a cabo labores de recolección, visualización, descripción y resumen de datos. En ellas no existe una hipótesis de partida, por lo que el objetivo es descubrir nuevos conocimientos mediante la reducción, diferenciación y clasificación en grupos homogéneos. Entre las principales técnicas descriptivas cabe destacar el análisis factorial y el análisis de componentes principales, permitiendo agrupar la información en un menor número de variables llamadas “variables latentes”. Otras técnicas, como el análisis clúster, tienen como propósito agrupar las diferentes unidades que componen la base de datos con características similares con el propósito de formar grupos lo más homogéneos posible.

Por su parte, las técnicas predictivas (confirmatorias) hacen uso de la inferencia estadística que se encarga de extraer patrones de una muestra y generalizarlos a la población, permitiendo así crear modelos y realizar predicciones. Al contrario que las primeras, estas si usan hipótesis de partida las cuales están basadas en un marco teórico previo (Ochoa Ramírez & Toscano Moctezuma, 2012). Las técnicas principales de carácter predictivo, y que aplicaremos en este trabajo, son el análisis discriminante (AD) y la regresión logit (RL).

El análisis discriminante tiene como objetivo analizar si existen diferencias entre grupos a través de una serie de variables, y de existir, explicar en qué sentido se dan, permitiendo clasificar posteriores observaciones. Dentro de la literatura de predicción, los modelos que usan AD representan el 30,3% del total (Korol, 2012). La técnica es especialmente utilizada en finanzas, desde su uso por Altman (1968) precisamente en el ámbito de la predicción de la quiebra empresarial. Por su parte, la regresión logit es un caso particular del modelo de regresión lineal tradicional, pero de carácter no lineal. Proporciona la estimación de la probabilidad de un suceso, en este caso, de quiebra empresarial, y permite identificar los factores que determinan dichas probabilidades y su importancia relativa sobre las mismas. A diferencia del AD, presenta restricciones estadísticas menos rígidas, ya que no tiene la necesidad de que los datos se distribuyan de acuerdo con una distribución normal.

El uso de ratios financieras tiene múltiples ventajas, pero presenta también varios inconvenientes a la hora de integrarse dentro de los modelos estadísticos. Así, según Gitman y Zutter (2012), en la selección de las ratios existe un componente subjetivo, los

datos deberían estar auditados – en particular pequeñas y medianas empresas sin obligación de someterse a controles externos como la auditoría. Además, el uso de ratios debería contrastarse con otros análisis (Cuervo, 1986). Debe existir además una homogeneidad en el tratamiento contable, por ejemplo, en el caso de la contabilización del inventario o la amortización (Fernández, 1986). La contabilidad no reconoce en sus cuentas elementos como los recursos humanos o capital intelectual, el mercado, el producto, etc. (Ochoa Ramírez & Toscano Moctezuma, 2012). Los estados no tienen en cuenta la inflación, por lo que deberían reexpresarse (Ibarra Mares, 2009). Y, por último, el uso de ratios suele presentar problemas estadísticos debido a que comparten componentes comunes y por ello presentan altas correlaciones (Nava Rosillón, 2009).

2.1.2 Revisión de la literatura de predicción de insolvencia

La afluencia de modelos de predicción de insolvencia durante el último siglo hasta ahora ha sido más que notable, como muestra el número de artículos encontrados mediante la búsqueda de “bankruptcy prediction models” en la base de datos Scopus (Gráfico 11). Por consiguiente, es conveniente una pequeña revisión de los modelos que han destacado en el ámbito de la predicción del fracaso empresarial.

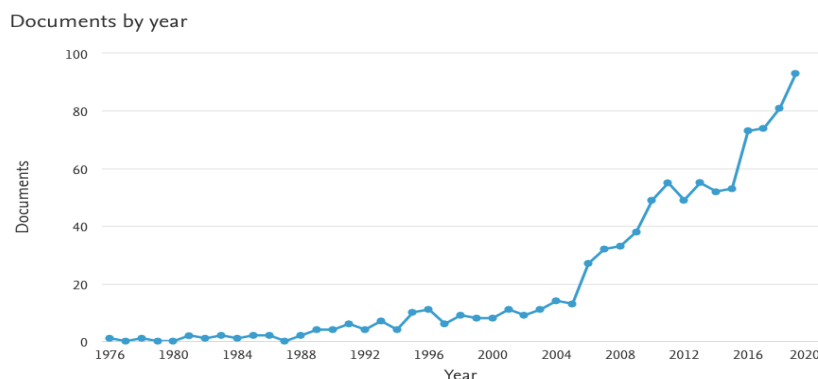


Gráfico 11: Documentos subidos a la base Scopus por año que contienen "bankruptcy prediction models"
Fuente: Scopus (2020)

Se puede clasificar los modelos que fueron surgiendo a lo largo del tiempo en: Modelos paramétricos y modelos no paramétricos. Los modelos paramétricos parten de la asunción de que el fenómeno objeto de estudio sigue una distribución de los datos conocida. Mientras tanto, en las técnicas no paramétricas se desconoce la distribución a priori, siendo su objetivo buscar cómo se distribuye dicho fenómeno.

Dentro de los modelos paramétricos destacan los ya mencionados modelos discriminante y logit / probit. El primer modelo multivariante paramétrico fue llevado a cabo por Altman en 1968. A pesar de la evolución en la complejidad de la metodológica

empleada en los últimos 30 años, posteriores modelos no han logrado tener unos resultados cualitativamente superiores (Correa Rodríguez, Acosta Molina, & Lorenzo González Pérez, 2003). En los últimos años se crearon variantes basadas en este modelo, como Bhandari (2013) o Alamamy et al. (2015), ambas con la inclusión del flujo de efectivo como variable explicativa. Otros autores, como du Jardín P. (2015), se enfocaron en la obtención de modelos con mejor desempeño en la predicción a largo plazo, ya que los modelos clásicos se sustentan en asunciones erróneas para su implementación en períodos largos, como el uso de variables históricas durante un periodo único o asumir que el proceso de quiebra es similar para todas las empresas. Finalmente, como alternativa al modelo discriminante nace la RL de la mano de Ohlson (1980). Este modelo fue mejorado por diversos autores con la introducción de variantes a lo largo de los años. Hernández y Wilson (2013) añadieron indicadores financieros clásicos combinados con la estructura de capital e indicadores macroeconómicos. Ben (2017) creo un modelo de regresión logística de mínimos cuadrados (RL-PLS) con el propósito de solventar deficiencias como la alta correlación entre variables y la no consideración de los *missing data* en la matriz.

La necesidad de modelos de predicción más complejos, posibilitando una mayor capacidad en el procesamiento de datos en entornos dinámicos, así como formalizar patrones ocultos en las estructuras de información, llevó a la formulación de los métodos no paramétricos. Korol (2013) compara técnicas no paramétricas, como los árboles de decisión y las redes neuronales (modelos de computación blanda), con el AD. Comprueba que todos los modelos presentan altas tasas de efectividad a la hora de predecir una quiebra, incluso en mercados inestables como los latinoamericanos, pero se aprecia cierta superioridad de la técnica del árbol de decisión frente a modelos estadísticos como el AD, e incluso las redes neuronales, a pesar de que estas dominan el campo de la predicción, representando el 9% del total de modelos. En general, los modelos no paramétricos resultan más eficaces, ya que no presentan tantas rigideces estadísticas y se adaptan a cualquier distribución de datos, pero ello trae consigo una mayor dificultad a la hora de su interpretación y realización.

Todos los trabajos mencionados analizaban empresas industriales, no comerciales, y pocos se han orientado al sector del *retail*. He y Kamath (2006) comprobaron que un modelo genérico como el de Ohlson pierde buena parte de su eficacia predictiva cuando se utiliza en el sector comercial. Por último, el modelo logit aplicado por Keener (2013) muestra que las empresas comerciales pequeñas y con menos trabajadores tenían una mayor probabilidad de quiebra. Además, bajos ratios de liquidez y bajos márgenes de *cash flow* anticipaban mayor probabilidad de bancarrota.

2.2 Modelos utilizados

Dado que las metodologías que aplicaremos en este trabajo son el AD y el RL, dedicamos un epígrafe específico a su explicación en mayor detalle.

2.2.1 Metodología de análisis discriminante

Como hemos visto, el AD es un método estadístico-econométrico que permite hacer una descripción, predicción y clasificación en grupos previamente establecidos (Perez Lopez, 2004). Es una técnica llamada de dependencia ya que en estos modelos se analizan las relaciones entre una variable dependiente categorías (Y_i) con un conjunto de variables independientes continuas (X_i). Partiendo de un conjunto de datos de N individuos (número de empresas) que se clasifican en q grupos (quiebra o sanas) utilizando para ello K variables explicativas (las distintas ratios), donde cada observación vendría representada por cada X_{ij} , podemos integrar estos elementos en una función discriminante $y = F(x_1, x_2, \dots, x_K)$ y resolver el problema de la clasificación en los q grupos mediante la ecuación:

$$D = u_1 \times X_1 + u_2 \times X_2 + u_3 \times X_3 \dots + u_k \times X_k$$

donde D representa el punto de discriminación, U_i es el factor de ponderación, y X_i son las variables independientes. Con esta metodología se pretende encontrar relaciones lineales entre las variables independientes para que consigan una mejor discriminación a la hora de pertenecer a una categoría u a otra. La variable dependiente va a establecer los distintos grupos previamente, pudiendo estos derivar de un análisis factorial o de conglomerados previo.

Considerando la existencia de n individuos pertenecientes al grupo de datos objeto de estudio, expresamos la ecuación de la siguiente manera:

$$D_i = u_1 \times X_{1i} + u_2 \times X_{2i} + \dots + u_K \times X_{ki} \quad i = 1, 2, \dots, n$$

donde D_i representa la puntuación discriminante de la observación i -ésima obtenida tras proyectar los puntos en el espacio K -dimensional formado por el valor de las variables de las n observaciones sobre el modelo. Expresado de forma matricial:

$$\begin{bmatrix} D_1 \\ D_2 \\ \cdot \\ \cdot \\ \cdot \\ D_n \end{bmatrix} = \begin{bmatrix} X_{11} & X_{21} & \dots & X_{k1} \\ X_{12} & X_{22} & \dots & X_{k2} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ X_{1n} & X_{2n} & \dots & X_{kn} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ \cdot \\ \cdot \\ \cdot \\ U_k \end{bmatrix}$$

En notación compacta: $d = Xu$

Formulando esta matriz en función de las desviaciones frente a la media y calculando la variabilidad de la función discriminante – es decir, la suma de cuadrados de las desviaciones respecto a su media – resulta, en notación compacta:

$$d'd = u'X'Xu$$

La matriz $X'X$, que vendrá representada por la asignación T , es una matriz simétrica que representa la suma de cuadrados total de las desviaciones de las variables independientes de la matriz X . Según la teoría del análisis multivariante, la varianza de la matriz $X'X$ se puede descomponer en matriz F , que nos indica la variabilidad entre grupos, y la matriz V , que indica la variabilidad intragrupo.

Por lo tanto $X'X = T = F + V$

$$d'd = w'X'Xw = w'(F + V)w = u'Fu + u'Wu$$

Tanto T como F y W son calculables con los datos de la muestra mientras que los pesos que representa cada variable en el modelo vienen representados por los coeficientes U_i , que serán hallados bajo el criterio de maximización siguiente:

$$\text{Maximización } \frac{\text{variabilidad entre grupos}}{\text{variabilidad intragrupos}}$$

Con este criterio se pretende encontrar el eje discriminante que permita una mayor dispersión posible entre los grupos (variabilidad entre grupos) y, al mismo tiempo, la menor dispersión dentro de las observaciones de cada grupo. En notación compacta aparecería como la siguiente formulación:

$$\lambda = \frac{u'Fu}{u'Wu}$$

Para obtener los u será necesario maximizar esta fórmula derivando λ respecto a u e igualando a cero. Esto nos va a proporcionar el primer eje discriminante $W^{-1}Fu = \lambda u$. Para la construcción del eje discriminante se tomará el vector asociado al mayor autovalor (λ), logrando así el máximo poder discriminante. El autovalor indicará la proporción de varianza total que está explicada por la función discriminante. La suma de los λ muestra la proporción total de varianza explicada por el modelo. El eje discriminante que se obtenga ayudará a clasificar un nuevo individuo en uno de los grupos formados por la variable dependiente, a través de la contrastación de su perfil de variables independientes con la función. Si el análisis discriminante es únicamente para dos categorías, como ocurre en este trabajo, no se necesita calcular otro eje.

Los centroides se encargan de resumir la información sobre el grupo al que pertenecen. Se expresan de la siguiente forma:

$$\bar{X}_I = \begin{bmatrix} \bar{X}_{1,I} \\ \bar{X}_{2,I} \\ \bar{X}_{K,I} \end{bmatrix} \quad \bar{X}_{II} = \begin{bmatrix} \bar{X}_{1,II} \\ \bar{X}_{2,II} \\ \bar{X}_{K,II} \end{bmatrix}$$

$$\bar{D}_I = u_1 \times \bar{X}_{1,I} + u_2 \times \bar{X}_{2,I} + u_k \times \bar{X}_{K,I} \dots$$

$$\bar{D}_{II} = u_1 \times \bar{X}_{1,II} + u_2 \times \bar{X}_{2,II} + u_k \times \bar{X}_{K,II} \dots$$

El punto de corte discriminante, C, se obtiene simplemente como la media aritmética de D_I y \bar{D}_{II} , resultando los siguientes criterios de clasificación: si $D_i < C$, se clasifica el individuo i en el grupo I, y si $D_i > C$, se clasifica al individuo i en el grupo II.

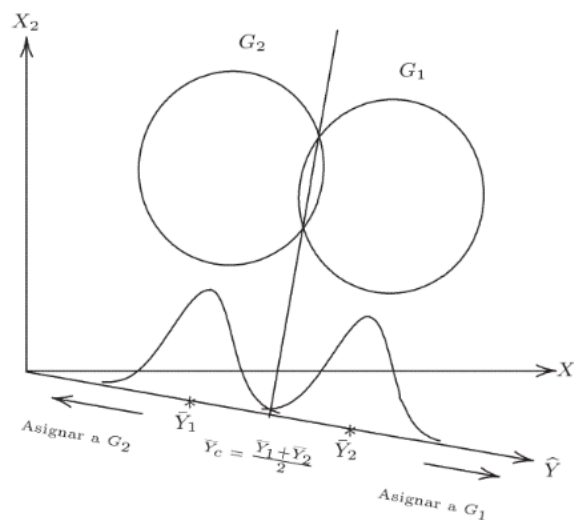


Gráfico 12: Punto de corte entre los grupos de un análisis binomial
Fuente: plepso.blogspot.com

Los objetivos del AD son principalmente tres (Cea D’Ancona, 2016). Primero, observar si la agrupación de datos es correcta y los grupos presentan características adecuadamente diferenciadas. Además, conseguir la función discriminante, consistente en la combinación lineal de variables independientes que haga máximas las diferencias entre grupos. Por último, obtener la probabilidad de dependencia de cada grupo a partir del valor de las variables que forman la función.

El AD tendrá la consideración de técnica descriptiva cuando se centre en la descripción de las variables que mejor diferencian los grupos, mientras que se considerará una técnica predictiva cuando se enfoque en determinar las funciones que sirvan para la clasificación de un nuevo individuo. Será necesario en todo caso un análisis descriptivo previo para tener claras las funciones de clasificación antes de llevar a cabo el análisis predictivo. La efectividad de la clasificación se evalúa posteriormente a través de una tabla cruzada donde se contrasta la clasificación que se hizo a priori y la pertenencia real al grupo. Por su parte, si el objetivo es inferir los resultados a la

población, será necesario el cumplimiento de una serie de hipótesis estadísticas. Cea D'Ancona (2016) menciona al menos cinco requisitos.

Primera, debe utilizarse una muestra de tamaño elevado. Existe controversia en cuanto a la consideración del número suficiente para un análisis efectivo, sin embargo, el más extendido es el que acepta 20 casos¹⁴ por cada variable independiente. Segunda, las variables dependientes deben ser cualitativas (no métricas), y las independientes cuantitativas (métricas). Además, no debe existir multicolinealidad entre las variables independientes, y deben seguir una distribución Normal, aunque el AD resulta robusto al incumplimiento de la normalidad si es por asimetría y no por atípicos. Otro aspecto es la presencia de homocedasticidad - es decir, que la varianza del error sea constante.¹⁵ Por último, cabe mencionar que para el AD se necesitan datos completos en las variables independientes, sino serán excluidos del análisis, mientras que, los denominados *missing data* repercutirán negativamente en los resultados.

2.2.2 Metodología de regresión logit

La segunda técnica tratada en el presente estudio es la regresión logística, la cual surgió de la necesidad de modelar fenómenos donde solo se permita dos tipos de respuestas. Este tipo de situaciones están muy presentes en la economía, jugando importantes roles en la teoría de decisión y el *management*.

Estos modelos pertenecen a los denominados modelos de variables dependientes limitadas o modelos de variable dependiente discreta, ya que la variable dependiente está acotada a un cierto número de valores, pudiéndose observar la distribución en su totalidad. Existen ciertas variantes, como el modelo logit multinomial, el modelo logit condicional, el modelo logit ordenado y el modelo logit anidado, que permiten que la variable discreta tenga más de dos respuestas. No obstante, en el presente trabajo nos

¹⁴ En esta proporción hay que tener en cuenta el procedimiento de validación, ya que para este se dividirá la muestra en 2: una parte servirá para la realización del análisis y otra para la validación. El porcentaje de la muestra que se dedique a cada parte varía en función del criterio del investigador.

¹⁵ Garantizar que los datos sigan una distribución normal se requiere para poder hacer uso de pruebas estadísticas como la F de Fisher y la t Student junto con la inferencia estadística. Cuanto mayor tamaño se trate la muestra, más probable es que cumpla este supuesto, por el teorema central del límite. En todo caso, para medir la normalidad de la distribución se hace uso de la prueba D de Kolmogorov-Smirnov (un valor próximo a 0,0 y p alto para rechazar la H0). Por último, la homocedasticidad suele relacionarse con el hecho de la normalidad y se da cuando la matriz de covarianzas intragrupos es la misma. Una de las consecuencias de no cumplir este supuesto es que los grupos tengan una mayor probabilidad de ser clasificados en grupos con mayor dispersión.

centraremos únicamente en el modelo logit binomial, ya que es el que mejor se ajusta al problema objeto de estudio (Pérez, Kizys & Manzanedo, 2010).

En esta función tendremos una variable respuesta (Y) que tomará valores 0 o 1 (proceso binomial), y una serie de variables independientes (X), cuantitativas o cualitativas, que explicarán el comportamiento de la variable dependiente. El análisis logit se asienta en las bases del modelo de regresión lineal múltiple (MRLM), con el objetivo de solucionar algunos problemas ocasionados por la variable dependiente binaria (o simplemente, categórica). Partiendo de la ecuación del MRLM donde la variable dependiente resulta continua tenemos:

$$Y = B_1 + B_2X_2 + \dots B_KX_K + u$$

Para el cual se estimarán los parámetros B, a través del método de Mínimos Cuadrados Ordinarios (MCO).

Consideremos el caso de que la variable dependiente resulte categórica binaria por lo tanto tendremos la siguiente ecuación:

$$E[Y] = 0P(Y = 0) + 1P(Y = 1) = P(Y = 1)$$

Utilizando el MRLM bajo el supuesto habitual de $E(u)=0$ e igualando las dos ecuaciones anteriores, obtendríamos la siguiente expresión, conocida como modelo de probabilidad lineal:

$$P(Y = 1) = B_1 + B_2X_2 + \dots B_KX_K = Y - u$$

de donde se deduce que la variable dependiente se puede expresar, por tanto, como la probabilidad de éxito más la perturbación:

$$Y = P(Y = 1) + u = E(y) + u$$

Este planteamiento implica varios problemas de tipo estadístico, que resumimos a continuación (Pérez et al., 2010). La perturbación se distribuye de forma discreta, no continua, por lo que no se cumpliría el supuesto de normalidad. Además, el término E presenta heterocedasticidad por lo que la estimación de los estimadores por mínimos cuadrados no será la más eficiente. El modelo lineal ajusta la nube de puntos alrededor de una recta, pudiendo proporcionar valores superiores a 1 e inferiores a 0, los cuales carecen de sentido a la hora de su interpretación.

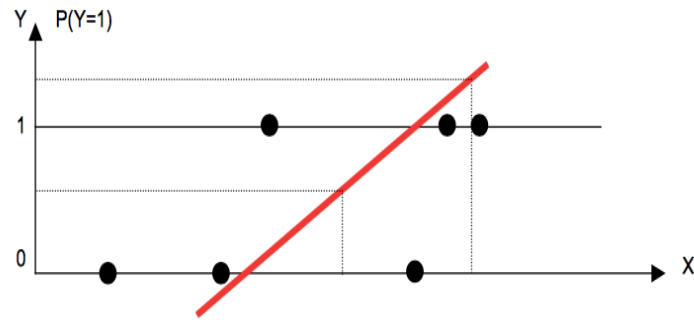


Gráfico 13: Recta del MRLM ajustada a los puntos
Fuente: Pérez, Kizys & Manzanedo (2010)

Además, la igualdad $P(Y = 1) = B_1 + B_2X_2 + \dots + B_KX_K$ expresa que la probabilidad de “éxito” viene determinada por la combinación de una serie de variables explicativas. Esta relación se podría expresar como $\frac{\partial P(Y=1)}{\partial X_i} = B_i \quad \forall i = 2,3,4, \dots, k$, con lo que estamos teniendo en cuenta que la respuesta de la variable dependiente va a estar marcada por incrementos o decrementos constantes de las variables independientes, sin tenerse en cuenta el punto de partida, lo cual resulta poco realista.

Debido a los problemas presentados por el modelo lineal se han introducido funciones no lineales a la MRLM:

$$Y = f(B_1 + B_2X_2 + \dots + B_KX_K) + u$$

Siguiendo el planteamiento anterior:

$$E[Y] = P(Y = 1) = f(B_1 + B_2X_2 + \dots + B_KX_K)$$

Por lo tanto, se buscó una función que pueda acotar los valores entre 0 y 1, correspondiéndose estos con las probabilidades. Para el caso de la regresión logística se utilizará la función logística o también conocida por curva sigmoide.

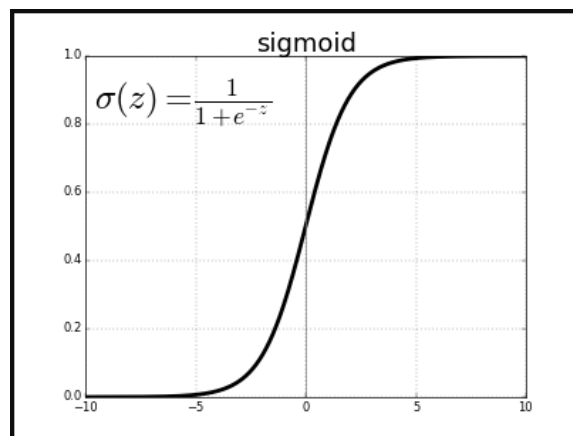


Gráfico 14: Curva sigmoide
Fuente: dspace.esPOCH.edu.ec

A diferencia de la regresión lineal múltiple que puede dar valores superiores a 1 o inferiores a cero, en la regresión logística se soluciona este problema utilizando una función cuyo resultado está comprendido entre 0 y 1, a través de la utilización de la función sigmoide (función logística). Por tanto, sustituiremos f por dicha función:

$$f(z) = \frac{e(z)}{1 + e(z)}$$

$$E[Y] = P(Y = 1) = \frac{e^{(B_1 + B_2x_2 + \dots + B_Kx_K)}}{1 + e^{(B_1 + B_2x_2 + \dots + B_Kx_K)}}$$

Los estimadores serán hallados mediante el método de Máxima Verosimilitud (MV), sin embargo, al contrario que sucedía con el modelo discriminante, estos no se podrán interpretar directamente, solo permitirán saber la dirección del efecto, pero no la magnitud de este. Para poder profundizar más en la interpretación de estos parámetros se ha recurrido al concepto conocido de las ratios *odds*.

$$Odds = \frac{P(Y = 1)}{1 - P(Y = 1)} = e^{(B_1 + B_2x_2 + \dots + B_Kx_K)}$$

Aplicando el logaritmo neperiano a la expresión anterior obtenemos:

$$Logit[P(Y = 1)] = \ln\left(\frac{P(Y = 1)}{1 - P(Y = 1)}\right) = B_1 + B_2X_2 + \dots + B_KX_K$$

Un modelo binomial tendrá entonces una probabilidad de éxito llamada P y una de fracaso llamada q ($1 - P$). En esta última ecuación podemos observar que a través de esto podremos interpretar los coeficientes B_k como variaciones en el logaritmo neperiano entre la probabilidad de éxito y la probabilidad de fracaso a partir de la variación en una unidad de los X_k . Así, los *odd* permitirán evaluar la influencia de cada variable dependiente sobre el fenómeno estudiado a través de razones.

2.3 Definición de insolvencia

Son múltiples las críticas que se atribuyen a las metodologías estadísticas de predicción de insolvencia. Entre ellas podemos mencionar las siguientes (Romero Espinosa, 2013). Primera, en su mayoría no están destinados a las empresas con mayor índice de fracaso y mayor peso en el tejido empresarial, como lo son las PYMES. Existe cierta arbitrariedad a la hora de escoger las variables de cada modelo, y la distribución de estas, rara vez satisface la hipótesis de normalidad requerida por la mayoría de las técnicas estadísticas. Son modelos principalmente cortoplacistas, por lo que para horizontes temporales amplios sus capacidades predictivas se reducen drásticamente. Con frecuencia no tienen en cuenta la proporción de empresas fracasadas y no

fracasadas en la población para luego replicar una proporción similar en la muestra o simplemente carecen de una validación posterior. Por último, no existe consenso a la hora de delimitar el concepto de “fracaso empresarial”, agravado por una normativa arcaica en materia concursal y falta de control sobre las PYMES a la hora de presentar sus estados contables por no tener la obligación de someterse a auditorías.

Nos centramos aquí, en particular, en este último hecho: buena parte de los trabajos académicos asumen un concepto de fracaso empresarial basado en la interpretación legal de quiebra. Sin embargo, esta no es la única alternativa posible – y, como veremos, en nuestro caso implicaría limitar la muestra y con ello la obtención de resultados. Por ello, dedicamos el epígrafe siguiente a abordar las distintas definiciones de insolvencia aplicadas, así como un breve repaso a la normativa concursal en España.

El fracaso empresarial se presenta como uno de los estados que la variable dependiente va a incluir en el modelo de predicción, por lo que delimitar este concepto es una pieza clave. En la literatura se observan diferentes perspectivas del fracaso empresarial, aunque existe un conjunto de términos que se utilizan como sinónimos de este, como lo son, bancarrota o quiebra. Esto quiere decir que, a pesar de no haber una definición exacta, si existe una terminología homogénea. Algunos autores como Bolaños (1997) distinguen dos tipos de fracaso empresarial: el fracaso económico y el fracaso financiero. Sin embargo, Según Graveline y Kokalari (2010), se contempla además una perspectiva más, la legal, existiendo 3 tipos de fracaso: “Fracaso económico, fracaso financiero o insolvencia técnica y fracaso legal”. Esta separación nos sirve para distinguir tres tipos de empresas: la empresa morosa, aquella que no satisface las cantidades después de 3 meses del vencimiento de una deuda; la empresa en situación de insolvencia técnica (es decir, con patrimonio neto negativo); y, la empresa que presenta concurso de acreedores, o en situación legal de suspensión de pagos o quiebra.

La mayoría de las definiciones recogidas en la Tabla 1 prestan especial importancia a la dimensión legal del concepto fracaso/quiebra. Otras interpretaciones son la incapacidad de pagar las deudas u obligaciones a corto plazo (e.g., Beaver, 1966), o la existencia de un patrimonio negativo (e.g., Altman, 1981). En la acepción de fracaso económico, es decir, la imposibilidad de hacer frente a deudas vencidas hace referencia a un estado que se presenta en el momento del estudio. Es decir, este concepto no tiene en cuenta la evolución futura, pudiendo ser fracasos temporales y no permanentes. La interpretación alternativa (e.g., Altman, 1981) hace referencia a un posible empeoramiento de las condiciones mediante la definición de quiebra técnica, más amplia que la quiebra legal (Espinosa, 2013).

Autor	Término	Definición
Beaver, 1966	Fracaso	Dificultad para atender deudas.
Altman, 1968	Quiebra	Catalogadas legalmente en quiebra
Altman, 1981	Quiebra	Insolvencia técnica-falta de liquidez
Taffler, 1982	Fracaso	Liquidación Voluntaria, orden legal de liquidación o intervención estatal
Laffarga, Martin y Vásquez 1987	Fracaso	Liquidación voluntaria o judicial.
García, Arques y Calvo-Flores, 1995	Fracaso	Empresa que haya solicitado apertura de expediente concursal de suspensión de pagos.
Correa, Acosta, González, 2003	Quiebra	Patrimonio neto negativo
Platt y Platt, 2004	Fracaso	Fracaso financiero
Calvo, García, Madrid, 2006	Fracaso	Riesgo financiero alto
Rubio Misas, 2008	Quiebra	Patrimonio negativo o quiebra técnica
Davydenko, 2010	Fracaso	Situación patrimonial refleja un valor reducido en los activos o escasez del flujo de caja.

Tabla 1: Clasificación por trabajo de las diferentes visiones acerca del fracaso empresarial
Fuente: Romero Espinosa (2013)

2.3.1 El concurso de acreedores en España

Aunque el concepto de fracaso empresarial que emplearemos en este trabajo no solo contemplará si la empresa presenta concurso de acreedores, entendemos necesario aportar una pequeña visión acerca de la situación de la figura del concurso de acreedores en nuestro país, mostrando así las ineficiencias que presenta y la necesidad de recurrir a otros avisos sobre el fracaso.

El actual concurso en España se presenta como un único procedimiento, si bien es verdad que existe el concurso ordinario y el abreviado, sin embargo, este último únicamente es una versión simplificada para pequeñas empresas. A pesar de la introducción de la nueva ley en 2003, acompañada del “principio de anticipación”, los concursos siguieron siendo escasos, tardíos y con la tendencia a acabar en liquidación. Además de estos factores, el *target* de empresas al que más afectaban era el de pequeñas empresas¹⁶ (Francisco, 2008).

¹⁶ El 85% de las empresas que se declaran en concurso tienen menos de 50 empleados. Fuente: www.elcomercio.es/v/20101116/economia/empresas-entran-concurso-acaban-20101116.html

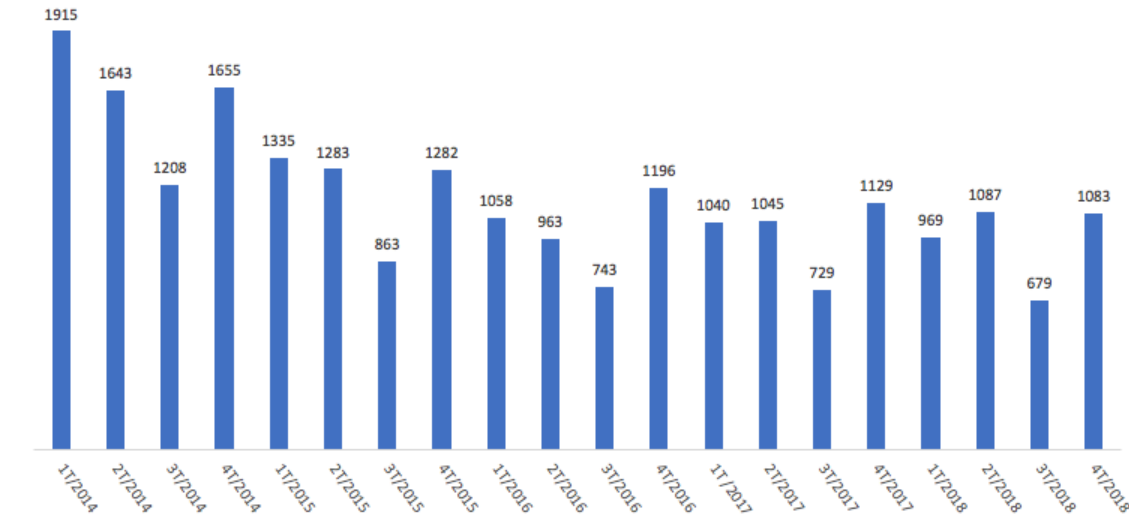


Gráfico 15: Número de Concursos en España por trimestre (2014-2018).
Fuente: Elaboración propia a partir INE

En 2016 España presentaba una de las tasas¹⁷ más bajas de empresas en concurso dentro de Europa según El Registro de Expertos en Economía Forense (REFOR). Incluso países como Portugal que tiene una menor dimensión económica (menor PIB y población), dobla el número de concursos de España.

Con la crisis surgida en 2008, los concursos aumentaron, pero no la cantidad que cabría esperar, evidenciado que aun en periodos de crisis, el mercado es reticente al uso del procedimiento concursal como un mecanismo más ante crisis empresariales. Cabe resaltar que el tipo de empresas mayoritario del tejido empresarial español, las microempresas (alrededor del 95% del total de empresas), ha presentado siempre las tasas más altas del uso del concurso, “en 2016 más de la mitad de los concursos (52,48%) son de micropymes. Además, si sumamos micro y pequeñas empresas llegarían hasta un 70% de los concursos en 2015 y 2016.” (REFOR, 2017, pág. 11)

Por otra parte, existen otras fórmulas que se llevan la atención del tejido empresarial español. Las pequeñas empresas optan por las renegociaciones privadas en vez de solicitar el concurso. Existen varios motivos para estas preferencias, el primero son los costes, como ya sabemos un concurso presenta un elevado coste para una empresa que ya se encuentra con problemas de liquidez, si a esto le sumamos que mayoría de los costes son fijos, hace que no sea una solución atractiva para pequeñas empresas. Otro motivo para la elección a favor de las renegociaciones es que las empresas pequeñas tienen, por lo general, un menor número de acreedores y una estructura de

¹⁷ Esta tasa fue hecha a través de dividir el número de empresa en concurso entre el número total de empresas en la economía. Se pueden consultar los datos en: <https://www.economistas.es/Contenido/Consejo/Estudios%20y%20trabajos/REFOR-%20Atlas%20concursal.pdf>

capital sencilla por lo que resulta más fácil llegar a acuerdos extrajudiciales (Banco de España, 2012).

En definitiva, el concurso sigue siendo una herramienta poco utilizada a pesar de considerarse, en teoría, un medio eficiente para la resolución de crisis patrimoniales, sobre todo de microempresas y PYMES. Este hecho se puede apreciar en la muestra que se maneja en este trabajo, donde el porcentaje de empresas en concurso en comparación con el de empresas “sanas” es relativamente bajo. Si bien es verdad, el periodo de los datos analizado en nuestro trabajo (2013-2018) presenta, en términos generales, una cierta estabilización y evolución positiva de la economía, haciendo que el número de empresas en quiebra se redujese en comparación al alcanzado en los años precedentes al 2008.

3. Análisis empírico

3.1 Criterios de elección de la muestra

3.1.1 Actividad y tamaño

Si bien el presente trabajo se orienta a las empresas del sector textil, no analizaremos el sector textil en su totalidad; en su lugar, nos centraremos en un número de compañías que cumplan una serie de pautas preestablecidas. En particular, tendremos en cuenta criterios referidos a tamaño de la empresa y a la actividad económica que esta realice.

Se ha puesto el foco en las denominadas “PYMES” (pequeñas y medianas empresas). El concepto PYME viene marcado por una serie de atributos: el número de trabajadores que emplea, el volumen de producción o de ventas, el valor del capital invertido y el consumo de energía (Castaño, 2000). Estos criterios son los más comunes; sin embargo, no son los únicos utilizados, por lo que para resolver esta cuestión se acudirá a la definición recogida en el BOE en el Anexo I Reglamento (UE) nº 651/2014 de la Comisión, el cual muestra como identificar de manera precisa que es considerado una PYME en España según una serie de límites financieros y número de trabajadores.¹⁸

Categoría de empresas	Efectivos	Volumen de negocio	Blance general
Mediana	<250	<= 50 mill. €	<= 43 mill. €
Pequeña	<50	<= 10 mill. €	<= 10 mill. €
Micro	<10	<= 2 mill. €	<= 2 mill. €

Tabla 2: Clasificación de la Comisión Europea de Mediana, Pequeña y Micro-empresa.
Fuente: Elaboración propia a partir de los datos del Reglamento 651/2014

La elección de este tipo de empresas tiene su justificación en que ocupan el mayor porcentaje de la red empresarial española. Son estas las que en mayor medida sufren

¹⁸ Información disponible en: www.ipyme.org/es-ES/UnionEuropea/UnionEuropea/PoliticaEuropea/Marco/Paginas/NuevaDefinicionPYME.aspx

los efectos de la crisis debido a su escasa capacidad de reacción y por ello, merecen especial atención dentro de este tipo de análisis (COLL-SERRANO, 2011).

Asimismo, pone su foco en la actividad del comercio textil y no de la creación de la prenda. Con la comercialización hacemos referencia a aquellas empresas que se centran en la venta de productos o servicios a los distribuidores, independientemente de su consideración como minoristas o otros mayoristas.

Esta elección se asienta en el hecho de que la mayoría de los modelos fueron creados para empresas de carácter industrial, las cuales difieren mucho, tanto en estructura económica como financiera, de las empresas no industriales (Mateos-Ronco, Marín-Sánchez, & Marí-Vidal, 2011). Un modelo pensado para una empresa industrial, en un principio, no debería ser extrapolable a una comercial, por lo que con este trabajo se podrá comprobar si los modelos multivariantes pueden resultar viables para empresas de esta índole.

En relación con la actividad se seleccionarán los siguientes códigos de la CAE resaltados en amarillo (Tabla 3):

Fabricación de la prenda	
Preparación e hilado de fibras textiles	131
Fabricación de tejidos textiles	132
Fabricación de otros productos textiles	139
Fabricación de calzado	162
Comercialización de la prenda	
461.- Intermediarios del comercio	
Intermediarios del comercio de textiles, prendas de vestir, peletería, calzado y artículos de cuero	4616
464.- Comercio al por mayor de artículos de uso doméstico.	
Comercio al por mayor de textiles	4641
Comercio al por mayor de prendas de vestir y calzado	4642
475.- Comercio al por menor de otros artículos de uso doméstico en establecimientos especializados.	
Comercio al por menor de textiles en establecimientos especializados	4751
477.- Comercio al por menor de otros artículos en establecimientos especializados	
Comercio al por menor de prendas de vestir en establecimientos especializados	4771
478.- Comercio al por menor en puestos de venta y en mercadillos	
Comercio al por menor de productos textiles, prendas de vestir y calzado en puestos de venta y en mercadillos	4782

Tabla 3: Códigos CNAE relativos al sector textil
Fuente: Elaboración propia a partir de datos del CNAE

3.1.2 Base de datos

Para la recolección de la información se ha escogido la base de datos SABI, la cual contiene datos financieros de más de 1,3 millones de empresas, españolas y portuguesas, de tal forma que cubre las necesidades del estudio al estar este dirigido a empresas de carácter nacional. La estandarización de los datos por medio de una misma estructura de balance para todas las empresas proporciona gran facilidad a la hora de tratar y depurar la muestra.

Un hecho reseñable es el tratamiento de los n.d. (datos no disponibles, *no data* o *missing data*) presentes en la muestra. Existe una amplia literatura acerca del tratamiento de los n.d. y como encontrar patrones para poder estimarlos; sin embargo, por simplicidad, y dado que no es objeto primordial de este trabajo, serán reemplazados por el valor “0”.

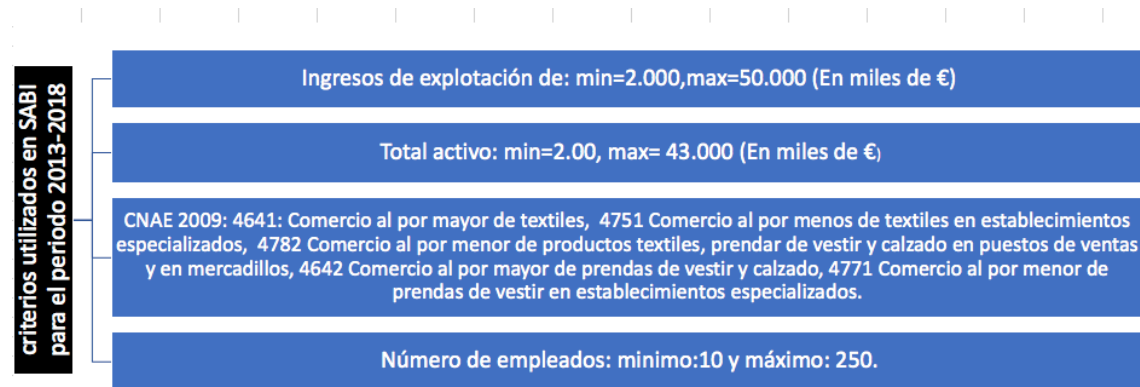


Tabla 4: Criterios establecidos para la selección de la muestra
Fuente: Elaboración propia

Aplicando los criterios de la Tabla 4, se ha obtenido un total de 789 empresas, las cuales se procedió a ordenar alfabéticamente y a enumerar. Del total de la muestra se han eliminado las compañías que presentasen como “Último año disponible” 31/12/2017 esto es así debido a que para realizar el modelo son imprescindibles datos completos de 2013 a 2017.

A la hora de depurar la base de datos, aparte de los criterios anteriores, también se introdujo una nueva restricción en el año utilizado para describir la muestra, 2018. Debido a la presencia de algunas microempresas (en algún año de la serie tuvieron la consideración de “pequeña”) que mostraban gran cantidad de datos no disponibles, se estableció que los Ingresos en 2018 debían ser superiores a 1 (los datos de la base se presentan en miles). Al realizar este proceso, se eliminaron prácticamente todas las empresas consideradas “micro” en 2018, permaneciendo solo 7 de esta categoría. Por proporcionalidad con el resto de los grupos, se decidió prescindir de ellas quedando la muestra conformada por 535 empresas clasificadas como “pequeñas” y “medianas”.

Finalmente, se procedió a realizar una media aritmética del periodo 2013 a 2017 con el objetivo de evitar posibles perturbaciones en los valores en años puntuales, pudiendo así suavizarlos y reduciendo el número de datos a tratar. Los datos del 2018 se reservaron para comprobar la eficacia del modelo una vez obtenido.

3.1.3 Elección de las variables

La primera variable por incluir en nuestro análisis es, obviamente, la condición de quebrada o fracasada para cualquier empresa de la muestra. Como vimos en apartados

anteriores, la elección de la definición de fracaso empresarial es otro de los aspectos controvertidos. En primera instancia, se optó por entender como empresa fracasada aquella que presente uno de los siguientes estados en SABI: concurso de acreedores, disuelta, liquidada o extinta. Sin embargo, la escasez de empresas resultante en nuestra muestra por la aplicación de esta definición de carácter estrictamente legal nos llevó a optar por una más amplia que Altman define como quiebra técnica, en la cual toda empresa con patrimonio neto negativo será también considerada como fracasadas.

En cuanto a las variables explicativas, nuestra muestra queda conformada por 15 ratios organizadas en 4 categorías principales de análisis (véase la lista completa y definiciones en el Anexo, Tabla 8). Hemos seguido a tal efecto trabajos como el de Tascón Fernández (2012) con el fin de comprobar las ratios utilizadas más frecuentemente. Las categorías son:

- Indicadores de liquidez y solvencia. La primera referida al efectivo necesario para amortizar las deudas a corto plazo; la segunda referida al conjunto de bienes y recursos que se posee para amparar toda deuda asumida.
- Indicadores de eficiencia y actividad. Miden cuan eficientemente se gestionan insumos, activos y procesos. La eficiencia relaciona el output (producto final) y los inputs necesarios utilizados en su creación. Los indicadores más utilizados son los de rotación (de activos totales, activos fijos, inventario, cuentas por cobrar, etc.). En particular, la rotación de inventario nos ayudará a evaluar el impacto de las *fast fashion* españolas sobre las pequeñas empresas. Asimismo, en la revisión teórica se ha observado una clara tendencia a la baja en el número de empleados. Las ratios de personal o de productividad ayudarán a determinar el nivel de eficiencia relativa de las empresas pequeñas y micro en este sentido.
- Indicadores de endeudamiento. Un endeudamiento excesivo podría provocar una incapacidad de pago futura. Se debe evaluar la cantidad de deuda contraída, tanto a corto (c/p) como a largo plazo (l/p). Entre los indicadores más utilizados tenemos: razón deuda, razón pasivo l/p y pasivo total, y cobertura de intereses.
- Indicadores de rentabilidad. Gitman (1997) define la rentabilidad como la relación entre costes e ingresos obtenidos por el uso de activos en actividades productivas. Estos beneficios se pueden medir en relación con las ventas, al capital o los activos totales, obteniendo los indicadores de rendimiento sobre ventas, sobre activos (rentabilidad económica), y sobre el capital aportado por los propietarios (rentabilidad financiera).

3.2 Análisis descriptivo de la muestra

Se partirá del año 2018 como año de referencia para proceder a la caracterización de la muestra. Por tanto, los gráficos que se exponen a continuación emplean, preferentemente, la información contable de las empresas referida a dicho ejercicio. El objetivo de esta breve caracterización es describir los trazos principales que poseen las empresas que componen nuestra base de datos.

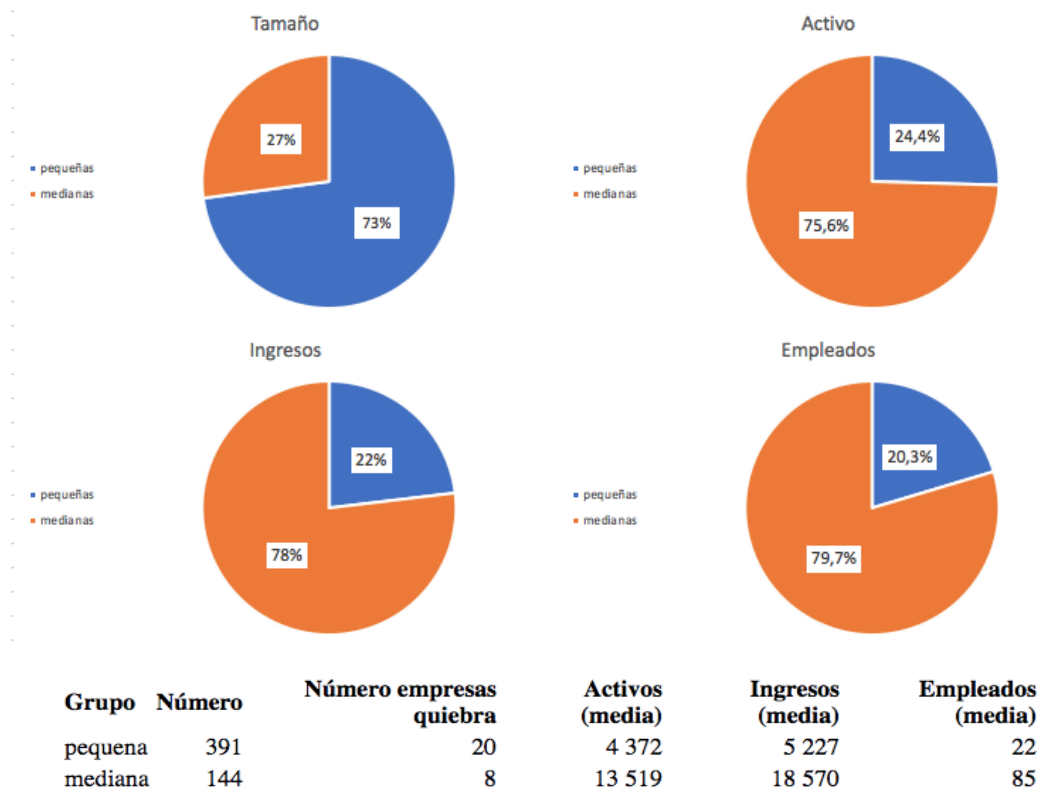


Gráfico 16: Composición de la muestra en el año 2018 según la media por grupos de diferentes magnitudes (Tamaño, activo, ingresos y empleados)
 Fuente: Elaboración propia a partir de datos de SABI

Se puede observar (Gráfico 16) que las empresas catalogadas como “pequeñas” representan el grueso de la muestra, con un 72,9% (452 empresas) del total. Esto es lógico, ya que como se ha visto anteriormente, el tejido empresarial en España está compuesto fundamentalmente por microempresas y pequeñas empresas. En cuestión de ingresos por grupo, las empresas “pequeñas” obtiene 4 veces menos ingresos de media que las empresas “medianas”, muy acorde con la inversión en otras magnitudes como empleados y activo donde la proporción se mantienen prácticamente semejantes.

A continuación, se procedió a dividir la muestra en función su estado, consideradas en “quiebra” o consideradas “sanas” para analizar la rentabilidad media de cada categoría.

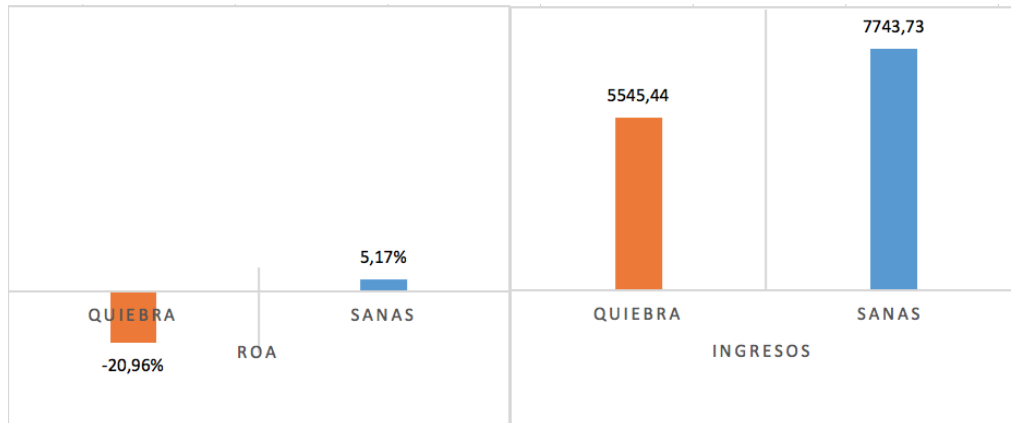


Gráfico 17: Muestra dividida en función de empresas sanas o en quiebra para las variables ingresos y ROA en 2018
Fuente: Elaboración propia a partir de datos de SABI

Atendiendo a la Gráfico 17 observamos que las empresas en quiebra muestran rentabilidades negativas, con cifras promedio 5 veces menores (un 26,13% menos) que las empresas sanas, mientras que la diferencia de ingresos entre ambas categorías únicamente alcanza un cuarto (28,39%).

Como pudimos observar en el marco teórico, el análisis Dupont analiza el ROA a través del Margen neto multiplicado por rotación de activos. Por ello, para indagar en el hecho de que unas empresas obtengan un ROA negativo y otras positivo, llevamos a cabo el Gráfico 18. Este está compuesto por las empresas que presentaron mayores márgenes netos¹⁹ en el año 2018. Así, analizaremos que diferencias existen en las magnitudes, activo, ingresos y empleados, para que algunas empresas resulten viables (ROA positivo) y otras no (ROA negativo) presentando unos márgenes netos semejantes y mejores que la media muestral.

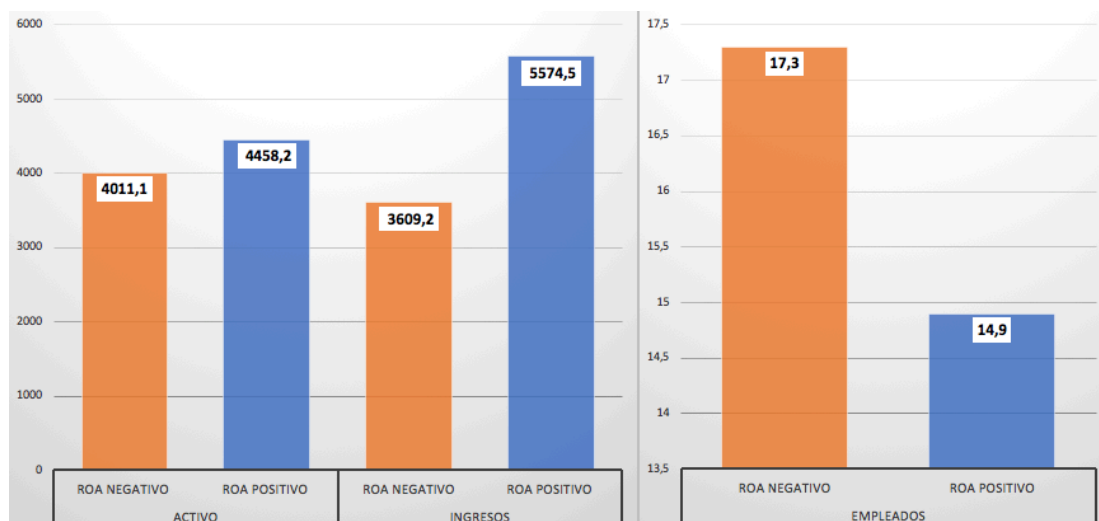


Gráfico 18: Perfil de la muestra según su activo, ingresos y empleo en el año 2018
Fuente: Elaboración propia a partir de datos de SABI

¹⁹ Se procedió a ordenar las empresas de mayor a menor margen neto, tomando la mitad superior de la muestra.

Como se puede contemplar, las empresas con ROA positivo lo hacen por una alta diferencia en el número de ingresos, ya que, con semejantes recursos (activo y empleados) consiguen unos ingresos medios superiores en 1,7 millones € frente a las empresas no viables (ROA negativo). Además, las empresas con ROA positivo utilizan de media 2,5 empleados menos que las empresas con ROA negativo

Haciendo una comparación con la muestra en conjunto (Gráfico 16) y este Gráfico 18, las empresas con mayor Margen Neto presentan un menor número de empleados de media. Se puede advertir que las medias de ambas categorías (ROA positivo y negativo) son relativamente bajas, teniendo en cuenta que la media de empleados total de empresas “pequeñas” es de 22 empleados (vista en el Gráfico 16).

Para poder comprobar si las diferencias entre empresas viables y no viables con un margen neto semejante se dan por el segundo elemento de la ecuación de Dupont, tenemos el Gráfico 19:

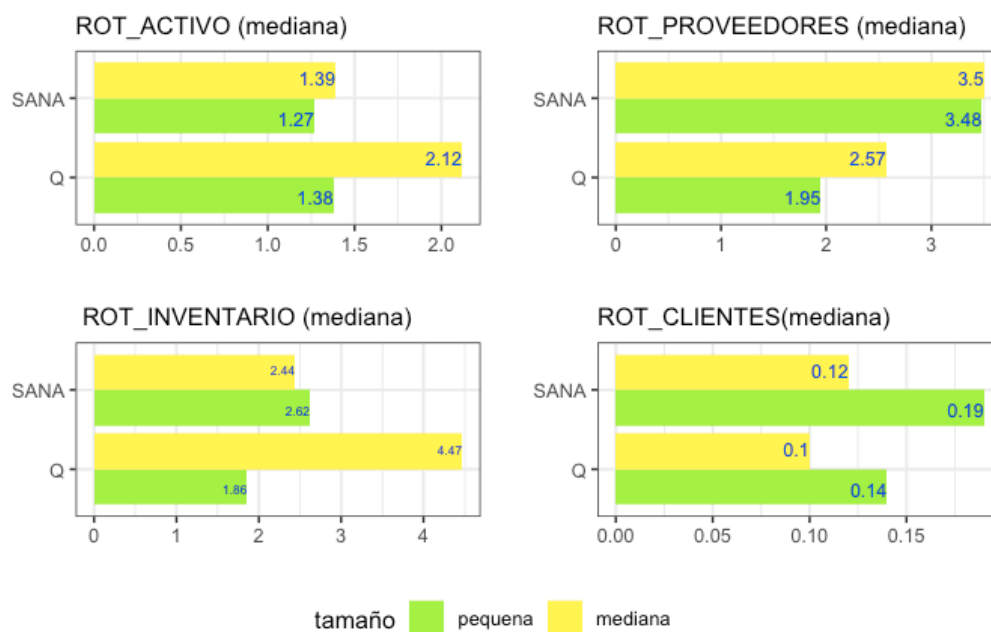


Gráfico 19: Muestra dividida en función de quiebra y su tamaño en el año 2018
Fuente: Elaboración propia a partir de datos de SABI

Las rotaciones en el sector *retail* pueden estar estrechamente relacionadas diversos elementos, entre ellos: el fenómeno des *fast fashion* en la gestión de inventarios (rotación de inventarios), la efectividad de su inmovilizado material la hora de obtener ingresos (rotación de activo) o incluso el nivel de poder sobre proveedores (rotación de proveedores). Con todo, contrastamos en el Gráfico 19 las empresas sanas y en quiebra a lo largo de las diferentes rotaciones, con el objetivo de encontrar algún patrón que las diferencie.

Atendiendo a la rotación de activo de las empresas “pequeñas, se pueden apreciar pocas diferencias entre empresas sanas y empresas en quiebra; sin embargo, si podemos apreciar una diferencia notable en las empresas “medianas”. Para poder ahondar en este fenómeno se llevaron a cabo una serie de observaciones llegando a la conclusión de que este suceso era debido a la drástica reducción del Activo, manteniendo unos ingresos constantes o ligeramente en declive, con lo que, la ratio aumentó su valor.

En cuanto a la rotación de inventario sucede un caso parecido al antes mencionado, apreciando altos valores en las empresas medianas en quiebra, con lo que también se procedió con su análisis. En este caso, la media obtenida es debida a dos empresas que presentan valores extraordinariamente altos; no obstante, estos valores no son puntuales: ambas empresas experimentaron un crecimiento en la ratio de inventario a lo largo del periodo, por lo que se decidió tenerlas en cuenta igualmente.

La rotación de proveedores presenta dos interpretaciones posibles: puede ser vista como el nivel de eficiencia de la empresa a la hora de cumplir con sus obligaciones a c/p, es decir, un endeudamiento excesivo podría entenderse como una situación de iliquidez, o una segunda interpretación con connotaciones positivas, referente al aprovechamiento del fenómeno del apalancamiento mediante la financiación a través de proveedores, es decir, sustituir un posible préstamo a c/p con una entidad bancaria por una deuda con proveedores la cual carece de intereses. Atendiendo al caso particular de nuestra muestra, la rotación de proveedores muestra diferencias considerables entre ambos estados, percibiéndose valores más elevados en empresas “sanas”, por lo que la lógica apuntaría a la segunda interpretación expuesta.

Concluimos el análisis descriptivo con una breve referencia al comportamiento temporal de las empresas durante el periodo analizado. En primer lugar, el Gráfico 20 muestra la evolución temporal de los ingresos medios de las dos categorías de empresas – pequeñas y medianas – entre 2013 y 2018.

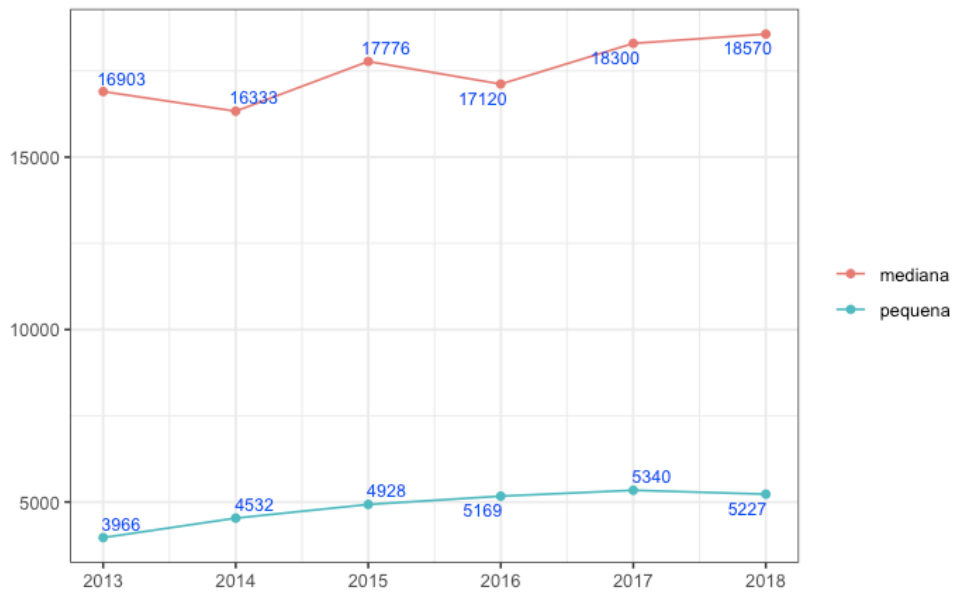


Gráfico 20: Número de ingresos según tamaño de empresa para el periodo 2013-2018
 Fuente: Elaboración propia a partir de datos de SABI

Se aprecia un ligero aumento de estos, más evidente en las empresas “pequeñas” que en las empresas “medianas”. Esta evolución se ralentiza con los años, pudiéndose examinar un cierto estancamiento en el último año de ambas series. La tendencia aquí mostrada resulta muy acorde a los gráficos del sector en conjunto, los cuales indicaban el comienzo de una pequeña recuperación.

Con la finalidad de observar si dicha recuperación vista en el Gráfico 20 también se corresponde con mayores rentabilidades tenemos el Gráfico 21. Este procede a dividir las empresas sanas y en quiebra a lo largo del periodo (2013-2018) y mostrar la evolución de su rentabilidad económica (ROA).

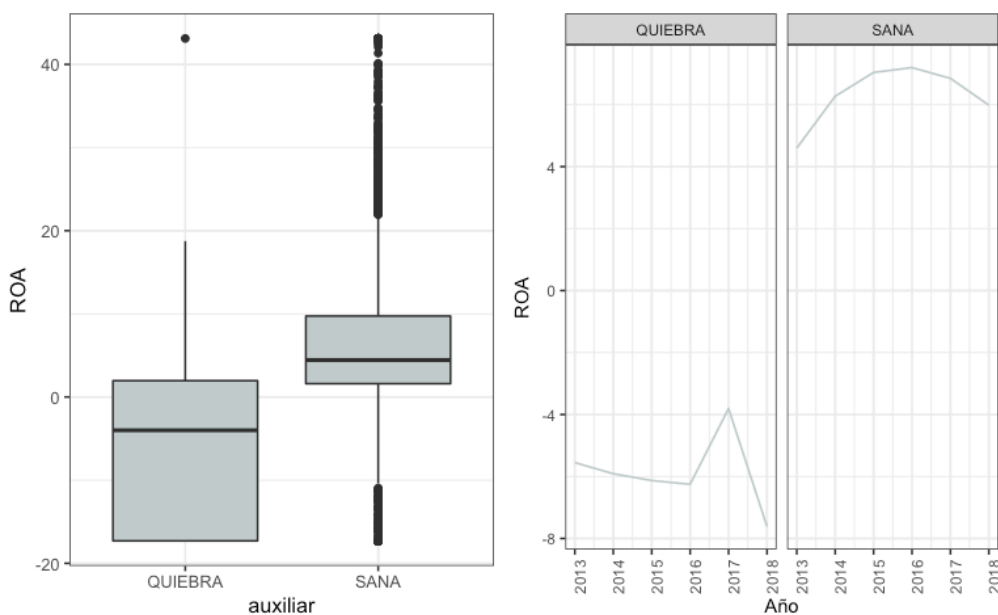


Gráfico 21: Evolución del ROA entre empresas en quiebra y sanas para el periodo 2013-2018
 Fuente: Elaboración propia a partir de datos de SABI

Completando la información aportada por el Gráfico 20 podemos observar que en lo referente a las empresas “sanas” dicho aumento en los ingresos se tradujo en una mejora en la rentabilidad; sin embargo, su estancamiento fue más acusado, llegando a un punto de inflexión en el año 2016.

En cuanto a las empresas en “Quiebra”, apreciamos que su caída se produce de forma abrupta, obteniendo en el año inmediatamente anterior al fracaso su valor más alto (- 4%). Esto puede ser debido a que otros factores actúan sobre el estado de la empresa de manera más temprana y dan como consecuencia inmediatamente anterior a la quiebra una bajada en la rentabilidad. En conclusión, la caída de esta ratio más que un hecho desencadenante de dicho fenómeno resulta una consecuencia de este.

3.3 Factores del fracaso empresarial: modelo AD

El objetivo del Análisis Discriminante es la creación de un modelo que permita la efectiva clasificación a priori entre empresas en Quiebra o Sanas para nuevos individuos externos a la muestra y la identificación de las variables clave a la hora de anticipar lo que en este trabajo se entiende por fracaso empresarial.

El análisis discriminante crea un criterio de clasificación a través de la obtención del llamado eje discriminante, el cual, se construirá haciendo uso de la base de datos que hemos extraído de SABI y del programa estadístico R-Studio. La aplicación del AD requiere validar en primer lugar que se cumplen los requisitos descritos en el marco teórico de este trabajo. Por tanto, dividimos este análisis en dos partes: primero se comprobarán dichos requisitos, y posteriormente analizaremos los resultados del análisis discriminante aplicado a nuestra muestra.

3.3.1 Hipótesis requeridas en el AD

Como se puede observar (Tabla 5) los máximos y mínimos de gran cantidad de variables presentan gran dispersión a la hora de compararse con su media, con lo que podríamos estar en presencia de los denominados datos atípicos. Esto puede ser debido a múltiples factores, entre ellos: errores de digitalización, empresas con rendimientos superiores a la media, hechos extraordinarios o incluso a la propia naturaleza de las ratios. Estas están compuestas por numerador y denominador por lo que si el numerador presenta valores cercanos a 0 la ratio se dispara, ocasionando valores atípicos o *outliers*.

	Min.	Q1	Median	Mean	Q3	Max.
current	0,00	1,20	1,70	2,50	2,74	68,35
ROA	-188,83	1,37	4,15	5,15	9,31	93,40
ROE	-8.144,22	1,43	7,25	5,12	18,53	2.812,04
quick	0,00	0,51	0,92	1,53	1,65	68,35
Mneto	765,38	0,39	1,80	1,95	4,88	898,99
rota	0,00	0,97	1,34	1,50	1,84	15,32
rotacli	0,00	0,05	0,17	0,20	0,30	6,80
Rotinv	0,00	1,60	2,70	24,02	5,12	26.139,94
Rotaprov	-0,01	2,24	3,40	5,25	5,49	194,98
Salario	0,00	22,49	28,39	30,52	36,80	303,20
Prodemp	0,00	125,55	196,20	304,80	355,90	13.072,80
Gtopersonal	0,00	9,14	14,40	15,80	20,70	207,83
Cobertint	-40.776,30	-37,30	-9,10	86,70	-0,20	416.511,80
Mbruto	-519,01	26,65	37,99	37,01	48,01	100,10
Endeudamiento	0,00	32,39	55,39	57,69	76,89	1.273,80

Tabla 5: Parámetros característicos de cada variable
Fuente: Elaboración propia a partir de datos de SABI

Estos valores pueden ocasionar interferencias a la hora de interpretar o crear el modelo, por lo que deben ser tratados con anterioridad. Después de una exploración pormenorizada, se llegó a la conclusión de que la eliminación²⁰ de *outliers* ocasionaba la completa exclusión de la muestra de las empresas medianas en quiebra. Alternativamente, se ha optado por tratar los *outliers* con la denominada técnica que recibe el nombre de su creador Charles P. Windsor como *winsorizing*.²¹

²⁰ Otra de las técnicas más conocidas es la eliminación de datos o también denominado *trimming*.

²¹ El *winsorizing*, permite substituir los *outliers* por sus correspondientes límites inferiores o superiores, clasificando los valores de menor a mayor y asignándole el valor límite equivalente al valor del cuartil establecido al resto de observaciones que lo sobrepasen.

En la gran mayoría se ha establecido un 2%, exceptuando algún caso como ROE y ROA, ya que, al tratarse de ratios sobre activos y capital propio, puede dar lugar a casos muy extremos cuando los valores del denominador tienden a cero. Podemos visualizar de manera más clara los datos en diagramas de cajas.

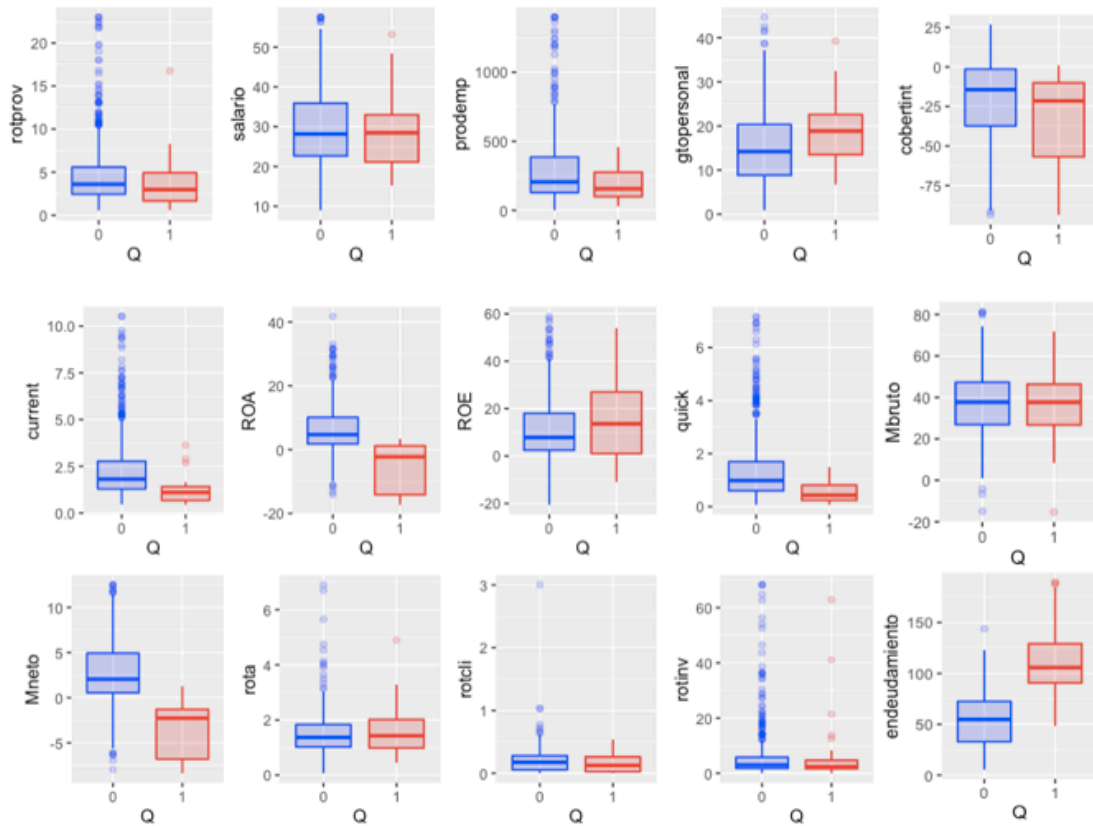


Gráfico 22: Diagramas de cajas por niveles de factor de cada variable tras eliminar valores atípicos
Fuente: Elaboración propia a partir de datos de SABI

A pesar de llevar a cabo el tratamiento de atípicos, nos encontramos con variables que continúan presentando una alta dispersión en las observaciones (Gráfico 22), pudiendo advertir: la rotación de inventario y de proveedores y las ratios current y quick.

Por otra parte, la correlación entre variables dificulta la medición del poder predictivo de manera individual, afectado en mayor medida a la hora de hacer inferencias a la población, más que a la obtención del modelo en sí. Desafortunadamente, es muy difícil que las variables no muestren algún grado de dependencia entre ellas, puesto que, al tratarse de información financiera, ciertas ratios como las de liquidez o rotaciones comparten componentes, creando así altas dependencias. Por lo tanto, se buscará que presenten la menor multicolinealidad posible, ya que, de no ser así, la inclusión de una variable que presente elevada correlación no aportaría un mayor poder explicativo al modelo.

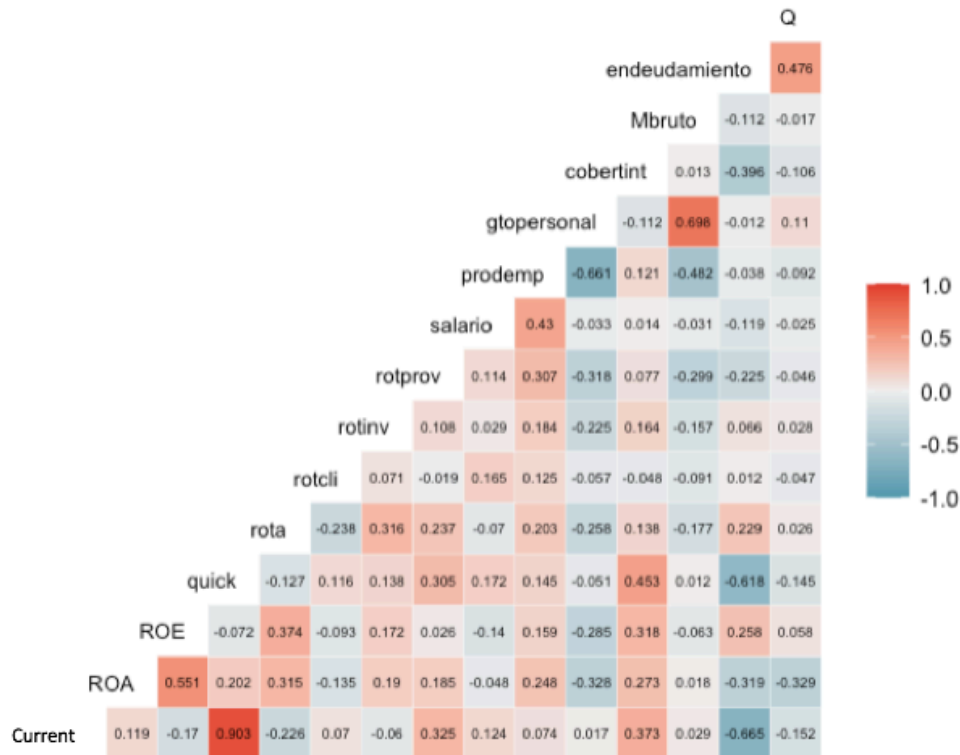


Gráfico 23: Matriz de correlaciones
Fuente: Elaboración propia a partir de datos de SABI

A la vista del Gráfico 23 existen altas correlaciones positivas entre la ratio quick (liquidez) y la ratio current (solventia) 0,903, seguido por la relación existente entre gtopersonal y Mbruto (0,698). La primera relación es debida a que ambas ratios se construyen con las mismas partidas contables de balance, en el denominador el PC, mientras que el en numerador el AC. Como ya se ha comentado de manera teórica, la ratio de liquidez no integraría en el AC las existencias, centrándose así en los elementos más líquidos de la empresa – y, por tanto, es razonable que presenten dicho resultado.

En cuanto a la siguiente relación, si bien es verdad que comparten los ingresos de explotación, los otros componentes difieren. El Margen bruto es construido antes de restarle gastos de personal, que si están reflejados en el Margen Neto. Con todo esto, podríamos estar ante una relación sin causalidad - es decir, una relación casual o propiciada por variables externas a estas ratios. Por último, la correlación negativa entre endeudamiento y current resulta fácilmente interpretable, ante disminuciones de liquidez existirían altas probabilidades de aumentos en el endeudamiento, o viceversa (no siendo posible aclarar con este análisis la relación causa-efecto).

Como el resto de los grados de correlación no resulta excesivo y la preocupación se encuentra principalmente en 3 relaciones, seguiremos con nuestro análisis y se decidirá sobre su eliminación una vez obtenidos los modelos.

Una de las propiedades deseables de los modelos discriminante es la existencia de la homocedasticidad. Para ello, la varianza del error debe ser constante o, en consecuencia, las matrices de varianzas-covarianzas correspondientes a cada grupo deben ser iguales entre si. Sin este hecho, el modelo perdería fiabilidad y eficiencia, por lo que se analizará mediante el test de *M Box*. En este contraste la hipótesis nula representa la igualdad entre las matrices de varianzas-covarianzas, mientras que H1 representa la no igualdad de estas. Atendiendo al el p-value y estableciéndose un nivel de significación del 5%, rechazamos Ho y aceptamos H1, lo que comprende aceptar que no existe homocedasticidad – véase Gráfico 36 del Anexo.²²

Test de Normalidad VARIABLES	p. value	
	0	1
ROA	3,380E-08	1,053E+03
ROE	8,690E-09	1,603E-01
Quick	5,869E-20	9,495E+02
Mneto	4,982E-01	8,722E+03
Rota	8,436E-15	1,414E+03
Rotcli	8,502E-24	8,522E-24
Rotinv	6,532E-29	2,013E-02
Rotprov	4,647E-21	1,556E+02
Salario	1,567E-04	5,872E+04
Prodemp	4,072E-20	3,582E+04
Gtopersonal	1,423E-03	1,373E+05
Cobertint	1,819E-08	6,792E+03
Mbruto	0.496	0.307
Endeudamiento.	7,035E+00	7,407E+03
Current	3,831E-20	3,944E-01

Tabla 6: Resultados para cada variable del contraste de normalidad
Fuente: Elaboración propia a partir de datos de SABI

A través del *Shapiro Test* se observará si las variables independientes se distribuyen como una normal. El contraste establece que la hipótesis nula representa la distribución normal de la variable mientras que la alternativa indica lo contrario. Con un nivel de significancia en 0,05 observamos (Tabla 6) que la gran mayoría de variables no presentan dicha distribución, exceptuando casos como el del Margen bruto.

La normalidad es muy recomendable en el análisis discriminante. Sin embargo, como ya se expuso en el marco teórico de este trabajo, el AD es robusto a la falta de normalidad de las variables explicativas si esta es causada por asimetría y no por

²² Si bien trabajamos con datos divididos por un periodo de tiempo – menos propensos a mostrar heterocedasticidad - no se tiene en cuenta su estructura temporal (mediante series temporales), tratándolos como observaciones independientes entre si (datos de corte transversal). Asimismo, la prueba de *M Box* es tachada de conservadora por múltiples estadísticos debido a su gran sensibilidad en tamaños muestrales grandes y a la no normalidad de las distribuciones de las variables, la cual resulta el tercer supuesto para el análisis discriminante.

atípicos. Siendo los *outliers* tratados en un primer momento, atenderemos que el modelo sigue siendo aceptable desde el punto de vista de la predicción. Para poder comprobar este hecho, se recurrió a los histogramas con los datos tipificados para una mejor visualización (Gráfico 24). Teniendo en cuenta que una variable normal debería mostrar una disposición de los datos en forma de campana y que la media, mediana y moda deben coincidir en el valor 0, se puede llegar a diferentes conclusiones a través de su visualización, entre ellas, la asimetría.

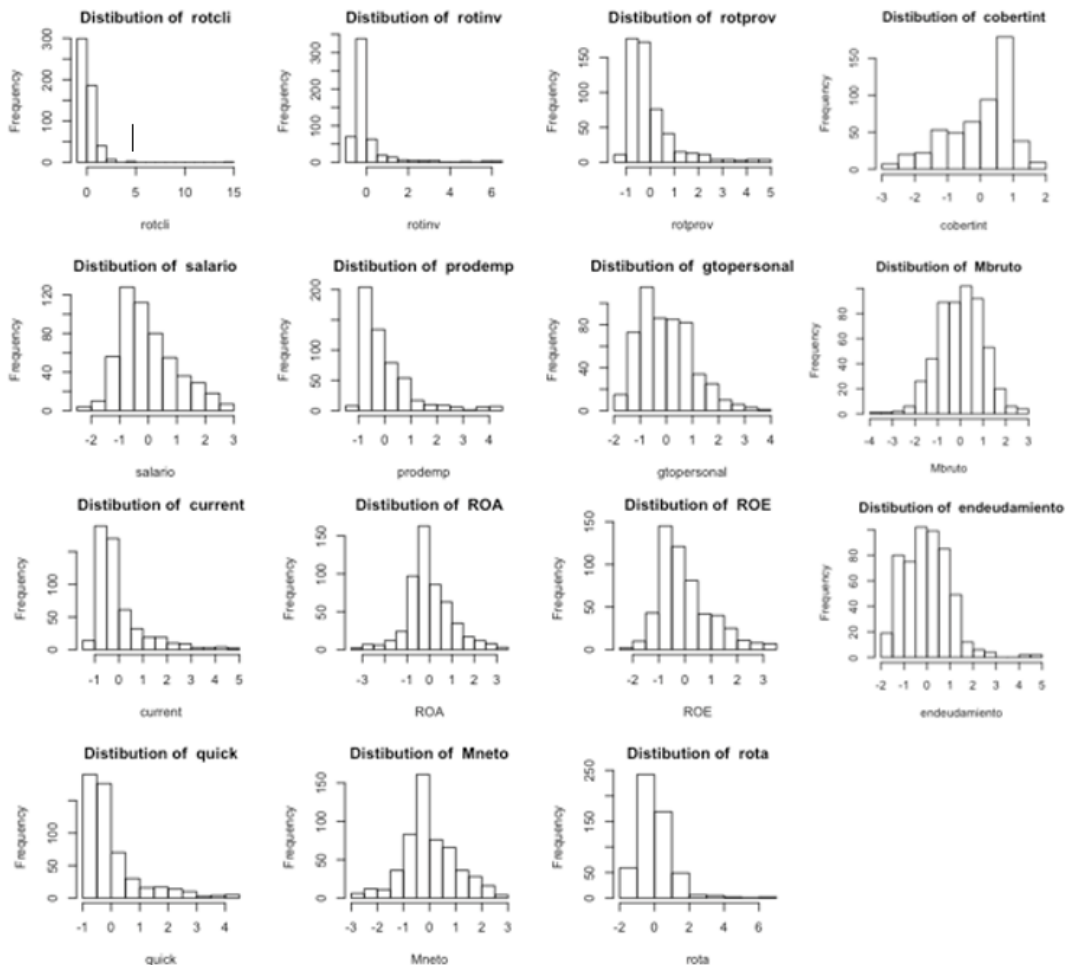


Gráfico 24: Histogramas de cada variable tipificada
Fuente: Elaboración propia a partir de datos de SABI

La denominada asimetría negativa, o sesgo a la derecha, se puede advertir en las ratios de productividad de los empleados, quick, rotaciones de proveedores, gastos en personal y salario. En estos casos la moda se encuentra en la barra más alta, mientras que la media se encontrará hacia la izquierda, y entre ambas se situará la mediana. Por su parte, la distribución asimétrica positiva o lo que es lo mismo, sesgo hacia la izquierda, se encuentra, únicamente, en la cobertura de intereses. Además, se puede apreciar simple vista que todas las distribuciones se presentan como leptocúrticas, con pesadas colas, especialmente visibles en las rotaciones. Esto indica un mayor número

de empresas con resultados cercanos a la media, pero al mismo tiempo, el porcentaje de empresas que nos se ajustan a esta media muestran resultados diversos, mas extremos, y alejados de lo que le correspondería a una distribución normal.

3.3.2 Resultados del análisis discriminante

Previo a la creación del algoritmo es la elección de las variables que lo van a componer. Para comprobar que indicadores ayudan a la hora de discernir si una empresa se dirige hacia un estado de fracaso, se llevará a cabo un análisis de varianzas. Primeramente, se constatará que la media de las poblaciones para cada categoría (fracaso o sana) es distinta. Este contraste establece que la hipótesis nula representa la igualdad de medias, mientras que la alternativa nos indica que existencia de diferencias entre ellas.

Además de verificarse la diferencia entre ambas poblaciones en el modelo general (Gráfico 37 del Anexo), también se comprobará para cada variable independiente (Tabla 7) con el objetivo de descartar ratios que no aporten poder discriminador al modelo.

Manova		
VARIABLES	PR(>F)	RESULTADO
ROA	5.578E-15	SIGNIFICATIVO
ROE	0.182	NO SIGNIFICATIVO
Quick	0.001	SIGNIFICATIVO
Mneto	2,211E-13	SIGNIFICATIVO
Rota	0.55	NO SIGNIFICATIVO
Rotcli	0.279	NO SIGNIFICATIVO
Rotinv	0.523	NO SIGNIFICATIVO
Rotprov	0.293	NO SIGNIFICATIVO
Salario	0.559	NO SIGNIFICATIVO
Prodemp	0.034	SIGNIFICATIVO
Gtoperonal	0.011	SIGNIFICATIVO
Cobertint	0.014	SIGNIFICATIVO
Mbruto	0.689	NO SIGNIFICATIVO
Endeudamiento.	2,267E-13	SIGNIFICATIVO
Current	0.000	SIGNIFICATIVO

*Tabla 7: Resultados del contraste Manova para cada variable
Fuente: Elaboración propia a partir de datos de SABI*

A la hora de hacer el contraste para todas las variables en conjunto podemos afirmar que este es significativo (Gráfico 37 del Anexo), y que el conjunto general de variables muestra diferencias significativas en sus medias; sin embargo, cuando se establece el foco en cada variable en particular no ocurre lo mismo (Tabla 7) y alrededor de la mitad de las variables que se han propuesto resultan no significativas.

Una de las diferencias en la elección de ratios con otros trabajos de la misma índole fue la introducción de las rotaciones por tratarse de medidas interesantes la hora de

analizar una empresa perteneciente al sector servicios. A pesar de ello, resultaron no ser de gran relevancia a la hora de distinguir el fracaso empresarial según el MANOVA.

Otra serie de variables que tampoco ayudarían en la diferenciación entre empresas en quiebra o sanas serían el salario proporcionado a los empleados, el margen bruto y ROE. Este último indicador resulta muy atractivo para los accionistas de una compañía, ya que les permite medir los beneficios que le reporta ser propietario de una parte de los fondos propios de la empresa; no obstante, la muestra pertinente contiene empresas pequeñas y medianas que no cotizan en bolsa. Se constata que las empresas utilizan una media del 30% de sus recursos en patrimonio neto, frente al 70% en pasivos – mayormente préstamos o créditos de proveedores – y no suelen requerir de grandes inversiones iniciales. Por lo tanto, la atención en empresas de este sector se concentrará en el ROA, la rentabilidad que obtiene con sus activos, que puede implicar elementos como las marcas, establecimientos o incluso gestión de inventarios.

Una vez descartadas, se procedió a la creación del modelo a través de un criterio *forward* de selección de variables, en base al poder de separación entre grupos. Como resultado se obtuvieron 3 variables significativas, dispuestas por orden de aportación a la hora de separar los grupos:

- Endeudamiento: 0,9453
- Endeudamiento + current: 0,9567
- Endeudamiento + current + ROA: 0,9652

Este último valor representa el poder separador que alcanza el modelo conformado por las 3 variables. Se puede reparar que el endeudamiento es el criterio que mayor porcentaje aporta a la hora de distinguir un grupo de otro, mientras que la ratio current y ROA mejoran esta ligeramente. De forma visual apreciamos la separación conseguida con las 3 variables a través de un gráfico de dispersión (Gráfico 25).

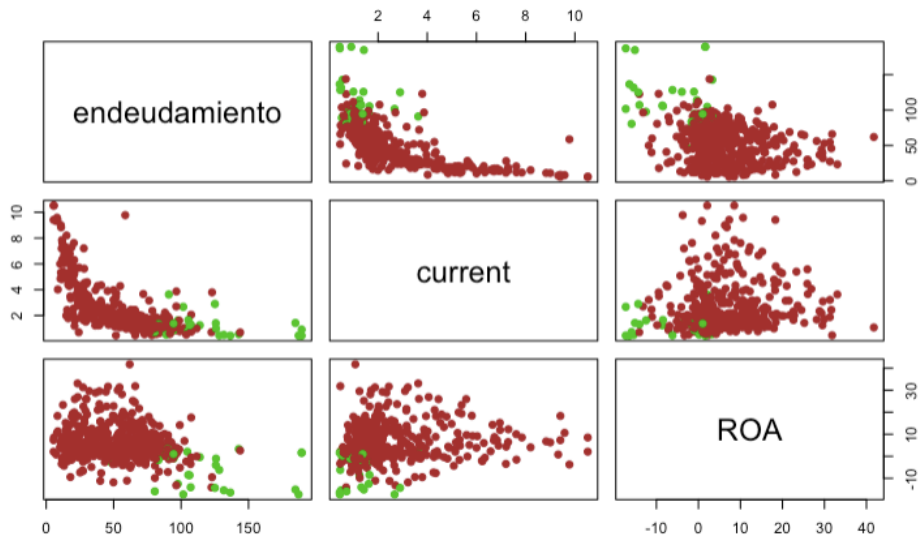


Gráfico 25: Dispersión entre las variables endeudamiento, current y ROA
Fuente: Elaboración propia a partir de datos de SABI

```
Call:
lda(cualit ~ endeudamiento + current + ROA, data = datos)

Prior probabilities of groups:
      0      1
0.94766355 0.05233645

Group means:
  endeudamiento  current      ROA
0      53.72554  2.415636  6.109780
1     116.11945  1.221714 -5.628739

Coefficients of linear discriminants:
              LD1
endeudamiento  0.04380729
current        0.32705629
ROA           -0.05355496
```

Gráfico 26: Resultados del modelo discriminante
Fuente: Elaboración propia a partir de datos de SABI

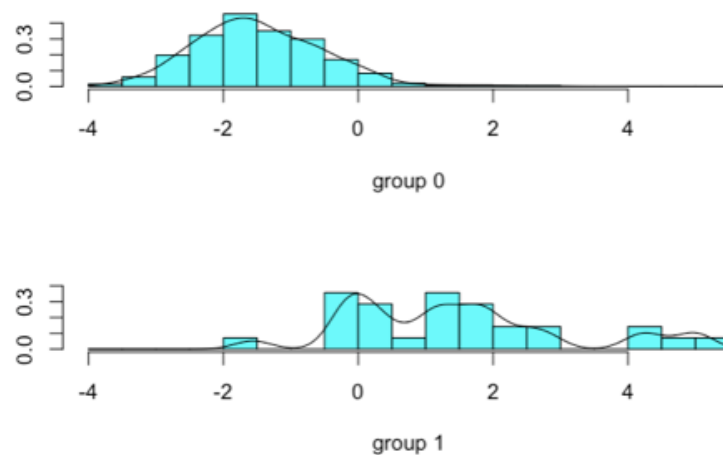
Lo primero que podemos advertir en el Gráfico 26 es *prior probabilities of groups*, que muestra la probabilidad de que un individuo pertenezca a un grupo de acuerdo con las proporciones existentes en la base de datos. Así, existe un 94,7% de probabilidad de que una empresa en nuestra muestra se encuentre en una buena situación, mientras que, como complementario, existe un 5,3% de que se encuentre en quiebra.

El segundo indicador, *Group means*, muestra la media de cada variable del modelo. Esto nos permite observar que el endeudamiento promedio de las empresas en quiebra es más alto, mientras que el de los ratios current y ROA resulta más bajo – resultados poco sorprendentes y acordes con la teoría económica. Sin embargo, el signo negativo de la ratio current es difícil de interpretar, ya que indicaría que una empresa tendría más

probabilidades de llegar al denominado fracaso económico cuanto mas alto fuese la ratio de liquidez. Entendemos que es posible que la presencia de empresas en fase de liquidación esté distorsionando este resultado. En todo caso, en el gráfico de dispersión podemos ver que la ratio current es efectivo a la hora de distinguir empresas en quiebra únicamente cuando se encuentra en combinación del ROA o endeudamiento, sin embargo, por si solo, muchas empresas sanas se encuentran en el mismo nivel en el que la ratio establece que una empresa esta en quiebra. Vista su correlación con la función del modelo (Anexo-Gráfico 39), se tomó la decisión de excluirlo de la ecuación por su falta de significado económico, su baja aportación al modelo y su alta correlación negativa con el endeudamiento.

Por tanto, la ecuación final del modelo se mostraría de la siguiente forma:

$$Y = 0,0438 \times Endeudamiento - 0,0535 \times ROA$$



qualit <fctr>	frecuencia <int>	centroidesLD1 <dbl>
0	507	-0.153842
1	28	2.785639

Gráfico 27: Valor de los centroides por grupo
Fuente: Elaboración propia a partir de datos de SABI

Atendiendo a los centroides (Gráfico 27) vemos que su punto de corte es 1,21 y que el centroide de las empresas en quiebra se encuentra en la zona positiva (2,78) lo cual denota que el grupo de estas empresas tiende a obtener ponderaciones positivas en la ecuación discriminante. A mayor endeudamiento, mayores serán las probabilidades de que la empresa se encuentre en quiebra, mientras que a mayor Rentabilidad en los Activos menores serán las probabilidades de que esta empresa se encuentre en una situación de fracaso, ya que el signo negativo que precede a la variable ROA hace que los movimientos en esta actúen de forma inversa al endeudamiento.

Por último, comprobamos la eficacia del modelo utilizando los datos del año 2018 que se habían reservado en un principio (*test set*). La *performance* del modelo se expondrá través de la *confusión matrix*, la cual está distribuida de la siguiente forma:

		Confusion Matrix	
		Predict group	
		0	1
Actual group	0	504	3 (error tipo I)
	1	4 (error tipo II)	24
Accuracy		98,69%	
Precision		88,89%	
Sensitivity		85,71%	
Specifity		99,41%	

Gráfico 28: Matriz de confusión para el modelo discriminante
Fuente: Elaboración propia a partir de datos de SABI

La matriz de confusión (Gráfico 28) está compuesta por “*Actual group*” y las dos categorías 0 (Sana) y 1 (Quiebra) de forma vertical y los valores predichos divididos también ambas categorías de forma horizontal. Atendiendo a esta disposición, el nivel de desempeño del algoritmo vendrá dado por la diagonal principal, es decir los puntos (0,0) y (1,1), siendo estos los puntos de acierto del modelo. Por contraste, las coordenadas (1,0) y (0,1) serán las empresas que han sido clasificadas erróneamente por el algoritmo. La primera fila muestra que 504 empresas sanas han sido clasificadas de forma correcta según el modelo, mientras que 3 empresas que se encuentran sanas han sido clasificadas como en quiebra. La siguiente fila muestra la eficacia a la hora de clasificar las empresas en quiebra: 4 empresas que se encuentran en quiebra dentro del *training set* fueron predichas como sanas, mientras que 24 han sido correctamente clasificadas por el modelo como en quiebra.

La matriz de confusión permite comprobar el desempeño del modelo a través de una serie de medidas centradas en distintos aspectos de su cualidad como clasificador. Entre ellas se encuentra la exactitud (*Accuracy*), la cual mide el porcentaje casos que el modelo ha clasificado correctamente. Comprobamos que se obtiene una elevada exactitud, de un 98,69%, pero hay que tener en cuenta la disparidad de proporción de integrantes en cada grupo afecta a la hora de analizar dicho porcentaje. En el caso de que el presente modelo clasificase a todas las empresas como sanas, incluyéndose las 28 en quiebra, seguiría obteniendo un alto porcentaje de *Accuracy*. En estadística este fenómeno se trataría como una reducción del error tipo I a costa de aumentar el error tipo II, tratándose a ambos errores con la misma consideración. Por ello, esta medida a

pesar de aportar datos interesantes debe ser analizada junto a otras tales como la precisión (*Precision*) y exhaustividad (*Sensitivity o recall*).

La precisión hace referencia a la calidad del modelo (Error tipo I), estando formada por el número de verdaderos positivos (24), dividido entre el número total de empresas clasificadas en estado de quiebra por el modelo (24+3), alcanzando un 88,89%. Por otra parte, la exhaustividad (Error tipo II) muestra la cantidad de empresas que, dirigiéndose a un estado de quiebra (4+24), el modelo fue capaz de identificar (24), alcanzando un 85,71%. En otras palabras, ambas medidas responden a uno de los objetivos principales del modelo, que es el porcentaje que alcanza este a la hora de predecir el fenómeno Quiebra, por lo que se analizarán de forma conjunta a la hora de compararlas con el modelo logit.

3.4 Factores del fracaso empresarial: modelo logit

Haciendo uso de la misma base de datos del epígrafe anterior, obviamos ya el análisis de los *outliers* y las gráficas descriptivas del *training set* – ya que resultaría redundante – realizamos la selección de las variables predictoras del estado de una empresa en base a la metodología de regresión logit. Para ello, llevamos a cabo un método distinto al forward por resultar este menos eficiente para los modelos logit. En lugar de ello se utilizó la función *bestglm* que, a diferencia de los métodos *stepwise*, *forward* y *backward*, limitados a añadir o eliminar variables a cada paso, este analiza todas las combinaciones posibles entre variables y escoge la mejor de acuerdo con un criterio preestablecido, que en este caso será el modelo que menor AIC²³ presente.

²³ El Aikake Information Criterion es una herramienta estadística que permite elegir entre varios modelos teniendo en cuenta la complejidad, dada por el número de variables, y la calidad del ajuste, medida a través del criterio de máxima verosimilitud, con la ecuación: $AIC=2K-2\ln(L)$.

```

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.43866 -0.17725 -0.07395 -0.02428  2.44558

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -8.34479    1.24280  -6.715 1.89e-11 ***
endeudamiento  0.06520    0.01338   4.872 1.10e-06 ***
Mneto        -0.41182    0.10106  -4.075 4.60e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 219.712 on 534 degrees of freedom
Residual deviance: 91.642 on 532 degrees of freedom
AIC: 97.642
    
```

Gráfico 29: Resultados modelo logit utilizando la función best subsets
Fuente: Elaboración propia a partir de datos de SABI

A vista de los resultados mostrados, la única variable que se presenta en ambos modelos es el endeudamiento. La razón de la aparición de diferentes ratios puede ser debida a la decisión de llevar a cabo diferentes métodos de selección de variables

El análisis logit goza de algunas ventajas sobre el AD en lo referente a las hipótesis a satisfacer para validar el desempeño del modelo, como la no necesidad de que las variables sigan una distribución normal y la igualdad de matriz de varianza-covarianza. Sin embargo, su principal hándicap es la interpretabilidad: mientras que en el AD los coeficientes se podían explicar directamente como incrementos o decrementos fijos de acuerdo con el aumento o disminución de 1 unidad en nuestras variables, en el análisis logit no ocurre así. En estos modelos solo serán interpretables los signos.

De acuerdo con el Gráfico 29, un incremento del endeudamiento actuaría como factor que intensificador de la probabilidad de que una empresa se encuentre en una situación de fracaso, mientras que, el Margen Neto funcionarían en sentido inverso. Debido a la poca información aportada por estos coeficientes, se recurre al concepto de los denominados *odd ratios*, los cuales se han obtenido aplicando el logaritmo neperiano a los coeficientes del Gráfico 29. Ya tratados en el apartado modelo del logit, un *odd ratio* no es más que la razón de ocurrencia de un suceso respecto a su no ocurrencia.

Intercepto	Endeudamiento	Margen Neto
0,0035	6,7201	0,19982

Gráfico 30: Valores de los Odd ratio para cada variable
Fuente: Elaboración propia a partir de datos de SABI

A la vista de los datos mostrados (Gráfico 30), a medida que aumenta en ratio de endeudamiento en una unidad la empresa tendrá 6,72 veces mas probabilidad de

encontrarse en quiebra. Mientras que, las otras dos variables, Mneto, al ser menor a uno, se realizará su interpretación obteniendo su inversa, siendo esta 5. Esto quiere decir que, ante aumentos unitarios del margen neto, existirán 5 veces menos probabilidades de que la empresa esté en situación de quiebra.

		Confusion Matrix	
		Predict group	
		0	1
Actual group	0	504	3 (error tipo I)
	1	7 (error tipo II)	21
Accuracy		98,13%	
Precision		87,50%	
Sensitivity		75,00%	
Specifity		99,41%	

Gráfico 31: Matriz de confusión para el modelo logit
Fuente: Elaboración propia a partir de datos de SABI

Atendiendo a la *Confusion Matrix*, podemos denotar una *Accuracy* casi idéntica (98,13%) a la proporcionada por el modelo discriminante. Esta semejanza viene propiciada por la baja proporción de empresas en quiebra frente a sanas, conllevando que errores en el número de clasificación de empresas en quiebra se vea reflejado de forma ínfima en la exactitud del modelo en conjunto.

Sin embargo, si se pone el foco en las siguientes dos medidas, *Precision* y *Sensitivity*, podemos analizar las diferencias respecto a nuestro anterior modelo. Así, la precisión cae ligeramente, en un 1,39%, respecto al AD. La exhaustividad también disminuye, pero esta en mayor medida, en un 10,71%. Este modelo presenta un mayor error de tipo I y de tipo II que alcanzado por el AD. Sin embargo, el análisis logit mejora el nivel de un error en detrimento del otro. Este intercambio entre errores se podría observar a través de la curva ROC, (Gráfico 31) representado en el eje vertical *Sensitivity* y en el eje horizontal la ratio de falsos positivos.

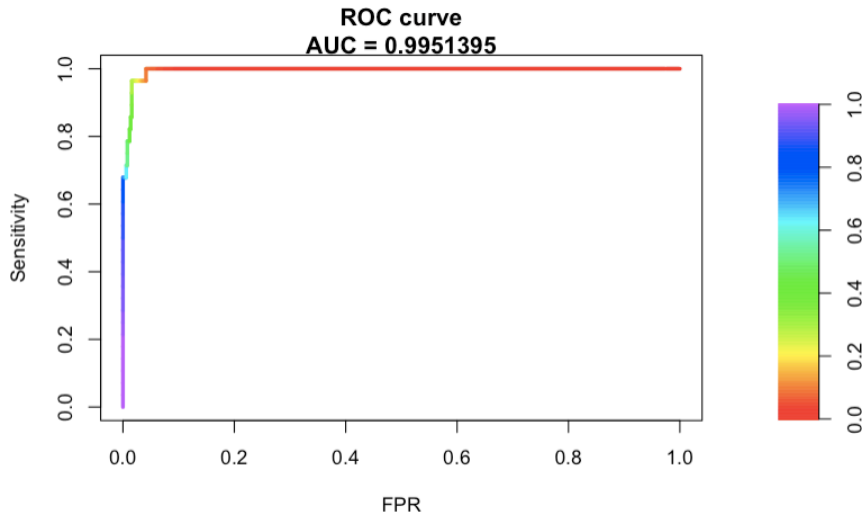


Gráfico 32: Curva Roc

Fuente: Elaboración propia a partir de datos de SABI

La curva ROC también es una buena medida de rendimiento del modelo, pudiendo observarse este a través del área que recae debajo de la curva denominado AUC. Esta nos proporciona el rendimiento agregado de los diferentes umbrales que se podría dar, estando comprendida entre 1, poder discriminatorio perfecto y 0,5, haciendo referencia a la pura aleatoriedad del algoritmo a la hora de clasificar.

A vista de los resultados obtenidos podemos concluir que nuestro modelo es fuertemente satisfactorio, comprendiendo el área AUC un 0,995, muy cerca al poder discriminatorio perfecto. Este transvase de aciertos entre los dos tipos de errores está estrechamente relacionado con el concepto *cut off* óptimo establecido. Como sabemos, la regresión logística clasifica en una categoría u otra de forma binaria a través de probabilidad de pertenencia a un grupo que nos proporciona la función logit; no obstante, recae en manos del investigador decidir cual es el punto a partir del cual una empresa pertenece a un grupo u otro. En un primer momento se optó por el 50%, otorgando así la misma importancia a ambos errores.

Por otra parte, resulta de gran interés analizar como evoluciona nuestro modelo al penalizar el hecho de clasificar una empresa en quiebra como sana. Por lo que a continuación, se muestra el coste de maximizar la exhaustividad en detrimento de la precisión, dando una importancia 5 veces mayor al hecho de presentar falsos negativos.

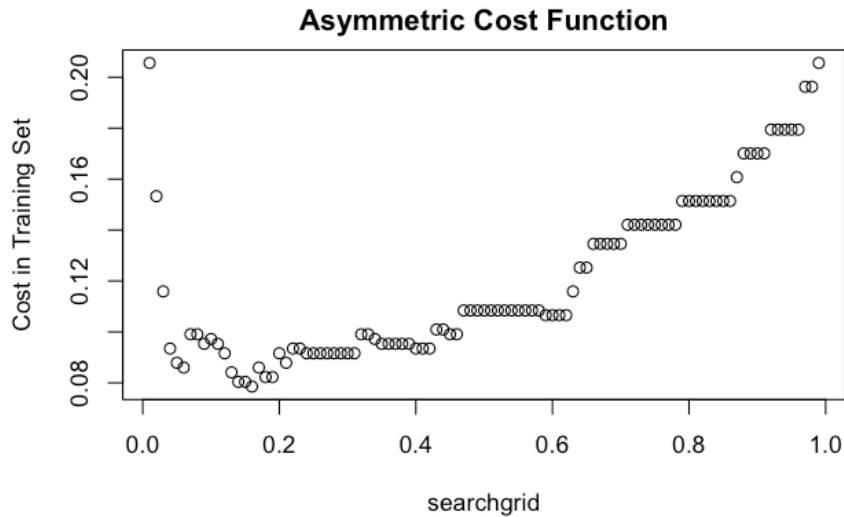


Gráfico 33: Función de coste asimétrica

Para establecer el punto de corte óptimo en la función antes vista, tendríamos que localizar el punto donde a función alcanza un mínimo, el cual se posiciona en el 0,11. Utilizando dicho cut off, la *confusion matrix* resultante sería la siguiente (Gráfico 34).

		Confusion Matrix	
		Predict group	
		0	1
Actual group	0	486	21 (error tipo I)
	1	1 (error tipo II)	27
Accuracy		95,89%	
Precision		56,25%	
Sensitivity		96,43%	
Specifity		95,86%	

Gráfico 34: Matriz de confusión con un punto de corte de 0,16
Fuente: Elaboración propia a partir de datos de SABI

3.5 Análisis de los resultados obtenidos

Este último apartado sintetiza el análisis empírico realizado, en dos aspectos. Primero, realizamos un breve resumen y comparación de los pasos analíticos aplicados en cada modelo. Y segundo, comentamos los resultados más significativos para nuestro sector de estudio, en base a ambos modelos.

Primeramente, para evitar potenciales problemas a la hora de formas los modelos fue necesario tratar la gran multitud de *outliers* en las ratios. De entre la variedad de técnicas y procedimientos existentes en este campo se seleccionó el denominado *wisoring*. Se comprobaron los supuestos previos que debe satisfacer el AD, referidas a la homocedasticidad (varianza intragrupo constante), multicolinealidad (correlaciones

entre variables), y normalidad (las ratios se distribuyen de forma normal). Los resultados mostrados no han desvelado nada nuevo respecto a otros trabajos, como el de Jardín P. (2015), exponiendo la poca consistencia de estas hipótesis en la realidad. Siendo el problema de los atípicos tratado en un primer momento y verificada la simetría de las distribuciones a través de histogramas, entendemos que el modelo sigue siendo válido para los objetivos del estudio. Además, la otra alternativa propuesta – el análisis logit – resulta mucho más flexible a estas rigideces estadísticas.

En cuanto a las variables que compondrán los modelos fue necesario hacer un MANOVA y un análisis de correlaciones para poder descartar ratios que no fueran realmente útiles a la hora de discernir entre empresas “Sanas” o en “Quiebra”. Los resultados permitieron reducir las posibles candidatas a la mitad y se ha constatado la poca utilidad de las rotaciones y las variables relativas a empleados. Por último, se emplearon distintos métodos para la formación del algoritmo, por ser considerados más apropiadas para cada técnica. Así, en el análisis discriminante el criterio resultaba de seleccionar las variables con mayor poder separador, en el logit se trataba del modelo con menor AIK.

En un primer momento, el modelo discriminante quedó conformado por las variables ROA, endeudamiento y current; sin embargo, esta última ratio fue eliminada por diversas razones, entre ellas, la falta de lógica económica que presentaba su signo negativo, la baja correlación con la función y la ya mencionada elevada correlación negativa con el endeudamiento. Mientras tanto, el modelo logit quedaba conformado por las ratios de endeudamiento y margen neto.

Sin duda, la variable que más explica el fracaso empresarial en ambos modelos para el sector textil es el endeudamiento, con un poder del 94,53% a la hora de separar los grupos y el mayor valor de los *odd ratios* en el modelo logit. En concreto, la probabilidad de quiebra aumenta en 6,72 veces por unidad de endeudamiento. El hecho de que el endeudamiento sea el factor decisivo en las quiebras contrasta con trabajos como el llevado a cabo por Keener (2013), donde todas las variables integrantes de su modelo resultan significativas (inferiores al 5%) excepto la ratio de endeudamiento. Además, Keener (2013) verifica que el hecho de tener un menor número de empleados incrementa las probabilidades de quiebra, mientras tanto, en la presente investigación, ninguna ratio referida a empleados se presentó en el modelo. Las causas de que las demás variables difieran se pueden deber al uso de diferentes criterios de elección de las variables que conforman el modelo o su bajo poder discriminador. El ROA y margen neto ayudan obtener una mejor predicción, pero, por sí solas, no tendrían el suficiente poder predictivo a la hora de discernir entre ambos grupos.

Finalmente, se procedió comprobar el pronóstico de los modelos a través de la *confusión matrix*. Ambos presentaron alto porcentajes de acierto; a pesar de ello, debido su principal finalidad, el criterio de mayor peso será el error de tipo II, ya que, a la hora de predecir, clasificar una empresa sana como en quiebra podría traer algún que otro coste por investigaciones o sobreprotección innecesaria; en cambio, clasificar una empresa que se dirige hacia una quiebra inminente, como sana, podría traer duras consecuencias al no intervenir de forma anticipada en el rumbo de su estrategia.

Teniendo en cuenta lo anteriormente expresado, tendremos que establecer un mayor peso al porcentaje aportado por la *Sensitivity* o también denominado *recall*, por lo tanto, partimos de un 85,71% (AD) frente a un 75% (Logit). Por otra parte, en el análisis logit se pueden penalizar los errores a través del *cut off*. Estableciéndose este en un 0,11, estaríamos mejorando la exhaustividad del modelo en un 21,43% (96,43%), pero no sin perjuicio de la precisión, la cual desciende un 31,25% (56,25%).

En definitiva, si bien es verdad ante la falta del cumplimiento de las hipótesis del modelo discriminante, este presenta resultados eficientes en su pronóstico, además de que resulta el más equilibrado en ambos errores, mostrando una efectividad de detección de empresas en quiebra de un 85,71% para el año 2018. Por otra parte, aunque dentro del modelo logit se puede modificar el *cut off* a nuestro antojo, la caída en la precisión no compensa lo suficiente la escasa mejora en la exhaustividad. Con todo, podemos afirmar que el AD muestra unos resultados más eficientes que el logit para PYMES pertenecientes al sector del comercio textil.

Conclusiones

La industria de comercio textil en España atraviesa una época de gran dinamismo, propiciado por la demanda de los clientes, cada vez más exigentes e informados gracias a las nuevas tecnologías. La competencia es creciente, debido a la concentración de la demanda en unas pocas empresas líderes en el sector. Estas, presentan estrategias que no son adaptables al resto de actores del mercado, como las PYMES. La apertura de más y mayores establecimientos en mejores localizaciones o un fuerte poder de negociación sobre proveedores que hace abaratar los costes, resultan algunas de las fórmulas mencionadas.

En este contexto, este trabajo planteó la necesidad de un modelo de predicción que permita anticiparse a un inminente fracaso empresarial, y al mismo tiempo, sirva como base para posteriores investigaciones en la búsqueda de indicadores de quiebra para modelos centrados en el sector *retail*. Ante la gran cantidad de técnicas y trabajos referenciados dentro del campo de los modelos de predicción de insolvencia, nos hemos decantado por el contraste de dos técnicas, de fácil comprensión, pero de gran efectividad y popularidad: el análisis discriminante y la regresión logit.

El estudio emplea una muestra de 535 empresas perteneciente al sector del comercio de prendas de vestir extraídas de la base de datos SABI. A través de un primer acercamiento a la muestra se pudo comprobar que las empresas en quiebra obtenían una rentabilidad sobre activos (ROA) mucho menor que las empresas sanas. Teniendo en cuenta este hecho y habiendo una clara relación inversa entre ROA y fracaso empresarial, se procedió a la realización de un análisis descriptivo en el año 2018 que nos permitiese indagar acerca de las causas que ocasionan que unas empresas obtengan un mejor aprovechamiento de sus activos que otras. Se concluye que: las empresas con mayor margen neto presentaban un menor número de trabajadores de media que la muestra en su conjunto; las empresas medianas se aprovechaban mejor del fenómeno de aplacamiento a través de los créditos a proveedores; una ligera recuperación a lo largo del periodo (2014-2018) respecto a los ingresos acompañada de una mejora en la rentabilidad y, por último, la caída del ROA como consecuencia inmediatamente anterior al fracaso empresarial.

Una vez caracterizada la muestra pasamos al eje principal de trabajo, el análisis empírico, con la aplicación de dos técnicas de estadística multivariante: análisis discriminante y regresión logit. Escogemos el primero por ser una de las técnicas tradicionales en este ámbito. Sin embargo, uno de los problemas del AD son las rígidas asunciones estadísticas que este poseía, como la necesidad de que las variables se distribuyan de acuerdo con una normal o la homocedasticidad de los datos. Por ello, se decidió contrastar los resultados con otra técnica que presentase menos premisas, como el análisis logit.

Los resultados del AD para el criterio de mayor poder de separación entre grupos (“stepwise”) señalan que las variables endeudamiento y ROA son las que mejor anticipan una eventual situación de insolvencia. En el caso de la regresión logit – suando la función “bestglm” por ser esta más adecuada para modelos lineales generalizados – se obtiene evidencia en favor de los ratios endeudamiento, margen neto y current. No obstante, esta última fue eliminado por diversos motivos, como una baja correlación con la función, una alta correlación con otra de las variables clave (ratio de endeudamiento), y un impacto sobre la solvencia empresarial de difícil interpretación económica.

A diferencia de otros trabajos centrados en el sector *retail*, la variable por excelencia resultó ser, por tanto, el endeudamiento, seguido por otras como Margen Neto y ROA. Estas últimas ayudan obtener una mejor predicción, pero, por sí solas, no tendrían el suficiente poder predictivo a la hora de discernir entre ambos grupos. El endeudamiento se presentaba como la variable clave de las ecuaciones, mostrando un gran poder separador en las gráficas de dispersión del AD y el valor más elevado de los odd ratios en el logit.

Finalmente, con el fin de averiguar si los modelos presentan resultados realmente eficientes, se comprobó la *performace* de ambos a través de la *confusion matrix* y una serie de medidas derivadas de esta: *Accuracy*, *Precision*, *Sensitivity* y *Specificity*. La *Accuracy* muestra niveles similares en ambas técnicas, cercanos al 98% de aciertos. No obstante, debido a la cercanía de resultados y a los objetivos del trabajo se hizo hincapié en las métricas de precisión y exhaustividad. El AD ha presentado porcentajes 88,89% y un 85,71% respectivamente, mientras que el modelo logit mostró unos resultados ligeramente inferiores de un 84,62 y 78,57%. Para poder decantarnos por un modelo u otro, examinamos el modelo logit a través de cambios en su *cut off* y considerando el error de tipo II (exhaustividad) el más grave. Los resultados alcanzados a través de cambiar el *cut off* alcanzaron un 96,43% de exhaustividad, sin embargo, la abrupta caída que se produjo en la precisión (60%) no compensó esta ligera mejora.

Analizando los resultados, nos preguntamos si la definición de insolvencia escogido fue la correcta o, por otra parte, se deberían haber tomado otras consideraciones al definir a una empresa en e estado de “fracaso empresarial”. También nos planteamos si las altas tasas de predicción se podrían conservar para horizontes temporales más largos o, por el contrario, caerían abruptamente. En todo caso, debemos ser conscientes de que este trabajo cuenta con ciertas limitaciones. Entre las principales, señalamos el tratamiento simplista de los datos no disponibles (n.d.), una definición de fracaso empresarial no generalmente aceptada, y un modelo únicamente nutrido de ratios contables. Pese a ello, el trabajo ha contribuido a la extensión de este tipo de técnicas a un sector tal como el comercio, mostrando que pueden alcanzarse resultados de predicción elevados y suscitando interés por los actores más vulnerables del mercado, las PYMES.

En definitiva, este estudio contribuyó a reforzar conceptos ya estudiados en el máster y ampliar conocimientos en el paquete estadístico Rstudio. Asimismo, nos ayudó a evidenciar la importancia de las técnicas multivariantes aplicadas al análisis de insolvencia y su potencial extensión a otros campos del mundo empresarial.

Bibliografía

- Almamy, J., Aston, J., & Ngwa, L. (2015). An evaluation of Altman's Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: Evidence from the UK. *Corporate Finance*, 278-285.
- Altman, E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*.
- Altman, E. (1981). *Corporate Financial Distress and Bankruptcy*. John Wiley & Sons, Inc.
- Arribas Barreras, V., Josa, E., Bravo Durán, S., García Hiljding, I., & San Miguel Arregui, P. (2016). *El Sector de la Moda en España: Retos y desafíos*.
- Aylmer Fisher, R. (1963). *The Architect of Multivariate Analysis*.
- Asociación Empresarial de Comercio Textil, Complementos y Piel. (2018). *El Comercio Textil en Cifras*.
- Banco de España. (2012). *El uso de los concursos de acreedores en España*.
- Beaver, W. (1966). Financial ratios as predictors of failure. Empirical Research in Accounting: Selected Studies. *Journal of Accounting Research*.
- Bolaños, C. (1997). *El fracaso empresarial y el Contador Público y Auditor como Liquidador de empresas*. Guatemala: Universidad Francisco Marroquín, Facultad de Ciencias Económicas.
- Ben Jabeur, S. (2017). Bankruptcy prediction using Partial Least Squares Logistic Regression. *Journal of Retailing and Consumer Services*, 197-202.
- Bhandari, S. (2013). Predicting business failure using cash flow statement based measures. *Managerial Finance*, 667-676.
- Castaño, R. G. (2000). Creación de PYMES: Objetivo emprendedor .
- Cea D'Ancona, M. Á. (2016). *Análisis discriminante*. Madrid: CIS.
- CESCE. (2019). Informe Sectorial CESCE 2019 Textil . Madrid.
- COLL-SERRANO, V. B.-B. (2011). ANÁLISIS COMPARATIVO DE LA EFICIENCIA DE LA PYME TEXTIL ESPAÑOLA: UNA APROXIMACIÓN AL EFECTO DE LA LIBERALIZACIÓN EN EL PERIODO 2004-2006. *Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA*, 33-52.
- Conferación Española de Organizaciones Empresariales. (2019). *El sector comercio en la economía española*. Madrid.
- Correa Rodríguez, A., Acosta Molina, M., & Lorenzo González Pérez, A. (2003). La insolvencia empresarial: un análisis empírico para la pequeña y mediana empresa. *Revista de Contabilidad*, 47-79.
- Costa Néstor Duch, T. M. (2004). La renovación del sector textil-confección en España. *Revista Economía Industrial*, 263-272.
- Cuervo, A. R. (1986). El análisis económico-financiero de la empresa. *Revista española de financiación y contabilidad*, 15-33.
- Du Jardin, P. (2015). Bankruptcy prediction using terminal failure processes. *European Journal of Operational Research*, 286-303.

- EL COMERCIO JUSTO EN ESPAÑA. (2014). *El mercado global del textil y sus desequilibrios comerciales*.
- Espinosa, F. R. (2013). Alcances y limitaciones de los modelos de capacidad predictiva en el análisis del fracaso empresarial. *Red de Revistas Científicas de América Latina, el Caribe, España y Portugal*, 45-70.
- Fernández, A. I. (1986). El diagnóstico financiero de la empresa: Nuevas tendencias en el análisis. *Revista española de financiación y contabilidad*, 113-132.
- Francisco, V. C. (2008). TRES AÑOS DE LEY CONCURSAL: TEMAS DE REFORMA. *Revista de Derecho Concursal y Paraconcursal*, 113.
- García-Ayuso, M. y Jiménez, S.M. (1996): Una reflexión crítica sobre el concepto y el ámbito del análisis financiero y los objetivos de la investigación en material de análisis de la información financiera. *Revista Española de Financiación y Contabilidad*.
- Gitman, L. (1997). *Fundamentos de administración financiera*. México: Editorial OUP.
- Gitman, L., & Zutter, C. (2012). *Principios de administración financiera*. México: Pearson.
- Goñi Arbide A. (2015). *Análisis de la distribución comercial del sector textil*. (Tesis de grado). Universidad del País Vasco, Bilbao.
- Graveline, J. y Kokalari, M. (2008). Credit risk. Working Paper, The Research Foundation of CFA Institute, November.
- He, Y., & Kamath, R. (2006). BUSINESS FAILURE PREDICTION IN RETAIL INDUSTRY: AN EMPIRICAL EVALUATION OF GENERIC BANKRUPTCY PREDICTION MODELS. *Academy of Accounting and Financial Studies Journal*, 97.
- Hernandez Tinoco, M., & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International Review of Financial Analysis*, 394-419.
- Ibarra Mares, A. (2009). *Desarrollo del análisis factorial multivariable aplicado al análisis financiero actual*. S.I,Argentina: B.EUMED.
- Irastorza Eduardo. (enero de 2019). *Cada español gastó de media 429 euros en ropa en 2018*. EAE.es. Recuperado de: <https://www.eae.es/actualidad/noticias/cada-espanol-gasto-una-media-de-429-euros-en-ropa-en-2018>
- Keener, M. H. (2013). Predicting The Financial Failure Of Retail Companies In The United States. *Journal of Business & Economics Research*, 373-380.
- Korol, T. (2013). Early warning models against bankruptcy risk for Central European and Latin American enterprises. *Economic Modelling*, 22-30.
- Mateos-Ronco, A., Marín-Sánchez, M. M., & Mari-Vidal, S. y. (2011). Los modelos de predicción del fracaso empresarial y su aplicabilidad en cooperativas agrarias. CIREC-España. *Revista de Economía Pública, Social y Cooperativa*, 179-208.
- Mckinsey & Company & BoF. (2018). *The State of Fashion 2018*.
- Nava Rosillón, M. A. (2009). Análisis financiero: una herramienta clave para una gestión financiera eficiente. *Revista Venezolana de Gerencia (RVG)*, 606-628.
- Nogueira, D., Medina, A., Hernández, A., Comas, R., Medina, & Daylin. (2017). Análisis económico-financiero: talón de Aquiles de la organización. Caso de aplicación. *Ingeniería Industrial*.

- Ochoa Ramírez, S. A., & Toscano Moctezuma, J. A. (2012). Revisión crítica de la literatura sobre el análisis financiero de las empresas. *Revista de Ciencias Sociales y Humanidades*, 73-100.
- Ohlson, J.A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*.
- Pérez, C. (2004). *Técnicas de análisis multivariante de datos*. Aplicaciones con SPSS. Madrid: Pearson educación.
- Pérez, A., Kizys, R., y Manzanedo, L. (2010). *Regresión Logística Binaria*. Recuperado el 20 de abril de 2021, del Sitio web del Departamento Publicaciones de la Universidad de Córdoba: <https://econometria.files.wordpress.com/2010/01/reg-logistica.pdf>
- REFOR. (2017). *Atlas Concursal*.
- Rome Espinosa, F. (2013). Alcances y limitaciones de los modelos de capacidad predictiva en el análisis del fracaso empresarial. *Red de Revistas Científicas de América Latina, el Caribe, España y Portugal*, 45-70.
- Sojo Calvo, F. J. (2012). Evolución y mejora en la competitividad de las empresas del sector textil-confección. *Economía Industrial*, 39-46.
- Stevens, Stanley. (1946). On the Theory of Scales of Measurement. *Science, New Series*, Vol. 103, No. 2684, pp. 677-680. American Association for the Advancement of Science.
- Subdirección General de Estudios y Modernización del Comercio. (2007). *El comercio textil en España*.
- Wild, J., Subramanyam, K., & Halsey, R. (2007). *Análisis de Estados Financieros*. En J. j. Wild, K. Subramanyam, & R. F. Halsey. México: McGraw-Hill Interamericana.

Anexo

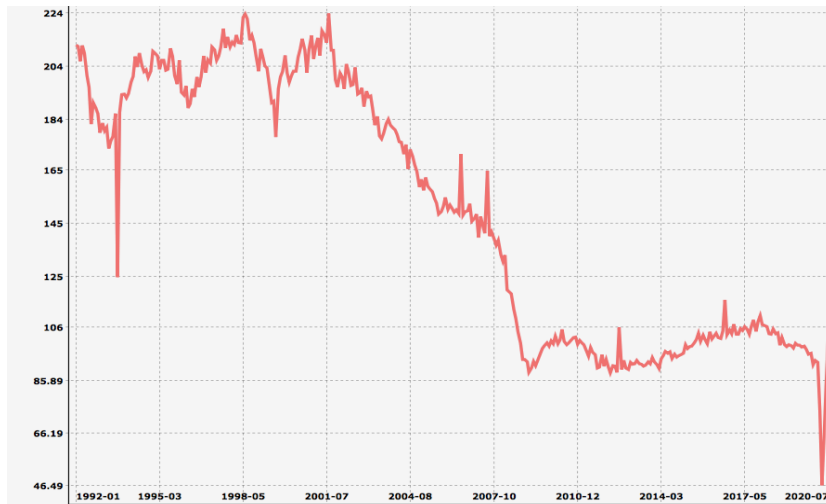


Gráfico 35: Evolución del Índice de Producción Industrial (IPI) industrial textil
Fuente: temáticas.org²⁴

```

Box's M-test for Homogeneity of Covariance Matrices

data: training.set[, 1:15]
Chi-Sq (approx.) = 463.99, df = 120, p-value < 2.2e-16
    
```

Gráfico 36: Test de homocedasticidad.
Fuente: Elaboración propia a partir de datos de SABI

```

          Df Pillai approx F num Df den Df   Pr(>F)
cualit    1 0.33586  17.497   15  519 < 2.2e-16 ***
Residuals 533
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

Gráfico 37: Resultados del contraste Manova para el modelo en conjunto
Fuente: Elaboración propia a partir de datos de SABI

²⁴ Fuente: tematicas.org/sintesis-economica/indicadores-de-produccion-y-demanda-nacional/ipi-industria-textil-correctas-de-calendario-y-estacionalidad/. Último acceso: febrero de 2021.

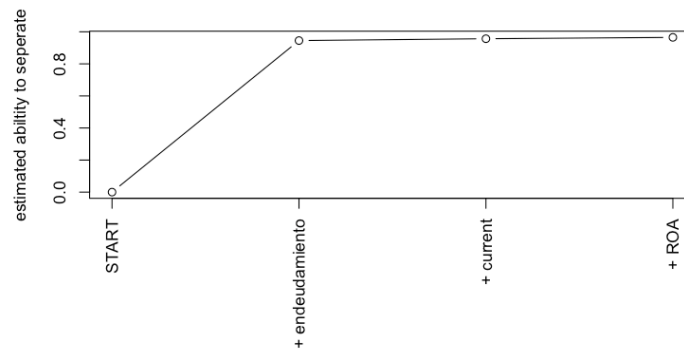


Gráfico 38: Poder de separación de las variables endeudamiento, current y ROA.
Fuente: Elaboración propia a partir de datos de SABI

	LD1
current	-0.5736328
ROA	-0.6747043
endeudamiento	0.9210491

Gráfico 39: Correlación de las variables con la función discriminante
Fuente: Elaboración propia a partir de datos de SABI

SOLVENCIA (LIQUIDEZ)	Quick ratio (prueba ácida)	$\frac{(AC - existencias)}{PC}$
	Current ratio (liquidez general)	$\frac{AC}{PC}$
	Ratio sobre EBITDA	$\frac{(Pasivo l/p + deudas c/p)}{EBITDA}$
RENTABILIDAD Y MÁRGENES	Ratio garantía (solvenca)	$\frac{AT}{PT}$
	Capital Ratio	$\left(\frac{PN}{AT}\right) \times 100$
	Rentabilidad económica (ROA)	$\left(\frac{EBIT}{ATN}\right) \times 100$
	Margen bruto	$\left(\frac{Ingresos explotación - consumo de materiales}{Ingresos explotación}\right) \times 100$
	Margen operativo	$\left(\frac{EBIT}{Ingresos explotación}\right) \times 100$
ACTIVIDAD Y EFICIENCIA	Margen neto	$\frac{BN}{Ingresos explotación} \times 100$
	Rentabilidad financiera (ROE)	$\left(\frac{BN}{FP}\right) \times 100$
	Gastos operativos sobre ingresos	$\left(\frac{Gastos operativos}{Ingresos explotación}\right) \times 100$
	Rotación activos	$\frac{Ingresos explotación}{AT}$
	Rotación cuentas por cobrar	$\frac{Ventas a crédito}{Ingresos Aprovisionamientos}$
	Rotación inventario	$\frac{PT}{Existencias}$
	Rotación cuentas por pagar	$\frac{Cuentas por pagar}{Coste de ventas}$
	Rotación ventas	$\frac{Ventas}{AT}$
	Gastos de personal sobre ingresos	$\left(\frac{Gastos personal}{Ingresos explotación}\right) \times 100$
	Salario	$\frac{Gastos personal}{Número empleados}$
ENDEUDAMIENTO	Productividad empleados	$\frac{Ingresos explotación}{Número empleados}$
	Endeudamiento total (apalancamiento)	$\frac{PT}{AT}$
	Cobertura intereses	$\left(\frac{EBIT}{Gastos de financieros}\right) \times 100$

Tabla 8: Ratios utilizadas
Fuente: Elaboración propia