ENFOQUES DE ANÁLISIS DE LOS DETERMINANTES DE LA OLEADA DE QUIEBRAS BANCARIAS EN LOS ESTADOS UNIDOS DURANTE LA CRISIS FINANCIERA INTERNACIONAL DE 2008

Presentada por
Manuel Merck Martel

Codirectores
Alvaro Hidalgo Vega

Agustín Álvarez Herranz

Diciembre 2015
UNIVERSIDAD DE CASTILLA–LA MANCHA

ABSTRACT

Determinantes de la oleada de quiebras bancarias en los Estados Unidos durante la crisis financiera internacional de 2008

Autor

Manuel Merck Martel

Codirectores

Álvaro Hidalgo Vega
Agustín Álvarez Herranz

Esta tesis desarrolla diversos enfoques para analizar los principales determinantes de las quiebras de entidades bancarias norteamericanas durante la reciente crisis financiera. El estudio inicial –desarrollado a lo largo del primer capítulo– plantea una exhaustiva explotación de la base de datos contables FFIEC. A continuación, el segundo capítulo ofrece una comparación entre un modelo clásico de alerta temprana EWS-CAMELS y un indicador naïve como el Ratio de Texas. El tercer capítulo propone la utilización del método de componentes principales, para incorporar el máximo de variables CAMELS en la especificación del modelo, evitando el problema de la multicolinealidad. El cuarto capítulo cierra el compendio condensando las principales conclusiones derivadas de las investigaciones realizadas. Globalmente, los resultados obtenidos a través de las diversas metodologías aplicadas sugieren que dichas técnicas son alternativas válidas para complementar los modelos insitu aplicados en la práctica supervisora.
TABLA DE CONTENIDOS

Agradecimientos ........................................................................................................................................ X
Introducción............................................................................................................................................... X

Capítulo I: Análisis de los determinantes de las quiebras bancarias en los Estados Unidos durante el periodo 2008-2012 ........................................................................................ X
 I.0 Abstract ................................................................................................................................................ X
 I.1 Introducción .......................................................................................................................................... X
 I.2 Antecedentes ....................................................................................................................................... X
 I.3 Metodología ......................................................................................................................................... X
 I.4 Resultados y discusión ................................................................................................................... X
 I.5 Conclusiones ....................................................................................................................................... X
 I.6 Referencias bibliográficas ................................................................................................................ X
 I.7 Anexo: definiciones de las variables de la ecuación 2 ................................................................... X

Capítulo II: Bank Failure Prediction Models: Less is more? ................................................................. X
 II.0 Objective, Abstract ............................................................................................................................ X
 II.1 Introduction ........................................................................................................................................ X
 II.2 Literature ........................................................................................................................................... X
 II.2.1 EWS implementation by U.S. Banking Supervisory Regulators ............................................ X
 II.2.1.1 FDIC model ............................................................................................................................. X
 II.2.1.3 OCC model ............................................................................................................................. X
 II.2.1.4 EWS literature review .......................................................................................................... X
 II.2.2 Texas Ratio ................................................................................................................................... X
 II.3. Counterfactual analysis .................................................................................................................... X
 II.3.1 Database management ................................................................................................................ X
 II.3.2 Econometric model specification ............................................................................................. X
 II.3.3 Variables selection ....................................................................................................................... X
 II.4 Findings ............................................................................................................................................ X
 II.5 Conclusions and policy suggestions .............................................................................................. X
 II.6 Statistical Annexes ........................................................................................................................ X
 II.7 References ........................................................................................................................................ X

Capítulo III: Análisis de quiebras mediante componentes principales ................................................. X
 III.1 Introducción .................................................................................................................................... X
 III.2 Revisión de la literatura .................................................................................................................. X
 III.3 Metodología empírica: construcción de los factores y especificaciones del modelo ................... X
 III.3.1 Análisis de componentes principales ...................................................................................... X
 III.3.1.1 ACP para las variables de Capital .................................................................................... X
 III.3.1.2 ACP para las variables de Asset Quality Summary ......................................................... X
 III.3.1.3 ACP para las variables de Earnings Summary ............................................................... X
III.3.1.4 ACP para las variables de Liquidity Summary ........................................X
III.3.1.5 ACP para las variables de Sensitivity Summary ......................................X
III.3.2 Planteamiento modelo logit de quiebras bancarias norteamericanas ...X
III.3.2.1 Modelo logit de quiebras bancarias en 2008 .........................................X
III.3.2.2 Modelo logit de quiebras bancarias en 2009 .........................................X
III.3.2.3 Modelo logit de quiebras bancarias en 2010 .........................................X
III.3.2.4 Modelo logit de quiebras bancarias en 2011 .........................................X
III.3.2.5 Modelo logit de quiebras bancarias en 2012 .........................................X
III.3.2.6 Conclusiones ...............................................................................................X

Capítulo IV: Resumen de hallazgos y conclusiones de ‘policy’ .........................X
Descripción de hallazgos .........................................................................................X
Principales conclusiones de ‘policy’ ........................................................................X
Quiero expresar mi más sincero agradecimiento a mis padres, María del Pilar y José Guillermo, por su apoyo incondicional. A mis Directores de Tesis, Álvaro Hidalgo y Agustín Álvarez, y a mi tutor de Master, Eloy González, por su amistad, y por su asistencia en la preparación de este documento.
INTRODUCCIÓN

El objetivo de este trabajo es la detección y el análisis de los principales factores subyacentes que podrían explicar la reciente oleada de cierres de entidades bancarias durante el período 2008-2012 en los Estados Unidos de Norteamérica.

Para alcanzar este propósito se han desarrollado tres enfoques metodológicos alternativos –y complementarios-, con la meta común de incorporar elementos novedosos a la literatura existente para identificar en qué medida las estructuras de los balances de los bancos quebrados presentan elementos diferenciales -con respecto al perfil medio de la industria y/o en relación con clusters específicos-, que caractericen el fracaso de estas entidades, o de sus modelos de negocio.

En consecuencia, el trabajo se ha dividido en tres apartados o capítulos, identificados con tres artículos en vías de publicación en revistas especializadas con factor de impacto JCR, medida de evaluación de la calidad científica de las revistas académicas proporcionada por el Journal Citation Report.

**En el primer capítulo** se utiliza la metodología de regresión logística como técnica estadística para la modelización de las quiebras bancarias. La particularidad –y el principal valor añadido- radica en el exhaustivo proceso exploratorio previo -inexistente hasta donde llega nuestra información- de las 1254 variables correspondientes a los estados contables de los 8.445 bancos existentes en la base de datos UBPR del FFIEC. Tras un proceso de filtrado y depuración -eliminando el truncamiento por la derecha y por la izquierda- se ha obtenido un panel balanceado que posteriormente se ha restringido a las primeras 30 variables con mayor inclusión en el total de entidades financieras, conservando únicamente las 6.997 entidades existentes durante el período completo de la muestra. A continuación, se ha realizado un procedimiento econométrico
tradicional bottom-up para la selección de las variables estadísticamente significativas a nivel individual y con signos esperados correctos, lográndose niveles de predicción similares a los modelos supervisores, a través de especificaciones que incorporan un número reducido de variables explicativas.

**En el segundo capítulo** se aborda un enfoque diferenciado con relación con la explotación innovadora de la base de datos que se ha realizado en el capítulo uno. Se exploran dos metodologías aceptadas en el ámbito financiero, que responden a diversas ópticas: un sistema tradicional de EWS inspirado en la metodología CAMEL, desarrollado desde el ámbito supervisor, y un enfoque procedente de la industria bancaria, basado en un indicador 'ingenuo' o heurístico, de carácter más intuitivo, el ratio de Texas. El objetivo es evaluar si el uso de relaciones heurísticas menos complejas puede ser la herramienta óptima para analizar eficazmente un entorno bancario cada vez más interconectado y concentrado.

**En el tercer capítulo** se han investigado las posibilidades de mejora del “modelo mixto” del capítulo anterior, planteando la introducción del mayor número posible de variables de cada grupo CAMEL, y evitando incurrir en el problema de la multicolinealidad, que surge cuando existe correlación lineal entre las variables explicativas del modelo provocando que los estimadores de los parámetros, aunque sean insesgados, sufran errores al aumentar o disminuir la muestra. Esto hace que sus varianzas y covarianzas crezcan al aumentar el grado de colinealidad, provocando una pérdida de fiabilidad en los contrastes de significación al incrementarse los intervalos de probabilidad, e invalidando la utilización del modelo para análisis estructurales. Nuestro estudio corrige la multicolinealidad a través de un análisis factorial mediante componentes principales.

**Por último, en el cuarto capítulo** se resumen los principales hallazgos y conclusiones, cerrando el planteamiento mediante unas breves sugerencias de ‘policy’ y señalando las futuras vías de investigación.
ANÁLISIS DE LOS DETERMINANTES DE LAS QUIEBRAS BANCARIAS EN LOS EE.UU. DURANTE EL PERÍODO 2008-2012
I.0. ABSTRACT

La industria bancaria ha experimentado períodos de crisis desde sus inicios, que han aglutinado múltiples quiebras y/o suspensiones de entidades financieras en episodios muy cortos, debido a múltiples factores pero con mayor frecuencia como resultado de ineficientes procesos de decisiones e inadecuadas estrategias de riesgos.

En 1984 el sistema bancario estadounidense alcanzó su maxima dimensión, al contar con una cifra de 15.000 entidades bancarias. Una década después, el tamaño de la industria disminuyó hasta alcanzar únicamente 10.741 entidades, como resultado de procesos de quiebras, fusiones y adquisiciones.

Tras un período de relativa estabilidad durante los años 90s, a finales de 2007 e inicios de 2008 se produjo una nueva oleada de distorsiones significativas en los mercados financieros que volvió a elevar los niveles de quiebras, provocando la preocupación de los clientes e inversores bancarios en relación con sus entidades financieras, lo que nuevamente motivó una particular atención en los medios especializados. En este contexto, se ha recuperado el interés por diferenciar qué entidades se encuentran próximas a un desequilibrio financiero o quiebra.

Para lograr este objetivo, este trabajo pretende identificar los principales factores subyacentes que explican las recientes quiebras de entidades bancarias en los Estados Unidos, a través de una exhaustiva explotación de la base de datos del Uniform Bank Performance Report (UBPR) del Federal Financial Institutions Examination Council (FFIEC), que aglutina la información procedente de los informes trimestrales del sector bancario de los EE.UU.

El proceso de identificación de los factores relevantes se ha realizado a través de una metodología cuantitativa basada en un modelo de regresión logístico, que ha
alcanzado un nivel de predicción de entidades quebradas superior a modelos ‘naive’, tales como el indicador conocido como “Ratio de Texas” desarrollado por Gerard Cassidy y otros (RBC Capital Markets).

Palabras clave: banca, finanzas, quiebras bancarias, sistemas de alerta temprana, crisis bancarias

I.1. INTRODUCCIÓN

Este trabajo presenta un análisis de los principales factores determinantes de la última oleada de quiebras de entidades bancarias en los Estados Unidos, durante el período 2008-2012. Se pretende detectar en qué medida la estructura de los balances individuales de los bancos quebrados puede presentar elementos diferenciales con respecto al perfil medio de la industria. De esta forma, se pretende identificar qué factores puedan contribuir a caracterizar el fracaso de sus planes de negocio, siguiendo la metodología de regresión logística como técnica estadística de modelización.

El término quiebra bancaria empleado en lo sucesivo sigue la definición de la Federal Deposit Insurance Corporation (FDIC)\(^1\), que hace referencia al “cierre de un banco por parte de una agencia regulatoria bancaria estatal o federal, generalmente cuando dicha entidad no es capaz de cumplir con sus obligaciones con los depositarios y otros acreedores”. El perímetro del análisis queda restringido al ámbito de las entidades aseguradas por la FDIC, compuesto por

\(^1\) [http://www.fdic.gov/consumers/banking/facts/](http://www.fdic.gov/consumers/banking/facts/). La creación de la FDIC como agencia federal independiente en 1933 respondió al objetivo de promover la confianza y estabilidad del sistema bancario de los EE.UU. El doble papel de esta institución ante un evento de quiebra es crítico: (i) en calidad de asegurador de los depósitos bancarios paga a los depositantes dentro de los límites establecidos (250,000$ por depositante y banco asegurado), y (ii) actúa como administrador del banco quebrado, asumiendo la gestión de sus activos y la pertinente liquidación de las deudas, atendiendo al orden de prelación legal.
entidades bancarias que disponen tanto de licencia gubernamental como estatal en los EE.UU..

Como principal fuente estadística se ha utilizado la base de datos del Uniform Bank Performance Report (UBPR) del Federal Financial Institutions Examination Council’s (FFIEC), que aglutina la información derivada de los informes trimestrales reportados por las entidades bancarias individuales aseguradas por el fondo de depósitos de los EEUU. De forma complementaria, en las especificaciones preliminares se han incluido datos macroeconómicos obtenidos de la Federal Reserve Economic Data (FRED) de la Reserva Federal del Estado de San Luis y los procedentes de los escenarios del stress test de la FED, datos de carácter público relativos a indicadores sobre la evolución del crédito, del sector de la construcción (residencial y comercial) y un grupo adicional de variables sociodemográficas que a priori pudieran ser relevantes.

El estudio plantea un indicador de vulnerabilidad bancaria, compuesto por una relación de ratios contables que han demostrado una significación previa en los principales modelos relevantes en la literatura de quiebras bancarias. El procedimiento de selección de variables explicativas responde a la metodología de análisis 'bottom-up & top-down'. Dicha investigación se complementa introduciendo la influencia de diversos factores del entorno macroeconómico –a nivel estatal- que podrían haber adelantado o amplificado los problemas acaecidos, afectando muy particularmente a los “community banks” ubicados en la franja del 'Sun Belt'. Adicionalmente, se evalúa el potencial de mejora del modelo mediante la inclusión de ratios de clasificación del modelo de negocio bancario, atendiendo al grado de identificación de los perfiles de actividad con la banca comercial (minorista) o de inversiones (mayorista).

El trabajo se articula en cuatro secciones además de la presente. La sección segunda repasa brevemente la literatura previa. La sección tercera describe la
metodología y la base de datos. En la cuarta sección se presentan los principales hallazgos empíricos del modelo, y la sección quinta plantea las principales conclusiones de nuestra investigación.

I.2. ANTECEDENTES

Tras la crisis financiera de la Gran Depresión, a finales de la década de los años 20s y principios de los 30s surgieron los primeros análisis de los factores determinantes de las quiebras bancarias (Spahr, 1932). A finales de los 60s se reavivó el interés por evaluar el grado de viabilidad de las empresas industriales a través del estudio de sus principales ratios financieros y contables, apareciendo modelos pioneros basados en una metodología de análisis discriminante univariante (Beaver 1966) y multivariante (Altman 1968).

El enfoque de análisis discriminante se trasladó al ámbito de las quiebras de entidades bancarias a comienzos de los 70s, considerando ratios específicos de la actividad bancaria (Altman et al., 1977, Meyer and Pifer, 1970, Sinkey, 1975, Stuhr and Van Wicklen, 1974). Al final de dicha década quedo corroborada la frecuente violación de las exigentes hipótesis de los modelos discriminantes por parte de los datos financieros reales (ej. distribución normal en partidas financiero-contables, o igualdad de las matrices de varianzas-covarianzas de los bancos quebrados y no quebrados).

Ante el fracaso de la metodología del análisis discriminante, surgió la regresión logística (Martin, 1977) como alternativa para explicar las quiebras bancarias, al no formular hipótesis restrictivas con respecto a las distribuciones de las variables predictivas. El profesor Martin verificó la robustez de los modelos logit, al examinar por primera vez el sistema de bancos comerciales miembros del Sistema
de la Reserva Federal en 1974 (23 quiebras, y 5.575 bancos activos). Su estudio incorporó 25 ratios procedentes de los estados financieros, divididos en 4 grupos según la naturaleza de la información incorporada: (C) adecuación del capital, (A) calidad de activos, (E) resultados y (L) liquidez. El modelo resultante incluyó cuatro variables relativas al capital (C), la calidad de activos (A) y los beneficios (E).


Al realizar un análisis global de los modelos logit más relevantes de las últimas tres décadas se comprueba que sus especificaciones han incorporado un número de ratios financiero-contables que oscila entre 9 y 27, resultando significativos -y con los signos correctos- un número similar de variables explicativas, que ha fluctuado entre 4 y 7 (Tabla 1).
Tabla 1. Variables significativas de modelos logit aplicados a las quiebras bancarias

<table>
<thead>
<tr>
<th>Estudio</th>
<th>Modelo</th>
<th>Ratios Incorporados</th>
<th>Ratios Significativos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin (1977)</td>
<td>Logit</td>
<td>25</td>
<td>GCARA, CI2LN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GCONI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NITA</td>
</tr>
<tr>
<td>Avery &amp; Hanweck (1984)</td>
<td>Logit</td>
<td>9</td>
<td>KTA, LNTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NLTA, CILNNL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NITA</td>
</tr>
<tr>
<td>Barth y otros (1985)</td>
<td>Logit</td>
<td>12</td>
<td>NWTA, LNTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ISFTF</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NITA, LATA</td>
</tr>
<tr>
<td>Thomson (1991)</td>
<td>Logit</td>
<td>16</td>
<td>NCAPTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NCOTA, INSLNTA, NLTA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OVRHDTA, ROA, LIQ</td>
</tr>
<tr>
<td>Andersen (2008)</td>
<td>Logit</td>
<td>27</td>
<td>CAR, ELOSS, CONS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RMGL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ROA, NBLI</td>
</tr>
</tbody>
</table>

Fuente: Elaboración propia a partir de Van der Ploeg (2010)

En las últimas décadas se han desarrollado y depurado los sistemas de alerta temprana (EWS) que pretenden detectar con antelación los eventos de quiebra, al objeto de evitar o al menos minimizar sus coste (Oct et al., 2011, Putnam, 1983, Rose and Kolari, 1985, Wheelock and Wilson, 2000). Igualmente se han desarrollado modelos de riesgos (Cox) aplicados al sector bancario (Lane et al.,

---

2 GCARA = Gross Capital/Adjusted Risk Assets; CI2LN = (Commercial and Industrial Loans + Loans to REITs and Mortgage Bankers + Construction Loans + Commercial Real Estate Loans)/Total Assets; KTA = (Equity Capital + Loan Loss Reserve Allowances)/Total Assets; LNTA = Natural Logarithm of Total Assets; NWTA = Net Worth/Total Assets; NCAPTA = (Book Equity + Loan Loss Reserves – Loans 90 Days Past Due - Non-Accruing Loans)/Total Assets; CAR = Capital Adequacy Ratio; RMGL = Residential Mortgages/Gross Lending; ELOSS = Expected Loss based on PD/Gross Lending; CONS = Herfindahl Index for Loan Portfolio.

3 GCONI = Charge-Offs/(Net Operating Income + Loss Provision); NLTA = Net Loans/Total Assets; CILNNL = Commercial and Industrial Loans/Net Loans. ISFTF = Interest Sensitive Funds/Total Funds; NCOTA = Net Charge-Offs/Total Assets; INSLNTA = Loans to Insiders/Total Assets; RMGL = Residential Mortgages/Gross Lending

4 OVRHDTA = Overhead/Total Assets.

5 NITA = Net Income/Total Assets; ROA = Return On Assets.

6 LATA = Liquid Assets/Total Assets; LIQ = (Federal Funds Purchased – Federal Funds Sold)/Total Assets; NBLI = Norges Bank’s Liquidity Indicator.

Es precisamente la necesidad de conjugar la metodología de los modelos logit, complejos pero con resultados robustos, con la sencillez de los modelos naive lo que fundamenta este trabajo. Nuestro modelo pretende ser una alternativa a los modelos naive, utilizando la base metodológica de los modelos logit. En la siguiente sección presentamos los objetivos del trabajo. En la tercera se muestra la metodología del modelo y los datos empleados. En la sección cuarta se presentan los resultados y la discusión de los mismos. Por último finalizamos con una serie de conclusiones.

I.3. METODOLOGIA

El período 2007-2008 ha venido caracterizado por una oleada de distorsiones significativas en los mercados financieros, con el resultado de una pronunciada elevación en los niveles de quiebras bancarias, y la consiguiente preocupación por parte de clientes e inversores bancarios.

En este contexto, la investigación desarrollada pretende identificar los indicadores más relevantes para detectar esos desequilibrios con el suficiente grado de anticipación. El trabajo elaborado en el presente apartado se centra en la presentación de un modelo preliminar, al objeto de monitorizar en qué medida la estructura de los balances de los bancos individuales quebrados puede presentar
elementos diferenciales con respecto al perfil medio de la industria, de forma que
puedan contribuir a caracterizar el fracaso de entidades específicas –o de sus
modelos de negocio–, empleándose como enfoque cuantitativo central la
metodología de regresión logística.

La opción seleccionada de un modelo de elección discreta frente a la econometría
tradicional (OLS)\(^7\) radica en su mayor eficiencia de cara a la modelización de
variables cualitativas, a través del empleo de técnicas específicas de tratamiento de
variables discretas\(^8\). Esta característica exige la codificación como paso previo a la
modelización, proceso por el cual las alternativas de las variables se transforman
en códigos o valores cuánticos, susceptibles de ser modelizados utilizando
técnicas econométricas.

En la literatura de interpretación estructural de los modelos de elección discreta
existen dos enfoques principales. El primero hace referencia a la modelización de
una variable latente a través de una función índice, que trata de especificar una
variable inobservable o latente. El segundo de los planteamientos permite
interpretar los modelos de elección discreta bajo la teoría de la utilidad aleatoria,
de tal manera que la alternativa seleccionada en cada caso será aquella que
maximice el nivel de utilidad esperado.

El presente estudio sigue el enfoque de la función índice, no limitada en su rango
de variación, \(\Gamma_i\). La variable latente \(\Gamma_i\) está relacionada con sus características a
través de un modelo de regresión:

\[ I_i = X_i \beta + \varepsilon_i \]

donde \(X_i \beta\) recibe el nombre de función índice.

\(^7\) Mínimos Cuadrados Ordinarios ('Ordinary Least Squares')
\(^8\) Se dice que una variable es discreta cuando está formada por un número finito de alternativas que miden cualidades.
Cuando la variable latente supera un determinado nivel, la variable discreta toma el valor 1, y si no lo supera toma el valor 0. Esta variable a su vez depende de un conjunto de variables explicativas que generan las alternativas existentes que se dan en el espectro real, permitiendo expresar el modelo dicotómico a través de la variable indicadora $Y_i$, definida por:

$$Y_i = \begin{cases} 
1 & \text{si } I_i^* > 0 \text{ lo que ocurre cuando } X_i \beta + \varepsilon_i > 0 \\
0 & \text{si } I_i^* < 0 \text{ lo que ocurre cuando } X_i \beta + \varepsilon_i < 0 
\end{cases}$$

En la anterior expresión el supuesto realizado sobre la distribución de $\varepsilon_i$ determina el tipo de modelo a estimar. El presente estudio asume una distribución de probabilidad ajustada a la curva de la función logística, tratándose por tanto de un modelo Logit. La hipótesis de que el umbral a superar por la variable latente sea cero puede ser sustituida por cualquier otro valor, sugiriéndose en determinados estudios que el valor crítico sea el definido por el término constante.

Bajo este enfoque el modelo probabilístico quedaría definido por:

$$P_i = \text{Prob}(Y_i = 1) = \text{Prob}(I_i^* > 0) = \text{Prob}(X_i \beta + \varepsilon_i > 0) = F(X_i \beta)$$

Siguiendo la definición anterior, la variable endógena del modelo dicotómico representa la probabilidad de ocurrencia del fenómeno analizado, en este caso la quiebra de entidades bancarias, siendo la probabilidad de la opción 1 más elevada cuanto mayor sea el valor de $I_i^*$.

---

[9] El modelo probit es otra forma de GLM también basado en una regresión binomial cuya variable dicotómica, y en una estimación máximo verosímil que generalmente produce resultados muy similares al logit. Sin embargo, su construcción se fundamenta en la distribución normal estándar, en lugar de la logística estándar.
A continuación se muestra el modelo logit general de partida, cuyo inicio teórico correspondiera a la inclusión de 1254 variables –clasificadas en 12 grupos principales– para el periodo comprendido entre el primer trimestre del año 2003 (2003.01) hasta el último trimestre de 2007 (2007.04):

\[
\text{Quiebra}_{it} = \alpha_0 + \sum_{j=1}^{262} \sum_{l=0}^{8.445} \beta_{j, Balance}_{it} + \sum_{j=1}^{70} \sum_{l=0}^{8.445} \delta_{j, Capital}_{it} + \sum_{j=1}^{51} \sum_{l=0}^{8.445} \phi_{j, Cartera}_{it} + \\
+ \sum_{j=1}^{28} \sum_{l=0}^{8.445} \gamma_{j, Crédito}_{it} + \sum_{j=1}^{189} \sum_{l=0}^{8.445} \eta_{j, Custodia}_{it} + \sum_{j=1}^{76} \sum_{l=0}^{8.445} \iota_{j, Derivados}_{it} + \sum_{j=1}^{87} \sum_{l=0}^{8.445} \lambda_{j, Morosidad}_{it} + \\
+ \sum_{j=1}^{77} \sum_{l=0}^{8.445} \mu_{j, Provisiones}_{it} + \sum_{j=1}^{41} \sum_{l=0}^{8.445} \theta_{j, Ratiocinos}_{it} + \sum_{j=1}^{153} \sum_{l=0}^{8.445} \rho_{j, Resultados}_{it} + \sum_{j=1}^{20} \sum_{l=0}^{8.445} \varphi_{j, Riesgo}_{it} + \\
+ \sum_{j=1}^{91} \sum_{l=0}^{8.445} \upsilon_{j, Titulizaciones}_{it} + \epsilon_{it} 
\]  \hspace{2cm} (1)

La fuente estadística utilizada para la exploración del modelo logit óptimo a partir del modelo general (1), correspondiente al período desde 2003.01 y hasta el 2007.04, que permita la detección temprana de las quiebras de las entidades bancarias que se sucedieron a lo largo del período posterior (2008-2012), es la base de datos UBPR de la FFIEC, utilizada por la FDIC.

Esta base de datos aglutina la información derivada de los informes trimestrales reportados por las entidades bancarias individuales aseguradas por el fondo de depósitos bancarios norteamericano. Esencialmente se trata de una herramienta analítica, creada y desarrollada para propósitos de supervisión, examen y gestión bancaria. Mediante un formato conciso, muestra el impacto de las decisiones de gestión y de las condiciones económicas sobre el rendimiento bancario y la composición del balance. Los datos de rendimiento y composición contenidos en el informe pueden ser empleados como soporte en la evaluación sobre el grado de adecuación de los beneficios, la liquidez, el capital, la gestión de activos y pasivos, y la gestión del crecimiento de los balances bancarios.
I.4. RESULTADOS Y DISCUSIÓN

Durante la fase inicial de la investigación se ha realizado un análisis cuantitativo exploratorio de las matrices de datos de las 1254 variables correspondientes a los estados contables de los 8.445 bancos existentes en la base de datos UBPR, al objeto de determinar el número de valores perdidos y la existencia de sesgos fuertes en la frecuencia de algún valor dentro de las variables -durante el período analizado-, que posteriormente pudieran condicionar los resultados del modelo logit de quiebras bancarias.

Como resultado de dicho análisis preliminar, hemos podido distinguir tres grupos de variables, denominadas y caracterizadas según el detalle que se acompaña:

a) **Variables descartadas**: Dicha clasificación se aplica a las variables que muestran un sesgo del valor 0 con frecuencia superior al 90% sobre el total de observaciones que la componen, y/o presentan un porcentaje de valores perdidos en el periodo temporal considerado superior al 50% del tamaño muestral.

b) **Variables incompletas**: Este grupo de variables no presenta un sesgo de un calibre tan importante en la frecuencia de algún valor incluido, presentando un porcentaje de valores perdidos considerable, aunque inferior al 50% de la muestra. Este grupo de variables queda reservado para su utilización en estudios posteriores, con menores requerimientos (corte transversal o paneles de datos de inferior tamaño).

c) **Variables aceptadas**: Son las variables consideradas en el presente estudio, por ser variables sin sesgos importantes en la frecuencia de sus valores, y cuyo porcentaje de valores perdidos es residual (despreciable
en términos cuantitativos), al corresponder únicamente a un grupo reducido de entidades financieras de las que no se dispone de un juego completo de todas las observaciones, para el período temporal comprendido entre el primer trimestre del 2003 y el cuarto trimestre del 2007.

Por consiguiente, las variables analizadas de la base de datos UBPR quedarían clasificadas -a efectos del análisis- tal y como aparecen en la Tabla 2.

Este primer trabajo se centra únicamente en las 302 variables aceptadas (Ecuación 2):

\[
\text{Quiebra}_t = \alpha_0 + \sum_{j=1}^{57} \sum_{i=1}^{8445} \sum_{l=0}^{20} \beta_{ijl} \text{Balance}_{ijl} + \sum_{j=1}^{52} \sum_{i=1}^{8445} \sum_{l=0}^{20} \delta_{ijl} \text{Capital}_{ijl} + \sum_{j=1}^{20} \sum_{i=1}^{8445} \sum_{l=0}^{20} \phi_{ijl} \text{Cartera}_{ijl} + \\
+ \sum_{j=1}^{9} \sum_{i=1}^{8445} \sum_{l=0}^{20} \gamma_{ijl} \text{Crédito}_{ijl} + \sum_{j=1}^{3} \sum_{i=1}^{8445} \sum_{l=0}^{20} \lambda_{ijl} \text{Morosidad}_{ijl} + \sum_{j=1}^{15} \sum_{i=1}^{8445} \sum_{l=0}^{20} \mu_{ijl} \text{Provisiones}_{ijl} + \sum_{j=1}^{39} \sum_{i=1}^{8445} \sum_{l=0}^{20} \beta_{ijl} \text{Ratios}_{ijl} + \\
- \sum_{j=1}^{96} \sum_{i=1}^{8445} \sum_{l=0}^{20} \rho_{ijl} \text{Resultados}_{ijl} + \sum_{j=1}^{26} \sum_{i=1}^{8445} \sum_{l=0}^{20} \zeta_{ijl} \text{Riesgo}_{ijl} + \sum_{j=1}^{5} \sum_{i=1}^{8445} \sum_{l=0}^{20} \nu_{ijl} \text{Titulizaciones}_{ijl} + \epsilon_{it} \tag{2}
\]
Tabla 2. Clasificación posterior al análisis exploratorio de las variables (Base de datos UBPR).

<table>
<thead>
<tr>
<th>Grupo</th>
<th>Total</th>
<th>Balance</th>
<th></th>
<th>Capital</th>
<th></th>
<th>Cartera</th>
<th>Crédito</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Balance $</td>
<td>Balance%</td>
<td>OffBalance</td>
<td>Capital</td>
<td>Capital</td>
<td></td>
</tr>
<tr>
<td>Aceptadas</td>
<td>302</td>
<td>27</td>
<td>21</td>
<td>9</td>
<td>18</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>Incompletas</td>
<td>141</td>
<td>64</td>
<td>2</td>
<td>13</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Descartadas</td>
<td>811</td>
<td>58</td>
<td>18</td>
<td>50</td>
<td>18</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td>Total</td>
<td>1254</td>
<td>149</td>
<td>41</td>
<td>72</td>
<td>40</td>
<td>30</td>
<td>51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grupo</th>
<th>Custodia</th>
<th>Derivados</th>
<th>Morosidad</th>
<th>Provisiones</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Custodia1</td>
<td>Custodia2</td>
<td>Derivado1</td>
<td>Derivado2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Morosidad1</td>
<td>Morosidad2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Provisiones1</td>
<td>Provisiones2</td>
</tr>
<tr>
<td>Aceptadas</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Incompletas</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Descartadas</td>
<td>125</td>
<td>64</td>
<td>40</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>125</td>
<td>64</td>
<td>40</td>
<td>36</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grupo</th>
<th>Ratios</th>
<th>Resultados</th>
<th>Riesgo</th>
<th>Titulaciones</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RdoOperativo</td>
<td>Resultados $</td>
<td>Resultado %</td>
<td>Titulaciones1</td>
</tr>
<tr>
<td>Aceptadas</td>
<td>39</td>
<td>26</td>
<td>38</td>
<td>32</td>
</tr>
<tr>
<td>Incompletas</td>
<td>1</td>
<td>8</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>Descartadas</td>
<td>1</td>
<td>8</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>42</td>
<td>62</td>
<td>49</td>
</tr>
</tbody>
</table>

Con el propósito de obtener un panel balanceado, partiendo del anterior grupo se han seleccionado las primeras 30 variables que cumplían una mayor inclusión en el total de entidades financieras, conservando en el modelo únicamente las 6,997 entidades existentes durante el período completo. Estas variables han sido relativizadas respecto a los activos totales, recogiendo su contribución con independencia de su tamaño del balance, ecuación (3).
El modelo especificado mediante la inclusión de este grupo inicial de 30 variables\textsuperscript{10} ha mostrado una fuerte correlación entre las diversas variables incluidas, pero ha permitido comprobar el grado de significación individual de cada una de ellas, y el contraste de signos esperados.

\begin{equation}
Quiebra_t = \alpha + \sum_{t=1}^{20} \sum_{j=1}^{20} \beta_{j,t} \frac{Balance_{j,t}}{ActivosTotales_t} + \sum_{t=1}^{20} \sum_{j=1}^{20} \gamma_{j,t} \frac{Capital_{j,t}}{ActivosTotales_t} + \sum_{t=1}^{20} \sum_{j=1}^{20} \delta_{j,t} \frac{Morosidad_{j,t}}{ActivosTotales_t} + \sum_{t=1}^{20} \sum_{j=1}^{20} \mu_{j,t} \frac{Provisiones_{j,t}}{ActivosTotales_t} + \sum_{t=1}^{20} \sum_{j=1}^{20} \rho_{j,t} \frac{Resultados_{j,t}}{ActivosTotales_t} + \epsilon_t
\end{equation}

Por lo tanto, el modelo se puede expresar como:

\begin{align*}
& Quiebra_t = -6661.257 + 159.9644 \text{UBPR3210/UBPR2170} + 11.59021 \text{UBPR3230/UBPR2170} + 6.864412 \text{UBPR3839/UBPR2170} + 10.97492 \text{UBPR4340/UBPR2170} + 7.932891 \text{UBPRB530/UBPR2170} - 12.48059 \text{UBPRE625/UBPR2170} - 1.338690 \text{UBPR0081/UBPR2170} \nonumber \nonumber \\
& + 6.613858 \text{UBPR1590/UBPR2170} - 0.018443 \text{UBPR2145/UBPR2170} + 21.82012 \text{UBPR3123/UBPR2170} - 2.628137 \text{UBPR6164/UBPR2170} - 16.49462 \text{UBPRE129/UBPR2170} 
\end{align*}

El modelo se ha estimado mediante el método ML - Binary Logit (Newton-Raphson) y el período temporal considerado es 2003.01-2007.04 (Ecuación 3).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeficiente</th>
<th>Error Std.</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-6661.257</td>
<td>3826.096</td>
<td>-1.741006</td>
<td>0.0817</td>
</tr>
<tr>
<td>UBPR3210/UBPR2170</td>
<td>159.9644</td>
<td>40.33229</td>
<td>3.966163</td>
<td>0.0001</td>
</tr>
<tr>
<td>UBPR3230/UBPR2170</td>
<td>11.59021</td>
<td>0.568943</td>
<td>20.37148</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPR3839/UBPR2170</td>
<td>6.864412</td>
<td>0.346348</td>
<td>19.81941</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPR4340/UBPR2170</td>
<td>10.97492</td>
<td>2.211827</td>
<td>4.961925</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPRB530/UBPR2170</td>
<td>7.932891</td>
<td>5.391117</td>
<td>1.471474</td>
<td>0.1412</td>
</tr>
<tr>
<td>UBPRE625/UBPR2170</td>
<td>-12.48059</td>
<td>3.140796</td>
<td>-3.973703</td>
<td>0.0001</td>
</tr>
<tr>
<td>UBPR0081/UBPR2170</td>
<td>-1.338690</td>
<td>0.818058</td>
<td>-1.636423</td>
<td>0.1018</td>
</tr>
<tr>
<td>UBPR1590/UBPR2170</td>
<td>-6.613858</td>
<td>0.514324</td>
<td>-12.85931</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPR2145/UBPR2170</td>
<td>-0.018443</td>
<td>1.153043</td>
<td>-0.015995</td>
<td>0.9872</td>
</tr>
<tr>
<td>UBPR3123/UBPR2170</td>
<td>21.82012</td>
<td>3.611672</td>
<td>6.041556</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPR6164/UBPR2170</td>
<td>-2.628137</td>
<td>0.838356</td>
<td>-3.134869</td>
<td>0.0017</td>
</tr>
<tr>
<td>UBPRE129/UBPR2170</td>
<td>-16.49462</td>
<td>1.955400</td>
<td>-8.435421</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

\textsuperscript{10} En el anexo se recogen los nombres y descripción de las 30 variables utilizadas en el modelo.
<table>
<thead>
<tr>
<th></th>
<th>Estimado/Hipótesis Nula</th>
<th>Signo</th>
<th>Significancia</th>
<th>Coeficiente</th>
<th>P-valor</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBPRK426/UBPR2170</td>
<td>1,650930</td>
<td>0,144179</td>
<td>11,45058</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPRK437/UBPR2170</td>
<td>4,415143</td>
<td>0,146519</td>
<td>30,13352</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPR1410/UBPR2170</td>
<td>1,198111</td>
<td>0,378613</td>
<td>3,164476</td>
<td>0,0016</td>
<td></td>
</tr>
<tr>
<td>UBPRE037/UBPR2170</td>
<td>-3,785762</td>
<td>1,006426</td>
<td>-3,761591</td>
<td>0,0002</td>
<td></td>
</tr>
<tr>
<td>UBPR1400/UBPR2170</td>
<td>11,14317</td>
<td>1,430272</td>
<td>7,790946</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPRD668/UBPR2170</td>
<td>3,867343</td>
<td>1,828858</td>
<td>2,608033</td>
<td>0,0091</td>
<td></td>
</tr>
<tr>
<td>UBPR0071/UBPR2170</td>
<td>-3,858300</td>
<td>0,69307</td>
<td>-6,332279</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPRG105/UBPR2170</td>
<td>-172,9033</td>
<td>40,6876</td>
<td>-4,252547</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPR3300/UBPR2170</td>
<td>6662,961</td>
<td>3826,996</td>
<td>1,741452</td>
<td>0,0816</td>
<td></td>
</tr>
<tr>
<td>UBPRB528/UBPR2170</td>
<td>-1,516669</td>
<td>0,38285</td>
<td>-4,233140</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPRD662/UBPR2170</td>
<td>-0,181650</td>
<td>3,935852</td>
<td>-0,046153</td>
<td>0,9032</td>
<td></td>
</tr>
<tr>
<td>UBPRD665/UBPR2170</td>
<td>-12,09070</td>
<td>0,605306</td>
<td>-19,97452</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPRE116/UBPR2170</td>
<td>-0,078954</td>
<td>0,415059</td>
<td>-0,190223</td>
<td>0,8491</td>
<td></td>
</tr>
<tr>
<td>UBPRE119/UBPR2170</td>
<td>3,207187</td>
<td>0,559104</td>
<td>5,736301</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPRE120/UBPR2170</td>
<td>-1,385029</td>
<td>0,269423</td>
<td>-5,140731</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPRE123/UBPR2170</td>
<td>-4,396192</td>
<td>0,600216</td>
<td>-7,324345</td>
<td>0,0000</td>
<td></td>
</tr>
<tr>
<td>UBPRC233/UBPR2170</td>
<td>-1,616964</td>
<td>14,77782</td>
<td>-0,109418</td>
<td>0,9129</td>
<td></td>
</tr>
<tr>
<td>UBPR2200/UBPR2170</td>
<td>-2,032141</td>
<td>0,180904</td>
<td>-11,23327</td>
<td>0,0000</td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>McFadden R-squared</td>
<td>0,175065</td>
<td>Mean dependent var</td>
<td>0,046286</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akaike info criterion</td>
<td>0,308555</td>
<td>Sum squared resid</td>
<td>5613,590</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwarz criterion</td>
<td>0,310736</td>
<td>Log likelihood</td>
<td>-21561,05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hannan-Quinn criter.</td>
<td>0,309207</td>
<td>Deviance</td>
<td>43122,09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LR statistic</td>
<td>9342,033Prob(LR statistic)</td>
<td>0,000000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Obs with Dep=0 | 133478 | Total obs | 139956 |       |       |
| Obs with Dep=1 | 6478 |       |       |       |       |

El análisis individual de variables en función de los signos estimados/esperados ha contribuido con una serie de argumentos de cara a su inclusión o exclusión en ecuaciones posteriores. La realización de un análisis de correlaciones bivariadas entre las variables explicativas de la ecuación ha permitido detectar la existencia -o no- de un grado de multicolinealidad que pueda modificar los niveles de significación del contraste del Z-Statistic, aceptando la Hipótesis Nula, y el signo de los coeficientes. Las variables que muestran un mayor nivel de colinealidad son mostradas en la Tabla 4:
Tabla 3. Tabla de correlaciones máximas en la Ecuación 3

<table>
<thead>
<tr>
<th></th>
<th>UBPRG105/UBPR2</th>
<th>UBPRD662/UBPR2</th>
<th>UBPR3210/UBPR2</th>
<th>UBPRB528/UBPR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBPRG105/UBPR2</td>
<td>1.000000</td>
<td>-0.999163</td>
<td>0.999949</td>
<td>-0.429398</td>
</tr>
<tr>
<td>UBPRD662/UBPR2</td>
<td>-0.999163</td>
<td>1.000000</td>
<td>-0.999065</td>
<td>0.428214</td>
</tr>
<tr>
<td>UBPR3210/UBPR2</td>
<td>0.999949</td>
<td>-0.999065</td>
<td>1.000000</td>
<td>-0.429274</td>
</tr>
<tr>
<td>UBPRB528/UBPR2</td>
<td>-0.429398</td>
<td>0.428214</td>
<td>-0.429274</td>
<td>1.000000</td>
</tr>
</tbody>
</table>

Dichas variables corresponden a:

- **UBPRB528**: Préstamos y leasing, netos de ganancias no realizadas
- **UBPR3210**: Capital bancario total
- **UBPRG105**: Capital total incluyendo intereses minoritarios en filiales consolidadas
- **UBPRD662**: Importe total de deudas
- **UBPR2170**: Activos totales

Comprobada la existencia de multicolinealidad en la ecuación 3, los criterios utilizados para la selección de las variables integrantes del modelo propuesto son:

**1) Introducción de variables CAMEL:**

\[ C = \text{Variables de capital. En principio tienen que tener signo negativo, si ello significa menos quiebra en ese modelo. (ubpr3210).} \]
A = Calidad de activos. La variable que incluiremos es ubprd668, Importe de préstamos y leasing impagados 30-89 días, cuyo tendrían que tener signo positivo.

M = Gestión.

E = Resultados.

L = Liquidez. Hemos considerado la variable UBPRE120, Títulos de Deuda del Tesoro y de las Agencias Públicas de los EEUU.

2) Partiendo de la primera premisa, variables CAMEL, la segunda condición es la de introducir en nuestras ecuaciones el mayor número de estas variables que pertenezcan al Modelo SCOR de la FDIC. En nuestra primera selección de 30 variables las variables que cumplen esta condición son:

- UBPRE123 : Activos productivos totales
- UBPRD668: Importe de préstamos y leasing impagados 30-89 días

Por último, la Tabla 5 recoge la ecuación estimada para el período temporal 2003.01- 2007.04, cumpliendo con las dos premisas establecidas:
Tabla 4. Modelo estimado

Dependent Variable: DQUIEBRA  
Method: ML - Binary Logit (Quadratic hill climbing)  
Sample: 2003Q1 2007Q4  
Included observations: 139956  
Convergence achieved after 7 iterations  
Covariance matrix computed using second derivatives

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-1.392329</td>
<td>0.291427</td>
<td>-4.777626</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPR1410/UBPR2170</td>
<td>3.330455</td>
<td>0.114185</td>
<td>29.16718</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPR3210/UBPR2170</td>
<td>-4.032095</td>
<td>0.408588</td>
<td>-9.868368</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPRE120/UBPR2170</td>
<td>-1.890158</td>
<td>0.180818</td>
<td>-10.45340</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPRK437/UBPR2170</td>
<td>5.110413</td>
<td>0.129830</td>
<td>39.36246</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPRE123/UBPR2170</td>
<td>-3.242165</td>
<td>0.319356</td>
<td>-10.15219</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPRD665/UBPR2170</td>
<td>-11.80571</td>
<td>0.494653</td>
<td>-23.86666</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

McFadden R-squared   0.146837  
Schwarz criterion  0.320410  
LR statistic  7703.698  
Prob(LR statistic)  0.000000  

Obs with Dep=0         133478  
Total obs                 139956  
Obs with Dep=1          6478

Estimated Equation Success cutoff: C = 0.034

<table>
<thead>
<tr>
<th>% Correct</th>
<th>Dep=0</th>
<th>Dep=1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Correct</td>
<td>60.76</td>
<td>83.71</td>
<td>61.82</td>
</tr>
<tr>
<td>% Incorrect</td>
<td>39.24</td>
<td>16.29</td>
<td>38.18</td>
</tr>
</tbody>
</table>

Se comprueba que el modelo alcanzan una predicción correcta de entidades quebradas del 83,71% (Inversos de los Errores Tipo I). En términos relativos, se trata de un nivel de eficiencia elevado, sobre todo en comparación con modelos más ‘naive’ como el indicador conocido como “Ratio de Texas”, desarrollado por Gerard Cassidy y otros (RBC Capital Markets). Dicho coeficiente alcanzó unos niveles de predicción del 82% en modelos a corto plazo (inferiores al año).

Uno de los factores principales diferenciales del modelo expuesto es su elevada capacidad como sistema de alerta temprana de mayor alcance (medio-largo plazo), al predecir satisfactoriamente una importante proporción de quiebras acumuladas durante un período total que se extiende a 16 trimestres (4 ejercicios completos), lo que representa la práctica totalidad de quiebras bancarias ocurridas
durante la reciente crisis bancaria en los EE.UU. Las definiciones de las variables incluidas en la ecuación 3 se recogen en la Tabla 6.

Tabla 6. Variables incluidas en la Ecuación 3

<table>
<thead>
<tr>
<th>Variables Modelo 3</th>
<th>Definición</th>
<th>signo estimado</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBPR1410</td>
<td>Préstamos Garantizados por Activos Inmobiliarios</td>
<td>+</td>
</tr>
<tr>
<td>UBPR3210</td>
<td>Capital Bancario Total</td>
<td>-</td>
</tr>
<tr>
<td>UBPRE120</td>
<td>Títulos de Deuda del Tesoro y de las Agencias Públicas de los EEUU</td>
<td>-</td>
</tr>
<tr>
<td>UBPRK437</td>
<td>Depósitos por encima del límite asegurado</td>
<td>+</td>
</tr>
<tr>
<td>UBPRE123</td>
<td>Activos Productivos Totales</td>
<td>-</td>
</tr>
<tr>
<td>UBPRD665</td>
<td>Préstamos a particulares</td>
<td>-</td>
</tr>
</tbody>
</table>

Como prueba adicional de robustez, de cara a la predicción más a corto plazo, se ha realizado una nueva estimación correspondiente a la totalidad de la muestra restringido al último trimestre de 2007, último trimestre de cierre previo al estallido de la crisis financiera internacional. Los resultados se presentan a continuación, Tabla 7:

Tabla 7. Modelo estimado a corto

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeficiente</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.366619</td>
<td>0.629138</td>
<td>0.582732</td>
<td>0.5601</td>
</tr>
<tr>
<td>UBPR1410/UBPR2170</td>
<td>3.850066</td>
<td>0.272620</td>
<td>14.12245</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPR3210/UBPR2170</td>
<td>-8.265096</td>
<td>1.091169</td>
<td>-7.574533</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPRE120/UBPR2170</td>
<td>-2.521038</td>
<td>0.471793</td>
<td>-5.343525</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPRK437/UBPR2170</td>
<td>5.309324</td>
<td>0.297031</td>
<td>17.87468</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPRE123/UBPR2170</td>
<td>-5.027063</td>
<td>0.698827</td>
<td>-7.193569</td>
<td>0.0000</td>
</tr>
<tr>
<td>UBPRD665/UBPR2170</td>
<td>-20.01870</td>
<td>1.471075</td>
<td>-13.60821</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
McFadden R-squared 0.186247 Akaike info criterion 0.305772
Schwarz criterion 0.307834 Hannan-Quinn criter. 0.306436
LR statistic 1954.374 Prob(LR statistic) 0.000000

| Obs with Dep=0 | 26676 | Total obs | 27972 |
| Obs with Dep=1 | 1296  |           |       |

<table>
<thead>
<tr>
<th>Estimated Equation Success cutoff: C = 0.034</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dep=0</td>
</tr>
<tr>
<td>% Correct</td>
</tr>
<tr>
<td>% Incorrect</td>
</tr>
</tbody>
</table>

Al comparar entre sí ambos resultados (largo plazo frente a corte plazo), se observa una mejora en las predicciones correctas de las entidades quebradas al pasar del 83.71%, modelo a largo plazo, al 85.73%, en el modelo a corto plazo.

De los resultados obtenidos en las tablas 3 a 7 podemos identificar que variables tienen una mayor influencia en la predicción de las quiebras:

1. **UBPR1410/UBPR2170: Préstamos garantizados por activos inmobiliarios / Activos Totales.** Variable significativa y con signo positivo. La principal justificación radica en el efecto perjudicial de una desproporcionada contribución de los préstamos hipotecarios en las carteras crediticias de los pequeños bancos regionales de la franja del ‘Sun Belt’, al acentuarse los niveles de morosidad tras el estallido de la crisis ‘subprime’, como consecuencia de la pérdida de valor patrimonial de los activos inmobiliarios, y del debilitamiento de la actividad económica, intensificándose las tasas de desempleo. Por tanto, en este contexto un incremento de la variable referida aumentaría la probabilidad de quiebra de una entidad bancaria.

2. **UBPR3210/UBPR2170: Capital Bancario Total / Activos totales.** Variable significativa y con signo negativo. Cumple con el resultado esperado, puesto que unos mayores niveles de capital suponen un mayor colchón de protección último contra las pérdidas no esperadas.
De forma acorde, un incremento de dicho ratio supondría una menor probabilidad de quiebra.

3. **UBPRE120/UBPR2170**: Títulos de Deuda del Tesoro y Ag. Públicas EE.UU. / Activos totales. Variable significativa y con signo negativo. De forma consistente con su signo esperado, porque este tipo de instrumentos financieros refuerzan los ratios de liquidez, al tiempo que permiten diversificar la gestión del riesgo por el nivel de solvencia del contraparte (garantizados por el Gobierno de los EE.UU.). Interpretación: menor probabilidad de quiebra.

4. **UBPRK437/UBPR2170**: Depósitos por encima del límite asegurado / Activos totales. Variable significativa y con signo positivo. Un exceso en su contribución al ‘mix’ de depósitos supone un menor grado de estabilidad en la financiación bancaria, al tiempo que un mayor coste de financiación, con respecto a los depósitos minoristas (por debajo de 100.000 dólares). De esta forma, se trata de otra variable que evoluciona en la misma dirección que la probabilidad de quiebra.

5. **UBPRE123/UBPR2170**: Activos productivos totales / Activos totales. Variable significativa y con signo negativo. Al contrario que los activos no productivos (NPAs), una positiva evolución de los activos productivos incrementan el nivel de beneficios de los bancos, los niveles de capital, y las posibilidades de crecimiento del préstamo, disminuyendo las probabilidades de quiebra.

6. **UBPRD665/UBPR2170**: Préstamos a particulares / Activos totales. Variable significativa y con signo negativo. Al contrario de lo que sucede con un exceso de préstamos hipotecarios sobre el total de la cartera crediticia, una significativa contribución de la cartera de
préstamos con respecto al balance total es favorable, por reflejar un modelo de negocio minorista y conservador, frente a las entidades más centradas en la gestión de las carteras de negociación, incluyendo instrumentos derivados.

I.5. CONCLUSIONES

El modelo propuesto sigue un enfoque metodológico novedoso en tres fases: 1) Depuración de la base de datos integral UBPR del FFIEC mediante análisis de frecuencias y eliminación de variables con excesivo porcentaje de observaciones perdidas, 2) Establecimiento de una secuencia de filtrado progresivo para construir una matriz balanceada, eliminando truncamiento por la derecha y por la izquierda y restringiendo nuevamente las variables a incluir con base en su mayor número de observaciones, y 3) Construcción de especificaciones siguiendo el criterio econométrico tradicional ‘bottom-up’, que conlleva las restricciones adicionales de incluir únicamente las variables estadísticamente significativas a nivel individual que presenten signos esperados correctos.

Conceptualmente, la especificación sugerida se alinea con el reciente paradigma de elección óptima bajo entornos de incertidumbre, enfoque seguido por Knight, Herbert, Simon, Kahneman, Amost Tversky, Hansen- Sargent, Kirman, y Gigerenzer entre otros autores, que defiende el empleo de modelos menos complejos y ratios heurísticos como herramienta óptima para un análisis efectivo de entornos interconectados y complejos, en contraposición con el enfoque de expectativas racionales y el marco moderno de macroeconomía y finanzas, dominado por modelos de toma de decisiones bajo entornos de riesgo de Arrow-Debreu y Merton-Markowitz.
Como se ha comentado, la robustez del modelo se constata mediante la comparación entre las estimaciones del panel de datos del período completo, frente a un panel más reducido que únicamente incorpora los datos correspondientes al 2007, observándose una disminución en los errores Tipo I del modelo a corto plazo (incremento de las predicciones correctas de quiebras del 83,71% al 85,73%).

En consecuencia, cabe reseñar la efectividad del modelo expuesto de cara a su utilización como sistema de alerta temprana de medio-largo plazo, al mostrar una satisfactoria capacidad de predicción de la práctica totalidad de las quiebras bancarias en los EE.UU. acumuladas durante la reciente crisis bancaria (cuatro ejercicios completos, hasta el último trimestre de 2012). Nótese que en general los EWS alternativos pierden notablemente su efectividad cuando el período de predicción es superior a los 24 meses.

En resumen, la especificación presentada supone una alternativa válida en el contexto de la modelización supervisora, como complemento a los existentes sistemas de predicción basados en datos financieros-contables, por mostrar una elevada capacidad de predicción en comparación con modelos econométricos (Ej. SCOR), y en relación con indicadores más sencillos (Ej. Ratio de Texas). La metodología es novedosa por explotar la base de datos FFIEC en su totalidad, y por realizar un riguroso proceso de filtrado en búsqueda de un grupo reducido de variables relevantes, para la producción de predicciones efectivas en el medio-largo plazo. En este sentido cabe señalar que el modelo SCOR del FDIC emplea 13 variables para lograr un porcentaje de error Tipo I similar. En consecuencia, menos es más.
I.6. REFERENCIAS BIBLIOGRÁFICAS


VAN DER PLOEG, S. 2010. Bank default prediction models: A comparison and an application to credit rating transitions. Erasmus University Rotterdam.


### I.7. ANEXO: Definiciones de las variables de la Ecuación 2

#### Grupo Balance (21 variables):

<table>
<thead>
<tr>
<th>Variable</th>
<th>Concepto</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBPR0071:</td>
<td>Saldos con interés adeudados por entidades depositarias</td>
</tr>
<tr>
<td>UBPR0081:</td>
<td>Saldos sin interés y en efectivo</td>
</tr>
<tr>
<td>UBPR1410:</td>
<td>Préstamos garantizados por activos inmobiliarios</td>
</tr>
<tr>
<td>UBPR1590:</td>
<td>Préstamos para la financiación de la producción agrícola</td>
</tr>
<tr>
<td>UBPR2145:</td>
<td>Edificios y Activos Fijos (incluyendo leasing capitalizable)</td>
</tr>
<tr>
<td>UBPR2170:</td>
<td>Activos totales</td>
</tr>
<tr>
<td>UBPR2200:</td>
<td>Depósitos totales</td>
</tr>
<tr>
<td>UBPR3123:</td>
<td>Provisión para pérdidas por préstamos y leasing</td>
</tr>
<tr>
<td>UBPR3300:</td>
<td>Deudas Totales, Acciones Preferentes no perpetuas y Capital Total</td>
</tr>
<tr>
<td>UBPR6164:</td>
<td>Importe agregado de la financiación concedida a directores, Ejecutivos y accionistas, e intereses vinculados</td>
</tr>
<tr>
<td>UBPR8528:</td>
<td>Préstamos y leasing, netos de ganancias no realizadas</td>
</tr>
<tr>
<td>UBPRD662:</td>
<td>Importe Total de Deudas</td>
</tr>
<tr>
<td>UBPRD665:</td>
<td>Préstamos a particulares</td>
</tr>
<tr>
<td>UBPRE116:</td>
<td>Préstamos a Empresas</td>
</tr>
<tr>
<td>UBPRE119:</td>
<td>Préstamos y leasing netos</td>
</tr>
<tr>
<td>UBPRE120:</td>
<td>Títulos de deuda del Tesoro y Agencias Públicas de los EE.UU.</td>
</tr>
<tr>
<td>UBPRE123 :</td>
<td>Activos productivos totales</td>
</tr>
<tr>
<td>UBPRE129:</td>
<td>Aceptaciones y otras deudas</td>
</tr>
<tr>
<td>UBPRG105:</td>
<td>Capital total incluyendo intereses minoritarios en filiales consolidadas</td>
</tr>
<tr>
<td>UBPRK426:</td>
<td>Depósitos por debajo del límite asegurado</td>
</tr>
<tr>
<td>UBPRK437:</td>
<td>Depósitos por encima del límite asegurado</td>
</tr>
</tbody>
</table>

#### Grupo Capital (4 variables):

<table>
<thead>
<tr>
<th>Variable</th>
<th>Concepto</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBPR3210:</td>
<td>Capital bancario total</td>
</tr>
<tr>
<td>UBPR3230:</td>
<td>Capital ordinario</td>
</tr>
<tr>
<td>UBPR3839:</td>
<td>Beneficios</td>
</tr>
<tr>
<td>UBPR530:</td>
<td>Otro Resultado Global Acumulado</td>
</tr>
</tbody>
</table>

#### Grupo Morosidad (2 variables):

<table>
<thead>
<tr>
<th>Variable</th>
<th>Concepto</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBPR1400:</td>
<td>Importe bruto de préstamos y leasing</td>
</tr>
<tr>
<td>UBPRD668:</td>
<td>Importe de préstamos y leasing impagados 30-89 días</td>
</tr>
</tbody>
</table>

#### Grupo Resultados (3 variables):

<table>
<thead>
<tr>
<th>Variable</th>
<th>Concepto</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBPR4340:</td>
<td>Resultado neto (pérdida)</td>
</tr>
<tr>
<td>UBPRE037:</td>
<td>Gastos generales totales</td>
</tr>
<tr>
<td>UBPRE625:</td>
<td>Dividendos / Cambios en el capital total</td>
</tr>
</tbody>
</table>

#### Grupo Provisiones (1 variable):

<table>
<thead>
<tr>
<th>Variable</th>
<th>Concepto</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBPRC233:</td>
<td>Ajustes</td>
</tr>
</tbody>
</table>
Capítulo 2

BANK FAILURE PREDICTION MODELS: LESS IS MORE?
II.0. Abstract

This paper analyzes the underlying factors behind the bankruptcies of U.S. banks during the recent financial crisis, to identify to what extent the accounting structures of the failed banks show significant differences with respect to the average profile of the industry. We apply two approaches, EWS standard CAMELS methodology and a heuristic indicator (Texas Ratio), to test the models of decision-making under risk \(^{11}\) versus the paradigm of optimal choice under uncertainty \(^{12}\).

Keywords: banking, finance, bank failures, early warning systems, crisis, Texas ratio, EWS, CAMELS, bankruptcy, logit.

\(^{11}\) (Arrow and Debreu 1954), (Markowitz 1952), and (Modigliani and Miller 1958)

\(^{12}\) (Knight 1921), (Simon 1956), (Khaneman and Tversky 1974), (Hansen and Sargent 2007), (Kirman 2010), (Gigerenzer 2007)
II.1 Introduction

The current study develops a dual approach in order to analyze the main determinant factors behind the last wave of bank failures in the United States, during the 2008-2014 period. Our key objective is to detect to what extent individual balance sheets of failed banks show differential elements with respect to the average profile of the industry, following two methodologies to identify the key drivers which may help to explain the failure of their business plans: (i) CAMELS logit model, built as a ‘proxie’ for the standard econometric modeling techniques implemented by most supervisory agencies, versus (ii) several versions of the Texas Ratio, introduced as ‘naive’ quotients developed as early warning systems (EWS) by the private sector.

The term *bank failure* -applied hereafter- follows the definition of the Federal Deposit Insurance Corporation (FDIC)\(^\text{13}\), which refers to ‘the closing of a bank by a federal or state banking regulatory agency, general when it is unable to meet its obligations to depositors and others. Nevertheless, although this issue remains outside of the scope of this paper, empirical studies reflect the lack of consensus on what constitutes ‘failure’. Hence, it can be worth mentioning the existence of alternative definitions of the event of bank failure in the most relevant literature about this subject, highlighting the following authors:

- (Beaver, 1966). Failure is defined as the inability of a firm to pay its financial obligations as they mature. Under operational terms, a firm fails when after any of these events: (i) bankruptcy, (ii) bond default,

\(^{13}\) [https://www.fdic.gov/consumers/banking/facts/](https://www.fdic.gov/consumers/banking/facts/). The creation of the Federal Deposit Insurance Corporation as an independent federal agency (1933) responded to the existing need in the U.S. to promote a high level of public confidence and financial stability in the nation’s banking system. In the event of a bank failure, -as the bank deposits insurer- the FDIC pays to the depositors under the established limits ($250,000 per depositor and insured bank), and -as the receiver- assumes the assets’ selling/collecting, and debts settling of the failed banks.
(iii), an overdrawn bank account, or (iv) non-payment of a preferred stock dividend.

(Ohlson, 1980). His definition of failure is purely legalistic: the failed firms in his sample must have filed for bankruptcy in the sense of Chapter X, Chapter XI, or some other notification indicating bankruptcy proceedings.

(Wheelock & Wilson, 2000). They apply two meanings: (i) failed banks are those closed by the FDIC (typically, capital/assets ratio equal to zero), and (ii) expanded definition, also including banks with total equity capital less goodwill, divided by total assets of less than two percent.

(Arena, 2005) Arena (2005). In this paper a financial institution has failed if it fits into any of the following categories: (i) recapitalization or liquidity injection by a central bank or public agency, (ii) temporary suspension of operations by the government, (iii) government closing, or (iv) absorption / acquisition by another financial institution.

(Van der Ploeg, 2010). Five events are defined in his study as ‘default events’: (i) bankruptcy filing, (ii) becoming in receivership, (iii), Federal Deposit Insurance Company (FDIC) closing, and (iv) a credit rating transition to a default rating class.
With respect to the perimeter of analysis, the database has been restricted to the subsample of U.S. regulated institutions, comprising all national or state chartered banks within the United States 14.

The statistical source used for this empirical investigation is the official database of the Uniform Bank Performance Report (UBPR) of the Federal Financial Institutions Examination Council (FFIEC), which brings together public information derived from the quarterly filings reported by individual banks insured by the U.S. deposit insurance fund (DIF).

Specifically, this report explores two different approaches to assess the degree of bank vulnerability: it compares several specifications of standard CAMELS models, consisting of different selections of accounting ratios among those which have shown previous statistical significance in the most relevant models in the literature of bank failures, with two variations of the Texas Ratio, the original quotient and a modified (adjusted) version.

The selection process of the explanatory variables is based on a ‘bottom-up & top-down’ methodology, starting with the total number of variables (65) included in the CAMELS indicators banking database, managed by SNL Financial 15 (See Table 2, Annex I).

The remainder of the paper is organized as follows: Section 2 presents a quick summary of the literature on bank failures, broken down into two main streams: (i) empirical work based on quantitative / econometric models, and (ii) The Texas ratio. Section 3 unfolds the methodological framework applied to perform the

---

14 A list of failed institutions can be found at Table 1, Annex I. A permanently updated list of failed banks is publicly available at: [https://www.fdic.gov/bank/individual/failed/banklist.html](https://www.fdic.gov/bank/individual/failed/banklist.html)

15 [www.snl.com](http://www.snl.com). SNL Financial is one of the leading sources in the United States for collection, standardization and dissemination of financial information.
counterfactual analysis, and it also lays out the specification of the different models. The main findings regarding the relative performance of these two approaches are shown in Section 4, and Section 5 concludes with some policy suggestions.

II.2 Literature

Turning to the background literature on bank failures, our analysis follows a breakdown between a well-established and orthodox approach, represented by several specifications of econometric models following the standard “CAMELS” methodology, and a more holistic approach, based on tracking a well-known naive indicator such as the popular “Texas Ratio”.

With respect to quantitative models, the earliest studies analyzing the determinant factors of bank failures arose as a result of the financial crisis of the Great Depression, at the end of the decade of the 20s and early 30s (Spahr, 1932).

The crisis of the U.S. industrial sector in the late 60s has resulted in a renewed interest in this type of research, pioneering models of univariate discriminant analysis (Beaver, 1966) and multivariate discriminant analysis (Altman E., 1968), in order to assess the degree of viability of firms through the analysis of financial and accounting ratios.

This methodology was transferred to the analysis of individual banks in the early 1970s, including specific accounting ratios which characterized the banking activity (Altman, Haldeman, R.G., & Narayanan, P., 1977), (Meyer & Pifer, 1970), (Sinkey, 1975) and (Stuhr, D.P. & Van Wicklen, R., 1974).
In the late 1970s, after finding that exigent (constraint) hypothesis inherent to discriminant models were frequently violated by real financial data (i.e. normal distribution functions in financial-accounting items, or equal variance-covariance matrices between failed v surviving banks), Martin introduced logistic regression, due to its lack of restrictive assumptions with regard to the distributions of the explanatory variables.

For the first time, (Martin, 1977) examined the system of commercial banks members of the Federal Reserve System, using a sample of 23 bankruptcies and 5,575 active banks. In order to achieve this, he introduced 24 ratios from financial statements, dividing them into 4 groups, according to the nature of the information: (C) Capital adequacy, (A) asset quality, (E) Earnings, and (L) Liquidity. The resulting model included four variables relating to capital (C), asset quality (A), and profits (E). Subsequently, other models obtained results consistent with those of Martin (Andersen, 2008), (Avery & Hanweck, G.A., 1984), (Barth, Brumbaugh, R.D.Jr., Sauerhaft, D., & Wang, G.H.K., 1985), (Estrella, Park, S., & Peristiani, S., 2000), (Kolari, Glennon, Shin, & Caputo, 2000), (Thomson, 1991), (Thomson, J.B., 1992), (Wheelock & Wilson, 2000).

The model built by (Hanweck & Avery, R.B., 1984) analyzed 100 failed banks against the background of a surviving sample of 1,190 institutions, during the period 1978-1983, including 9 significant variables from the previous literature, and resulting in five significant explanatory variables with correct signs. Consistently with (Martin, 1977), those variables were classified into capital, asset quality and earnings. (Barth, Brumbaugh, R.D.Jr., Sauerhaft, D., & Wang, G.H.K., 1985) confirmed the abovementioned factors: throughout their model of U.S. thrift institutions over the period 1981-1984, they introduced liquidity as a key additional explanatory factor, by using the ratio (liquid assets / total assets), and the natural logarithm of total assets (Barth, Brumbaugh, R.D.Jr., Sauerhaft,
D., & Wang, G.H.K., 1985). (Thomson, 1991) analyzed failures during the period 1984-1989, achieving a significant improvement in the structure of the previous logit models due to considering the contribution of quality management (M) as an additional explanatory variable, based on the analysis from (Graham & Horner, J.E., 1988), (Thomson, 1991). (Estrella, Park, S., & Peristiani, S., 2000) developed a model to compare the degree of efficiency of several capital ratios –referred to total and risk-weighted assets-, which resulted in a similar predictability power –at least equal- by simple versions of capital ratios. Moreover, several econometric models have assessed the performance of credit ratings, obtaining mixed results. (Andersen, 2008) has developed an analysis of failures in the Norwegian banking system, using a logit model.

All in all, the most relevant logistic models of bank failures during the last three decades have incorporated a number of financial and accounting ratios ranging from nine variables (Avery & Hanweck, G.A., 1984) to twenty seven variables (Andersen, 2008), with the result of a very similar number of significant explanatory variables –with correct signs-, ranging between four (Martin, 1977) and seven (Avery & Hanweck, G.A., 1984). Table 1 summarizes all significant variables.

In recent decades EWS (Early Warning Systems) have been developed and refined, sustaining the general purpose of an early detection of bank failures, in order to avoid–or at least mitigate- their associated costs (Oet, Eiben, R., Bianco, T., Gramlich, D., Ong, S.J., & Wang, J., 2011), (Putnam, 1983), (Rose & Kolari, J.W., 1985), (Wheelock & Wilson, 2000). Besides, there has been a gradual development of risk models (Cox) applied to the banking sector (Lane, Looney, S.W., & Wansley, J.W., 1986), (Wheelock & Wilson, 2000), while at the same time the application of neural networks’ methodologies to banking failures has
developed since the start of the 1990s (Odom & Sharda, R., 1990), (Tam & Kiang, M., 1992).

Table 5. Significant variables of logit models applied to bank failures

<table>
<thead>
<tr>
<th>Study</th>
<th>Model</th>
<th>Ratio (*)</th>
<th>Capital quality</th>
<th>Asset quality</th>
<th>Management</th>
<th>Earnings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin (1977)</td>
<td>Logit</td>
<td>25/4</td>
<td>GCARA</td>
<td>GCONI</td>
<td>CI2LN</td>
<td>NITA</td>
</tr>
<tr>
<td>Avery &amp; Hanweck (1984)</td>
<td>Logit</td>
<td>9/7</td>
<td>KTA</td>
<td>NLTA</td>
<td>LNTA</td>
<td>NITA</td>
</tr>
<tr>
<td>Barhy y otros (1985)</td>
<td>Logit</td>
<td>12/5</td>
<td>NWTA</td>
<td>ISFTF</td>
<td>LNTA</td>
<td>LATA</td>
</tr>
<tr>
<td>Thomson (1991)</td>
<td>Logit</td>
<td>16/7</td>
<td>NCAPTA</td>
<td>NCOTA</td>
<td>OVRHDTA</td>
<td>ROA</td>
</tr>
<tr>
<td>Andersen (2008)</td>
<td>Logit</td>
<td>27/6</td>
<td>CAR</td>
<td>RMGL</td>
<td>ELOSS</td>
<td>ROA</td>
</tr>
</tbody>
</table>

Source: (van der Ploeg, 2010)

In parallel, an alternative strand of the literature about bank failures has focused on the popularization of simpler (‘naive’) models, such as the Texas Ratio,

---

developed by (Cassidy, 1980s) at RBC Capital Markets. Most of the great success of this methodology responds to its computational simplicity, and its great degree of success to detect the proximity of an event of bank failure (Jesswein, 2009).

The key objective underpinning this work is to accomplish an assessment of the degree of robustness associated to both methodologies: logit econometric CAMELS models versus naive Texas ratios, in terms of their relative effectiveness to explain the determinants of bank failures. We will try to test if both techniques are complementary analytical tools.

II.2.1 EWS implementation by U.S. Banking Supervisory Regulators

CAMELS ratings were originally developed in the United States, to enhance the overall assessment of the condition of their banking system. These scores are assigned based on two main components: a) remote (off-site) monitoring through a ratio analysis of the financial statements, and b) on-site examinations (and follow-up), performed by a designated supervisory regulator. Along the rest of this subsection we will review the most widely used EWS by the U.S. federal banking regulators.

II.2.1.1 FDIC model¹⁷ ¹⁸

FDIC started to develop early warning system in the mid 80s.

Their current model is SCOR (Statistical CAMELS Off-site Rating). This model is composed of quarterly data from bank’s financial statements, and its main purpose is to identify the group of banks which could experience a downgrade.


during the next on-site examination, using an ordered logit analysis with a ‘step-wise’ approach. According to (Collier, Forbush, S., Nuxoll, D., & O'Keefe, J., 2003) this model has played a relevant role in the remote supervisory process, in terms of recourse allocation, and due to its monitoring power of industry trends. SCOR model is implemented to estimate the probability of each bank to be classified 1-5 within a compounded rating (1 being the best, and 5 the worst), and also to estimate the probability of a downgrading. The variables are:19

1. Equity
2. Loan-loss reserves
3. Loans past due 30-89 days
4. Loans past due 90+ days
5. Nonaccrual loans
6. Other real estate
7. Charge-offs
8. Provisions for loan losses and transfer risk
9. Income before taxes and extraordinary charges
10. Volatile liabilities
11. Liquid assets
12. Loans and long-term securities

All variables are ratios calculated as percentages of total assets. The calibration of the model is performed to signal banks with likelihood greater than 35% of being assigned a downgrade. The tests of goodness-of-fit (Collier, Forbush, S., Nuxoll, 19 The earlier SCOR specification used 13 variables, including dividends and showing other modifications, and being slightly less accurate than this more updated specification.)
D., & O'Keefe, J., 2003) indicate that the accuracy of this model has decreased, especially since 1993. Since this year SCOR had identified roughly 16% of banks downgrades (Type I accuracy), and approximately 27 percent of the banks not downgraded (Type II accuracy). SCOR replaced the earlier model (CAEL), due to higher levels of Type I and Type II accuracy during almost all periods.

II.2.1.2 Federal Reserve model 20 21

During the 90s there was a transition from a ‘model-based’ supervisory process, based on monitoring key financial ratios (‘screening’), towards an introduction of econometric models designed to forecast bank’s financial condition. Against that background, the current SEER (System to Estimate Examination Ratings) was developed.

The SEER model consists of two econometric models:

- **SEER ratings model.** Similar to SCOR, this system forecasts the likeliness of assigning one of the five ratings, applying a multinomial logit step-by-step analysis. The SEER rating is calculated as the sum of five grade levels, each one multiplied by their respective probabilities. Additionally, the model is often used to classify banks into satisfactory (grades 1-2) and unsatisfactory institutions (grades 3,4,5).

The model includes accounting and examination data from the two previous quarters, including approximately 45 financial and non-

---

http://www.bis.org/publ/bcbs_wp04.pdf  

financial variables. Through a step-by-step selection procedure, variables are incorporated into the final model each quarter. The variables that have consistently sustained their statistically significance in each of the quarters, are used as proxies for capital adequacy, asset quality, the level of profits, and the level of liquidity, which are the variables typically used by supervisors to determine the banks financial condition. Moreover, SEER model includes the previous CAMEL rating. Macroeconomic variables such as employment, GDP or income per capita are usually not statistically significant.

SEER risk rank model. Its key purpose is to estimate the likelihood of a bank becoming severely undercapitalized over the next two years (according to ‘PCA standards’, i.e. immediate corrective actions implemented by the Federal Reserve).

This model has been re-estimated quarterly, running a step-by-step probit regression based on data obtained from failed and undercapitalized banks during the most recent years. However, as the pace of bankruptcies felt considerably, the Fed maintained its panel model covering the period 1985-1991, in order to capture a sufficient number of financial institutions to estimate the model. As far as we know, this model was not being reestimated during the immediate previous years before the recent international financial crisis.

Both SEER models seem to show an acceptable level of performance. According to the original test validation, (Cole & Gunther, J.W., 1995), based on ten separate quarterly estimates using data from December 1989 to March 1992, SEER ratings
model showed a high degree of fit predicting banks with scores 1 and 2: 77.5% of banks rated 1, and 79.9% of banks rated 2 were correctly estimated in its forecast model.

Additionally, of the 262 failed banks within the sample, 97.7% received a SEER rating of 5; 1.9% received a rating of 4, and 0.4% received a rating of 3 (None of the banks were assigned a SEER rating of 1 or 2). Moreover, these models have been annually revalidated, with a significant number of enhancements with respect to the original model, back in 1995.

II.2.1.3 OCC model

The ‘Canary Sistem’ is the OCC’s monitoring methodology. It includes a set of models designed for economic and supervisory forecasting, organized into four main components:

- Benchmark
- Credit Scope
- Market Barometers
- Predictive models

As part of its global remote monitoring system, the Office of the Comptroller of the Currency applies this model in combination with FDIC SCOR abovementioned model.

Moreover, PGRM (Peer Group Risk Models) are a set of econometric models devised to measure the potential impact of different economic scenarios over ROAA (return over average banking assets) during the next three years, checking 11 sets of institutions with similar level of total assets, assorted in terms of their
loan portfolios, and differentiating newly created banks. This model includes in its specification several macroeconomic indicators (i.e. interest rates, wages, unemployment, GDP, bankruptcies) and a group of banking and financial variables (i.e. deteriorated loans, provisions for loan losses, and capital ratios). Additionally, several variables of specific segments are incorporated into the model, such as agricultural income and prices for institutions specialized in the agricultural segment. We haven’t found any information about the level of performance of this model.

II.2.1.4 EWS literature review

* (Whalen, 1991). This research analyzes a Cox proportional hazards model, which estimates the probability of certain institutions— with specific financial profiles— to survive beyond a previously set period. This model pulls apart failed and surviving banks with a high degree of accuracy. The author concludes that these sort of studies may be developed with relatively high efficiency and low costs.

* (Thomson, 1991). His study performs econometric models of bank failures of any size, performing logit analysis over public quarterly accounting and financial bank statements. The probability of failure is estimated as a function of capital adequacy, asset quality, management quality, earnings and the relative degree of liquidity of the assets portfolio (CAMELS proxy variables. Thomson found that these kinds of variables actually help explain bank failures up to four years in advance.

* (Cole & Gunther, J.W., 1995). They built a probit model to estimate the probability of bank failure during the next two years, applying a set of current financial ratios. Conclusion: The information contained in CAMELS ratings from on-site examination processes losses it effectiveness relatively quickly, showing
predicting power to identify bank failures over a remote EWS system model when data comes from inspections performed during the two previous quarters with respect to the forecast period. It also notes that the period of effectiveness from a EWS off-site model strongly depends on the reliability of accounting data, which clearly improves as a result of periodic inspections.

* (Gunther & Moore, R., 2000). They have revealed clear evidences of the close connection between the ‘revised Call Report data’ and bank examinations, which suggests an overvaluation of the quality of EWS models using real-time accounting. However, these kinds of data sustain their validity, in the lack of empirical evidences of a substantial impact from data revisions over the prediction power of EWS models.

* (Gilbert, Meyer, A., & Vaughan, M., 1999). They compare the power of failure forecasting between a group of univariate and multivariate models, concluding that the individual variable which best helps to explain the determinants of bankruptcy varies annually, and only univariate models are able to achieve consistently accurate predictions over time. One of the most significant variables in this test was the capital ratio.

* (Krainer & López, J., 2003). Their analysis was focused on the incorporation of market data in EWS, finding that share price changes tend to forecast ratings changes in their sample of bank holdings, at least nine months in advance, concluding that kind of information is useful to monitor.

II.2.2. Texas Ratio

This indicator is attributed to (Cassidy, 1980s) and his colleagues at RBC Capital Markets 22. In a nutshell, it compares the ratio of bad assets –problematic loans-

at a bank, with respect to its level of available capital to deal with this issue. A high Texas Ratio does not necessarily imply failure, but the ratio has shown a high degree of accuracy to assess bank’s ability to absorb future losses.

This measure was developed for an analysis of Texas banks during the recession of the early 1980s, during which Gerard Cassidy observed a solid pattern of bankruptcy linked to banks showing values for this ratio equal or greater than 1:1 (100%).

Calculation of the classic version of the Texas Ratio:

\[
\text{[Value of the lender’s NPAs (NPL + OREO) / (TCE + LLR)}
\]

Whereas:

NPAs: Nonperforming assets (Nonperforming Loans + Real Estate Owned)

TCE: Tangible Common Equity.

LLR: Loan Loss Reserves.

Until the recession of the 90s, many banks in Texas were considered some of the most solvent in the U.S., but that situation drastically turned around when the level of impaired assets soared. Hence, the mentioned analyst noted than when the problematic assets raised over 100% of the capital buffer, most of Texas banks ended up in bankruptcy filings (Annex 1 and 2). He also identified a

---

23 However, it should be noted that according to more recent data -extracted from the recent international financial crisis-, there are strong reasons to rename that ratio into the ‘Georgia Ratio’, considering the absolute number of failures in that state since 2008 (88), above Florida (71), Illinois (61) and California (40), as can be seen in Annex 1 and Annex 3, and also taking into account the percentage of banks with TR over 100. Conversely, Texas has more than 550 banks (more than any other state in the U.S. with less than 0.75% of this sample at risk of failure, according to this metric.

24 Source: SNL & Alip Artates. It includes all operating banks and thrifts that filed call reports for 2Q 2014. Adjusted Texas Ratio = NPA + Loans +90 days past due, net of delinquent government guaranteed loans and OREO covered by loss-sharing agreements with the FDIC, as % of tangible equity and loan loss reserves.
similar behavior in the New England banking sector during the 1990s recession. Cassidy is currently an analyst of RBC Capital Markets, and he is still monitoring the Texas Ratio: at the end of the 3rd quarter of 2009, 388 U.S. banks showed Texas Ratios above the threshold of 100%.

The recent international financial crisis (2008-2013) has not shown the same magnitude that the S&L (‘Savings and Loan’) crisis, in terms of the number of failures. Whereas in the 90s the FDIC ‘Troubled Watch List’ included more than 1,500 banks – almost 10% of a total of roughly 15,000 institutions, and subsequently falling more than 1,000 banks between 1988 and 1989, the ‘current’ crisis has resulted ‘just’ in 506 failed banks since January 2008, and only 17 of these failures have occurred during the current year (2014). However, current state requirements to open branches in another country has changed, meaning that the number of lenders – and the number of bank failures – is lower, but the size of those that have failed have been large, causing a detrimental effect on the U.S. economy.

According to official data released in the U.S., at the closing of the first semester of 2014 the number of banks with an adjusted Texas ratio above 100% has continue to decrease to 157, compared to 181 institutions in 1Q2014, with respect to an absolute peak during the recent financial crisis of 510 institutions in 4Q2010. The adjusted median TR (Texas Ratio) of the subsample of banks with a ratio above 100% has fallen to 155.38%, from a level of 156.14% in the previous quarter (Annex 3). This information is consistent with the steady reduction of the aggregate level of the Texas Ratio since 2011Q1, as reflected in Annex 4.

---


26 Source: SNL base on regulatory filings for commercial banks, saving banks and savings institutions. The number of institutions (red bars) represents number of banks and thrifts with adjusted Texas Ratio above 100%. The number of institutions failed (blue line) shows the number of banks and thrifts that have failed in each quarter. Adjusted Texas Ratio: NPA + Loans 90 days or more past due, net of delinquent
All in all, the indisputable success and popularity of the Texas Ratio has been founded on its high level of simplicity and availability, combined with a great degree of accuracy as EWS for bank failures (Annex 5 27 28). However, there is no ‘silver bullet’ in the field of forecasting bank failures, and this indicator is no exception to that rule due to several key limitations: (i) it doesn’t account for potential value of collateral, critical issue of PD (probability of default) and LGD (loss given default) analysis, (ii) it basically ignores potential overvalued assets in the balance sheet, and (iii) this ratio doesn’t incorporate information about the composition of the lending portfolio.

Against that background, this study tries to assess the analytical quality of this individual indicator –The Texas Ratio, hereafter TR- versus the traditional CAMELS methodology, as a starting point to accomplish a simple –but efficient- econometric specification to explain bank failures, which will be done throughout the next section.

II.3. Counterfactual analysis

II.3.1. Database management

The starting point of our study has been to analyze the aggregate database of the U.S. banking system, composed by 12,498 individual records containing accounting and financial data from all U.S. banking institutions regulated by three

---

27 Using FDIC data and based on the Texas Ratio and a Tier 1 leverage ratio, www.bankimplode.com released a list of problematic banks (>40% TR) in 3Q2008, identifying 82% of the failing banks during the following six months.

Federal Banking Agencies - FRB, FDIC and OCC respectively. Our database covers the period 2008-2014, and it can be divided into two subsamples according to the ‘state’ of each bank: more specifically, 546 failed banks and 11,952 surviving banks.

The second step of our analysis has been to extract all the institutions (records) with any missing data in each of the variables included in any of our models (CAMELS variables and three versions of the Texas Ratio: Traditional TR, Modified TR and Current TR), with the purpose of obtaining a balanced matrix with the same number of observations in all the variables of our database. This is a necessary condition to perform the logit regressions, because when there’s NA (not available) data the econometric software mixes up ‘integer’ with ‘chain’ (character) data.

After performing this progressive filtering process of NA data in all our variables, and subsequently deleting all records with a value of cero in any of the variables, this process resulted in a database comprised of a balanced matrix with 4,669 banking institutions, of which 368 were failed banks since 2008, without any missing observation in the matrix of explanatory variables (Note: additionally we have eliminated older failures, as they are outside the scope of this work).

Table 2. Balanced subsample (no missing data).

<table>
<thead>
<tr>
<th>Total failures:</th>
<th>368</th>
<th>7.88%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of failure:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>113</td>
<td></td>
</tr>
</tbody>
</table>
Then, our chosen temporal approach has been to focus on the 4th quarter of 2007, in an attempt to assess the predictive accuracy as a supervisory tool of the most immediate closing period before a banking crisis, considering that just 6 months later (September 15th, Lehman Brothers failure) the first wave of bankruptcies did take place.

**Table 3. Segmented Texas Ratios (Surviving v Failed banks / By year of failure)**

**AVERAGE TEXAS RATIO (2007Q4)**

(All bank failures during the aggregate period 2008-2014)

<table>
<thead>
<tr>
<th></th>
<th>Average Texas Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total banks</td>
<td>4669                12.29</td>
</tr>
<tr>
<td>Surviving banks</td>
<td>4301                   92.12%  10.91</td>
</tr>
<tr>
<td>Failed banks</td>
<td>368                    7.88%   28.43</td>
</tr>
<tr>
<td>Year of failure</td>
<td>Nº of failures</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>2008</td>
<td>15</td>
</tr>
<tr>
<td>2009</td>
<td>102</td>
</tr>
<tr>
<td>2010</td>
<td>113</td>
</tr>
<tr>
<td>2011</td>
<td>72</td>
</tr>
<tr>
<td>2012</td>
<td>38</td>
</tr>
<tr>
<td>2013</td>
<td>17</td>
</tr>
<tr>
<td>2014</td>
<td>11</td>
</tr>
</tbody>
</table>

Preliminary counterfactual analysis (after database refinement):

- The aggregate average Texas Ratio of all surviving banks in 2007 was clearly lower (10.91) with respect to the aggregate average of all failed banks during the period 2008-2014 (28.43).

- The sample is representative enough, as it contains a total of 4,669 institutions, 7.88% of which are 368 failures during the global period of the financial crisis (actually, 506 failures occurred -as explained in subsection 2.2-, but due to the database refinement we have lost 138 records relative to failed banks with missing data).
- Upon applying an EWS built from existing information at the end of 2007, it can be tested that the Texas Ratios associated to the most imminent bank failures (2008) were already showing average values over 70%.

- Moreover, after those events it can be checked (‘a posteriori’) that this Ratio gradually decreases when the distance to failure increases, which is consistent with the gradual process of deterioration observed in bank fundamentals. Thus, the subsequent 2009 failures at that time (4Q2007) were showing an average TR greater than 30%, to go down to a level slightly over 20% for the failures that occurred 5 years later (at the end of the crisis).

- Still, even the last group of failed banks from a financial crisis which ‘lasted’ more than 5 years, -using data from 4Q2007- already more than doubled the average TR of the subsample of surviving banks (21.81 versus 10.91).

**II.3.2 Econometric model specification**

First, we have implemented a classical bottom-up approach, starting by the estimation of individual logit models with each of the different explanatory variables, and building the different model specifications from here. Afterwards, we have developed a combination of top-down and bottom-up approach, following the mixed method advocated by Hansen (1996).

The key reasons for the selection of a ‘simple’ -not simplistic- model are:

a) To reduce model costs (i.e. to lower database building and processing costs),

23
b) To avoid introducing unrealistic –too complicated- assumptions, to help clarify the complex nature of nonexperimental –social- sciences,

c) To generate operating models: ‘relatively easy’ to understand, to set up, to communicate, and to recalibrate,

d) To mitigate data inconsistencies and high levels of error sensitivity, typically associated to sophisticated econometric models,

e) To facilitate the isolation and detection of the most relevant determinants of bank failures, because simple models are easier to analyze. Thus, these ‘naive’ models reduce the sources of blunders, subjectivity or oversights, as Keuzenkamp and McAleer (1995) have explained.

II.3.3 Variables selection

1. After the completion of the initial phase, during which we have performed individual tests of significance for all the explanatory variables (Annex 6). Based on the that previous results, we have selected several preliminary models whose specification only includes the variables pertaining to each of the components of the composite CAMELS ratio which have previously shown the highest levels of individual significance (t-student analysis) (Highlighted on Annex 6).

2. The second step of the estimation procedure comprises a comparison of these selected models, after performing a post-estimation analysis:

   a. Lroc. It computes an area under ROC curve, displaying this chart.

---

29 Stata SE 11.2 (Econometric software)
b. **Lsens.** This command calculates a table-graph of sensitivity and specificity levels, versus a selected probability cutoff (default level: 0.5).

### Note on the interpretation of EWS models applied to bank failures:

The level of accuracy of an early warning system based on a binary logistic function is strongly linked to its goodness of prediction fit, which—in turn— is closely associated to its degree of precision when assessing its binary classification performance, based on two main types of classification errors: Type I and Type II errors, respectively.  

In this context, the implementation of dichotomous methodologies into the field of bank failures its critically focused on achieving a reduced level of Type I error, also called ‘false positive rate, which equals [1 – Sensitivity]). The main reason for this ‘asymmetrical concern’ is the potential much higher cost derived from Type I error associated to an event of bankruptcy, which could have been avoided through earlier supervisory intervention. With regard to the costs relating to misclassifying as problematic an institution that is in fact solvent, in principle they should be limited to the resources ‘unnecessarily’ allocated to its previous monitoring.

### II.4 Findings

1. Econometric models based on CAMELS specifications have achieved slightly better results in our work, in terms of global explicability of the endogenous

---

30 The complementary event of **Type I error is Type I accuracy**, which is the percentage of actual failures that are identified in advance by the model (also called **true positive rate, recall rate or Sensitivity**). Analogously, the complementary event of **Type II error is Type II accuracy**, which is the percentage of actual surviving banks, that are identified in advance by the model (also called **true negative rate or Specificity**
variable -the event of banking failure-, when compared to ‘naive’ and less complex models, based on three analyzed versions of the Texas Ratio.

2. CAMELS models whose specification include the Texas Ratio -replacing capital adequacy and asset quality variables 31- show a clear improvement in their explanatory / forecasting power, tested by a decrease in % of Type I & Type II errors.

3. Through the selection of an optimal cutoff point on the model that includes just failures during 2009, based on the relationship between the number of failed banks (15) with respect to the total of banks in our sample (4316), our model accomplishes the maximum level of sensibility & specificity (lowest level of Type I & Type II errors).

4. It’s worth mentioning that this model overcomes –at least to the best of our information- not only the forecasting power of the Texas Ratio, but also the relating predicting capacity of the early warning systems applied by the federal supervisors in the United States, as our specification achieves –simultaneously- levels of sensitivity (true positive rate) and specificity (true negative rate) above 90%:

31 Those variables are already included in the denominator and the numerator of the TR, so we don’t include them twice in order to avoid multicollinearity).
```
.. logit failed_yn tr roatce ygsp_ass costfund

Iteration 8: log likelihood =  99.00416
Iteration 1: log likelihood =  63.163407
Iteration 2: log likelihood =  49.191313
Iteration 3: log likelihood =  44.120773
Iteration 4: log likelihood =  44.021626
Iteration 5: log likelihood =  44.026261
Iteration 6: log likelihood =  44.002609

Logistic regression
Number of obs =  4316
LR chi2(4) =  111.60
Prob > chi2 =  0.0000
Pseudo R2 =  0.0696

Log likelihood =  44.002609

+--------------+------------------+
<table>
<thead>
<tr>
<th>failed_yn</th>
<th>Coef.  Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>z  2</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
</tr>
<tr>
<td></td>
<td>tr</td>
</tr>
<tr>
<td></td>
<td>roatce</td>
</tr>
<tr>
<td></td>
<td>ygsp_ass</td>
</tr>
<tr>
<td></td>
<td>costfund</td>
</tr>
<tr>
<td></td>
<td>r_cons</td>
</tr>
</tbody>
</table>
+--------------+------------------+

.. estat classification, cutoff(0.50)

Logistic model for failed_yn

<table>
<thead>
<tr>
<th>Classified</th>
<th>True</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D</td>
<td>-D</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>14</td>
<td>349</td>
<td>363</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>3952</td>
<td>1</td>
<td>3953</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4066</td>
<td>350</td>
<td>4416</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Classified + it predicted Pr(DD) => 0.5048
True D defined as failed_yn != 0

Sensitivity    Pr(+|D) 99.33%
Specificity    Pr(-|D) 91.89%
Positive predictive value Pr(D +) 9.86%
Negative predictive value Pr(-|D) 99.99%

False + rate for true -0  Pr(+|D) 8.11%
False - rate for true D  Pr(-|D) 6.67%
False + rate for classified + Pr(-|D) 96.14%
False - rate for classified - Pr(D -) 0.53%

Correctly classified 91.89%
```
II.5. Conclusions and policy suggestions

The main policy recommendations arising from the positive results of this study are:

I. To reinforce the importance of monitoring and tracking approaches, techniques and ratios from the private industry, considering the incorporation of those measures and/or factors showing greater efficiency and robustness into the regulatory field,

II. To promote and encourage the development of simplified (not simple) quantitative models, to the extent that they are able to explain the main outlying drivers that influence the expectations of survival in the banking sector,

III. To fortify a dual supervisory approach, including remote monitoring systems to complement and strengthen on-site supervision, which is the most efficient tool to identify the degree of reliability of bank accounting and financial statements, and the only way to understand the processes of identification, assessment, control and mitigation of the main risks affecting banking institutions.
## II.6 Statistical Annexes

### ANNEX 1. GEOGRAPHICAL LOCATION OF U.S. BANK FAILURES (2008-2014)

<table>
<thead>
<tr>
<th>State</th>
<th>Failures</th>
<th>Percentage</th>
<th>Total Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>88</td>
<td>17.36%</td>
<td>17.36%</td>
</tr>
<tr>
<td>FL</td>
<td>71</td>
<td>14.00%</td>
<td>31.36%</td>
</tr>
<tr>
<td>IL</td>
<td>61</td>
<td>12.03%</td>
<td>43.39%</td>
</tr>
<tr>
<td>CA</td>
<td>40</td>
<td>7.89%</td>
<td>51.28%</td>
</tr>
<tr>
<td>MN</td>
<td>22</td>
<td>4.34%</td>
<td>55.62%</td>
</tr>
<tr>
<td>WA</td>
<td>18</td>
<td>3.55%</td>
<td>59.17%</td>
</tr>
<tr>
<td>MO</td>
<td>16</td>
<td>3.16%</td>
<td>62.33%</td>
</tr>
<tr>
<td>AZ</td>
<td>15</td>
<td>2.96%</td>
<td>65.29%</td>
</tr>
<tr>
<td>MI</td>
<td>13</td>
<td>2.56%</td>
<td>67.85%</td>
</tr>
<tr>
<td>NV</td>
<td>12</td>
<td>2.37%</td>
<td>70.22%</td>
</tr>
<tr>
<td>TX</td>
<td>11</td>
<td>2.17%</td>
<td>72.39%</td>
</tr>
<tr>
<td>MD</td>
<td>10</td>
<td>1.97%</td>
<td>74.36%</td>
</tr>
<tr>
<td>SC</td>
<td>10</td>
<td>1.97%</td>
<td>76.33%</td>
</tr>
<tr>
<td>CO</td>
<td>9</td>
<td>1.78%</td>
<td>78.11%</td>
</tr>
<tr>
<td>KS</td>
<td>9</td>
<td>1.78%</td>
<td>79.88%</td>
</tr>
<tr>
<td>WI</td>
<td>8</td>
<td>1.58%</td>
<td>81.46%</td>
</tr>
<tr>
<td>AL</td>
<td>7</td>
<td>1.38%</td>
<td>82.84%</td>
</tr>
<tr>
<td>NC</td>
<td>7</td>
<td>1.38%</td>
<td>84.22%</td>
</tr>
<tr>
<td>OK</td>
<td>7</td>
<td>1.38%</td>
<td>85.60%</td>
</tr>
<tr>
<td>PA</td>
<td>7</td>
<td>1.38%</td>
<td>86.98%</td>
</tr>
<tr>
<td>OR</td>
<td>6</td>
<td>1.18%</td>
<td>88.17%</td>
</tr>
<tr>
<td>UT</td>
<td>6</td>
<td>1.18%</td>
<td>89.35%</td>
</tr>
<tr>
<td>NJ</td>
<td>5</td>
<td>0.99%</td>
<td>90.34%</td>
</tr>
<tr>
<td>OH</td>
<td>5</td>
<td>0.99%</td>
<td>91.32%</td>
</tr>
<tr>
<td>TN</td>
<td>5</td>
<td>0.99%</td>
<td>92.31%</td>
</tr>
<tr>
<td>VA</td>
<td>5</td>
<td>0.99%</td>
<td>93.29%</td>
</tr>
<tr>
<td>NY</td>
<td>4</td>
<td>0.79%</td>
<td>94.08%</td>
</tr>
<tr>
<td>IN</td>
<td>3</td>
<td>0.59%</td>
<td>94.67%</td>
</tr>
<tr>
<td>NE</td>
<td>3</td>
<td>0.59%</td>
<td>95.27%</td>
</tr>
<tr>
<td>NM</td>
<td>3</td>
<td>0.59%</td>
<td>95.86%</td>
</tr>
<tr>
<td>PR</td>
<td>3</td>
<td>0.59%</td>
<td>96.45%</td>
</tr>
<tr>
<td>AR</td>
<td>2</td>
<td>0.39%</td>
<td>96.84%</td>
</tr>
<tr>
<td>IA</td>
<td>2</td>
<td>0.39%</td>
<td>97.24%</td>
</tr>
<tr>
<td>IA</td>
<td>2</td>
<td>0.39%</td>
<td>97.63%</td>
</tr>
<tr>
<td>ID</td>
<td>2</td>
<td>0.39%</td>
<td>98.03%</td>
</tr>
<tr>
<td>ID</td>
<td>2</td>
<td>0.39%</td>
<td>98.42%</td>
</tr>
<tr>
<td>LA</td>
<td>2</td>
<td>0.39%</td>
<td>98.82%</td>
</tr>
<tr>
<td>MS</td>
<td>2</td>
<td>0.39%</td>
<td>99.21%</td>
</tr>
<tr>
<td>CT</td>
<td>1</td>
<td>0.20%</td>
<td>99.41%</td>
</tr>
<tr>
<td>MA</td>
<td>1</td>
<td>0.20%</td>
<td>99.61%</td>
</tr>
<tr>
<td>SD</td>
<td>1</td>
<td>0.20%</td>
<td>99.80%</td>
</tr>
<tr>
<td>WV</td>
<td>1</td>
<td>0.20%</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Source: Author, FDIC
Source: SNL
# ANNEX 6. BOTTOM-UP APPROACH - INDIVIDUAL DEGREE OF SIGNIFICANCE

## Capital Adequacy Summary

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 1 Risk-based Ratio</td>
<td>4.69</td>
</tr>
<tr>
<td>Tier 1 Common Capital (CET1) Ratio</td>
<td>4.83</td>
</tr>
<tr>
<td>Risk Based Capital Ratio</td>
<td>4.28</td>
</tr>
<tr>
<td>Leverage Ratio</td>
<td>-4.83</td>
</tr>
<tr>
<td>Equity/Assets</td>
<td>-2.52</td>
</tr>
<tr>
<td>Equity Capital/Avg Assets</td>
<td>-2.59</td>
</tr>
<tr>
<td>Tangible Equity/Tangible Assets</td>
<td>3.96</td>
</tr>
<tr>
<td>Tangible Equity/Tangible Assets</td>
<td>4.06</td>
</tr>
<tr>
<td>Internal Growth Rate Ratio</td>
<td>2.53</td>
</tr>
<tr>
<td>Tier 1 Capital</td>
<td>0.11</td>
</tr>
<tr>
<td>Tier 1 Common Capital (CET1)</td>
<td>0.11</td>
</tr>
<tr>
<td>Total Risk Based Capital</td>
<td>0.11</td>
</tr>
<tr>
<td>Risk Weighted Assets</td>
<td>0.11</td>
</tr>
<tr>
<td>Avg Adj. Adjusted (Leverage Ratio)</td>
<td>0.1</td>
</tr>
</tbody>
</table>

## Asset Quality Summary

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPLs/Loans</td>
<td>5</td>
</tr>
<tr>
<td>Adjusted NPL/Total Loans</td>
<td>3</td>
</tr>
<tr>
<td>NPA/Assets</td>
<td>8.22</td>
</tr>
<tr>
<td>Adjusted NPA/Total Assets</td>
<td>8.24</td>
</tr>
<tr>
<td>Tier</td>
<td>7.96</td>
</tr>
<tr>
<td>Tier 1 Common Capital (CET1)</td>
<td>7.95</td>
</tr>
<tr>
<td>Tier 1 Capital</td>
<td>7.97</td>
</tr>
<tr>
<td>Tier 1 Loan Loss/Provl Tier 1 Capital</td>
<td>7.87</td>
</tr>
<tr>
<td>Reserves/NPAs</td>
<td>2.43</td>
</tr>
<tr>
<td>Net Chargeoffs/Avg Loans</td>
<td>2.84</td>
</tr>
<tr>
<td>Nonaccrual + 90PD+OREO</td>
<td>0.37</td>
</tr>
<tr>
<td>Adjusted Nonperforming Loans</td>
<td>0.49</td>
</tr>
<tr>
<td>Adjusted Nonperforming Assets</td>
<td>0.54</td>
</tr>
</tbody>
</table>

## Earnings Summary

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of Funds</td>
<td>3.18</td>
</tr>
<tr>
<td>Yield on Loans and Leases</td>
<td>0.65</td>
</tr>
<tr>
<td>Yield on Earning Assets</td>
<td>2.43</td>
</tr>
<tr>
<td>Yield/Cost Spread</td>
<td>2.3</td>
</tr>
<tr>
<td>ROAE</td>
<td>8.54</td>
</tr>
<tr>
<td>ROACE</td>
<td>8.59</td>
</tr>
<tr>
<td>ROATN</td>
<td>6.61</td>
</tr>
<tr>
<td>ROE</td>
<td>5.36</td>
</tr>
<tr>
<td>Noninterest Income/Operating Revenue</td>
<td>3.09</td>
</tr>
<tr>
<td>Net Interest Margin</td>
<td>3.67</td>
</tr>
<tr>
<td>Net Interest Margin (FTI)</td>
<td>2.1</td>
</tr>
<tr>
<td>Operating Exp/Operating Rev</td>
<td>4.25</td>
</tr>
<tr>
<td>Efficiency Ratio (FTI)</td>
<td>4.25</td>
</tr>
<tr>
<td>Noninterest Expense/Avg Assets</td>
<td>2.3</td>
</tr>
<tr>
<td>Net Income</td>
<td>0.37</td>
</tr>
<tr>
<td>Pre-Provision Net Revenue</td>
<td>0.29</td>
</tr>
<tr>
<td>Operating Revenue</td>
<td>0.15</td>
</tr>
</tbody>
</table>

## Liquidity Summary

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquid Assets/Assets</td>
<td>0.86</td>
</tr>
<tr>
<td>Liquidity Ratio</td>
<td>0.95</td>
</tr>
<tr>
<td>Net Noncore</td>
<td>2.86</td>
</tr>
<tr>
<td>On Hand Liquidity/ Liabilities</td>
<td>0.77</td>
</tr>
<tr>
<td>Reliance on Wholesale Funding</td>
<td>4.82</td>
</tr>
<tr>
<td>Leasing/Dep</td>
<td>1.56</td>
</tr>
<tr>
<td>Net Loans &amp; Leases/Total Deposits</td>
<td>1.37</td>
</tr>
<tr>
<td>Gross Loans &amp; Leases/Core Deposits</td>
<td>3.62</td>
</tr>
<tr>
<td>Total Loans &amp; Leases/Core Deposits</td>
<td>3.61</td>
</tr>
<tr>
<td>Brokered Deposits/Deposits</td>
<td>5.83</td>
</tr>
<tr>
<td>Total Liquid Assets</td>
<td>0.06</td>
</tr>
</tbody>
</table>

## Sensitivity Summary

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment Portfolio Depreciation</td>
<td>0.54</td>
</tr>
<tr>
<td>Long-term Assets/Assets</td>
<td>-3.06</td>
</tr>
<tr>
<td>Non-maturing Deposits/Long-term Assets</td>
<td>0.63</td>
</tr>
<tr>
<td>Residential Real Estate/Total Assets</td>
<td>-2.67</td>
</tr>
<tr>
<td>Rate sensitive Assets</td>
<td>0.06</td>
</tr>
<tr>
<td>Rate sensitive Liabilities</td>
<td>0.01</td>
</tr>
<tr>
<td>One Year Gap/Assets</td>
<td>5.59</td>
</tr>
<tr>
<td>One Year Cumulative Repricing Gap</td>
<td>0.17</td>
</tr>
</tbody>
</table>
II.7 References


Knight, F. (1921). Risk, uncertainty and profit.


Capítulo 3

ANÁLISIS DE QUIEBRAS MEDIANTE COMPONENTES PRINCIPALES
III.1. INTRODUCCIÓN

Este trabajo investiga las vías de mejora del “modelo mixto” desarrollado en un estudio anterior (“Bank Failure Prediction Models: Less is more?”, 2014), con el objetivo de introducir el mayor número posible de variables de cada componente del indicador CAMEL, al tiempo que se evita incurrir en el problema de la multicolinealidad, o elevada correlación entre las variables explicativas.

La base metodológica seleccionada ha consistido en la búsqueda de un modelo logit capaz de discriminar dos subgrupos a partir del total de entidades bancarias existentes en el momento temporal t: (i) entidades que quebrarán en el momento t+1, y (ii) entidades que sobrevivirán en dicho momento t+1.

Al modelo logit propuesto inicialmente se le exigirá la incorporación de la variable “ratio de Texas”, al tiempo que requerirá la inclusión del máximo número de variables CAMEL de todos los grupos. Se pretende especificar un modelo inicial suficientemente estable para prevenir las quiebras futuras con un año de anterioridad. El punto de partida ha sido el año 2007, al objeto de predecir las quiebras de las entidades bancarias producidas durante el año 2008. La robustez del modelo se comprobará a través del análisis de los niveles de predicción de las quiebras bancarias en los Estados Unidos durante el período 2009-2012.

III.2. REVISIÓN DE LA LITERATURA
III.3. METODOLOGÍA EMPÍRICA: Construcción de los factores y especificaciones del modelo

III.3.1. ANÁLISIS DE COMPONENTES PRINCIPALES

El problema de la multicolinealidad surge ante la existencia de correlación lineal entre las variables explicativas del modelo, que provoca unos estimadores de los parámetros que son insesgados pero se ven afectados por errores ante aumentos o disminuciones en el tamaño de la muestra, produciéndose incrementos en las varianzas y covarianzas conforme aumenta el grado de colinealidad, lo que su vez implica pérdidas de fiabilidad en los contrastes de significación al incrementarse los intervalos de probabilidad. Por tanto, la existencia de unos elevados niveles de multicolinealidad invalida la utilización del modelo de cara al análisis estructural.

Un procedimiento empleado en la corrección de la multicolinealidad es la utilización del método de Análisis Factorial mediante Componentes Principales. El método de Análisis de Componentes Principales (ACP) tiene como objetivo la transformación –sin pérdida de información– de un conjunto de variables originales en una combinación lineal de las originales, que son denominadas componentes principales (factores). Así, la metodología ACP intenta hallar estos componentes o factores, los cuales se caracterizan por estar incorrelacionadas entre sí, que sucesivamente expliquen la mayor proporción de la varianza total.

El primer factor -o componente- de un Análisis de Componentes Principales sería aquel que explica una mayor parte de la varianza total; el segundo factor explica la mayor parte de la varianza restante, es decir la no explicada por el primer factor, así sucesivamente. De este modo sería posible obtener tantos componentes como variables originales, aunque en la práctica no tiene sentido.
Es relevante destacar que las componentes principales se expresan en forma de una combinación lineal de las variables originales. Desde el punto de vista de su aplicación, el método de componentes principales es considerado como un método de reducción de datos, porque permite reducir la dimensión del número de variables inicialmente consideradas.

Respecto al interés operativo del ACP, éste viene asociado a procesos y situaciones en los que se dispone de información de multitud de variables que presentan correlación entre sí en mayor o menor grado, actuando como un velo que impide una adecuada evaluación del papel de cada variable en el fenómeno analizado. El ACP permite trabajar sobre un nuevo conjunto de variables, las componentes principales, que gozan de la ventaja de estar incorrelacionadas entre sí y que, además, pueden ordenarse de acuerdo con el nivel de información que llevan incorporada. Para medir la cantidad de información incorporada en una componente se utiliza la varianza: cuanto mayor sea su varianza, la información incorporada en dicha componente es superior. Por esta razón se selecciona como primera componente aquella que tenga mayor varianza, siendo la última la de menor varianza.

Una de las principales aplicaciones del ACP es la búsqueda de asociaciones de variables. Por este motivo, la asociación de variables aplicada sobre la metodología CAMEL, compuesta por variables de capital (C), calidad de activos (A), gestión (M), resultados (E) y liquidez (L), puede proporcionarnos una información muy valiosa de los distintos procesos asociados al fenómeno en estudio, en nuestro caso las quiebras bancarias.

III.3.1. 1. **Análisis de componentes principales para las variables de Capital**
Las variables que forman el grupo de Capital son las siguientes:

<table>
<thead>
<tr>
<th>Variables del Grupo Capital (Número: 14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1RBR</td>
</tr>
<tr>
<td>LEVRAT</td>
</tr>
<tr>
<td>RBCAPR</td>
</tr>
<tr>
<td>LEVRAT</td>
</tr>
<tr>
<td>TANG_EQ_TANGASS</td>
</tr>
<tr>
<td>TANG_EQ_TANGASS</td>
</tr>
<tr>
<td>INTGROWTHREQ</td>
</tr>
<tr>
<td>T1C</td>
</tr>
<tr>
<td>TRBC</td>
</tr>
<tr>
<td>AVGADJASS</td>
</tr>
</tbody>
</table>

Con carácter previo a la utilización del ACP es preciso comprobar si dicho método es necesario, contrastando si la correlación existente entre las variables explicativas analizadas es suficientemente significativa como para justificar la factorización de la matriz de coeficientes de correlación, Tabla1.

Tabla1: Matriz de correlaciones de las variables CAMEL de capital

<table>
<thead>
<tr>
<th>Correlación</th>
<th>T1RBR</th>
<th>LEVRAT</th>
<th>RBCAPR</th>
<th>LEVRAT</th>
<th>EQ_ASS</th>
<th>EQCA_AVASS</th>
<th>TANG_EQ_TANGASS</th>
<th>TANG_EQ_TANGASS</th>
<th>INTGROWTHREQ</th>
<th>T1C</th>
<th>T1CC</th>
<th>TRBC</th>
<th>RWA</th>
<th>AVGADJASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1RBR</td>
<td>0.95</td>
<td>0.95</td>
<td>0.77</td>
<td>0.79</td>
<td>0.77</td>
<td>0.802</td>
<td>0.001</td>
<td>0.006</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.018</td>
</tr>
<tr>
<td>LEVRAT</td>
<td>0.9</td>
<td>0.99</td>
<td>0.60</td>
<td>0.82</td>
<td>0.81</td>
<td>0.838</td>
<td>0.839</td>
<td>0.006</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>RBCAPR</td>
<td>0.9</td>
<td>0.99</td>
<td>0.60</td>
<td>0.82</td>
<td>0.811</td>
<td>0.837</td>
<td>0.837</td>
<td>0.006</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>15</td>
<td>-0.015</td>
</tr>
<tr>
<td>LEVRAT_A</td>
<td>0.7</td>
<td>0.60</td>
<td>0.60</td>
<td>0.72</td>
<td>0.725</td>
<td>0.729</td>
<td>0.728</td>
<td>0.003</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>22</td>
<td>-0.023</td>
</tr>
<tr>
<td>EQ_ASS</td>
<td>0.7</td>
<td>0.82</td>
<td>0.82</td>
<td>0.72</td>
<td>0.993</td>
<td>0.986</td>
<td>0.985</td>
<td>-0.003</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>17</td>
<td>-0.021</td>
</tr>
<tr>
<td>EQCA_AVASS</td>
<td>0.7</td>
<td>0.81</td>
<td>0.81</td>
<td>0.72</td>
<td>0.997</td>
<td>0.977</td>
<td>-0.003</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td>-0.021</td>
</tr>
<tr>
<td>TANGEQ_TANGASS</td>
<td>0.8</td>
<td>0.83</td>
<td>0.83</td>
<td>0.72</td>
<td>0.979</td>
<td>0.999</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>35</td>
<td>-0.035</td>
</tr>
<tr>
<td>TANG_EQ_TANGASS</td>
<td>0.8</td>
<td>0.83</td>
<td>0.83</td>
<td>0.72</td>
<td>0.977</td>
<td>0.999</td>
<td>1</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>35</td>
<td>-0.035</td>
</tr>
<tr>
<td>INTGROWTHREQ</td>
<td>0.0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>-0.003</td>
<td>0.002</td>
<td>0.002</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>21</td>
<td>-0.024</td>
</tr>
<tr>
<td>T1C</td>
<td>0.0</td>
<td>-0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.017</td>
<td>-0.035</td>
<td>-0.035</td>
<td>-0.021</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>18</td>
<td>0.98</td>
</tr>
<tr>
<td>RWA</td>
<td>0.9</td>
<td>0.99</td>
<td>0.99</td>
<td>0.94</td>
<td>0.98</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
La inspección visual de la matriz de correlaciones no es adecuada cuando existe un número elevado de variables. En dichos casos es necesario utilizar índices para dilucidar si existe un elevado grado de dependencia en la matriz de correlaciones, aspecto que permitirá la extracción de factores.

Los índices más habituales para determinar el grado de viabilidad de un análisis factorial, son los siguientes:

a) **Determinante de la matriz de correlaciones**: se utiliza como índice del tamaño de las correlaciones. Cuando su valor es elevado las correlaciones dentro de la matriz son bajas, mientras que un determinante reducido indica unos niveles de correlación elevados en la matriz.

b) **Prueba de esfericidad de Barlett**: contrasta la hipótesis nula de que la matriz de correlaciones observadas es en realidad una matriz identidad. Asumiendo que los datos provienen de una distribución normal multivariante, el estadístico de Barlett se distribuye aproximadamente según el modelo de probabilidad chi-cuadrado y es una transformación del determinante de la matriz de correlaciones. Un nivel crítico (Sig.) inferior a 0,05 indica el rechazo de la hipótesis nula al 95% de significación, lo que a su vez implica la existencia de correlación, pudiéndose afirmar –en consecuencia- que el modelo factorial es adecuado para explicar los datos.
Cuando las variables explicativas están correlacionadas se observan muchos valores elevados –en valor absoluto– fuera de la diagonal principal de la matriz de correlaciones, y el determinante es inferior a 1 (el valor máximo del determinante es 1, si las variables están incorrelacionadas). El test de Barlett realiza el siguiente contraste:

\[
H_0: |R| = 1
\]

\[
H_1: |R| \neq 1
\]

El valor de la determinante de la matriz de correlaciones aporta una orientación acerca del nivel de correlación generalizada entre todas las variables. El test se basa en una distribución chi-cuadrado en la que los valores elevados conducen al rechazo de H0. Dicho de otro modo, la prueba de esfericidad de Bartlett contrasta si la matriz de correlaciones es una matriz identidad, en cuyo caso el modelo factorial es inadecuado.

c) **Medida de adecuación muestral de Kaiser-Meyer-Olkin:** contempla las correlaciones parciales entre las variables explicativas. Es aconsejable obtener valores elevados, al menos superiores a 0,60, para considerar que el análisis factorial podrá realizarse con garantías. Normalmente se establece la siguiente escala de valores del test KMO:

- **Si KMO \( \geq 0,9 \):** implica que la relación entre las variables es muy elevada, y por tanto el resultado del test es muy satisfactorio para el empleo de ACP.

- **Para 0,9 > KMO \( \geq 0,8 \):** el test es satisfactorio.

- **Entre 0,8 > KMO \( \geq 0,7 \):** relación media entre las variables explicativas.
**Entre 0,7 > KMO ≥ 0,6:** relación débil entre las variables explicativas.

**KMO < 0,5:** variables explicativas muy levemente correlacionadas.

En la Tabla 2 se recogen los resultados de estos 3 índices para las variables del Grupo Capital.

**Tabla 2: Resultados de los Índices de idoneidad del análisis Factorial**

<table>
<thead>
<tr>
<th>Índice de idoneidad del Análisis Factorial</th>
<th>Resultados</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determinante de la Matriz de correlaciones</td>
<td>6,85E-023</td>
</tr>
<tr>
<td>Medida de adecuación muestral de Kaiser-Meyer-Olkin.</td>
<td>0,668</td>
</tr>
<tr>
<td>Prueba de esfericidad de Barlett Chi-cuadrado aproximado</td>
<td>390593,849</td>
</tr>
<tr>
<td>gl</td>
<td>91</td>
</tr>
<tr>
<td>Sig.</td>
<td>0,000</td>
</tr>
</tbody>
</table>

El valor del Determinante es bastante reducido y por tanto aceptable (6.85E-023), y la prueba de esfericidad de Barlett rechaza la hipótesis $H_0$, lo que significa el ACP es factible. No obstante, la Medida de adecuación muestral de Kaiser-Meyer-Olkin muestra un valor de 0,668, lo que indicaría una relación algo débil entre las variables.

Tras evaluar la adecuación de los datos de cara a la realización de un análisis factorial, se inicia el proceso de extracción de los factores por el método de componentes principales.

**Criterio 1:** Consiste en retener el número de factores necesario para lograr un alto porcentaje de explicación de la varianza total. Para ello, se utilizan los porcentajes
acumulados de los valores propios, con base en la varianza total del problema. La tabla 3 contiene los valores propios y las inercias explicadas por las componentes.

**Tabla 3:** Varianza Total explicada de los componentes de las variables de Capital.

**Varianza total explicada**

<table>
<thead>
<tr>
<th>Componente</th>
<th>Autovalores iniciales</th>
<th>Sumas de las saturaciones al cuadrado de la extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% de la varianza</td>
<td>% acumulado</td>
</tr>
<tr>
<td>Total</td>
<td>6,915</td>
<td>49,392</td>
</tr>
<tr>
<td>1</td>
<td>6,915</td>
<td>49,392</td>
</tr>
<tr>
<td>2</td>
<td>4,925</td>
<td>35,181</td>
</tr>
<tr>
<td>3</td>
<td>1,000</td>
<td>7,140</td>
</tr>
<tr>
<td>4</td>
<td>.593</td>
<td>4,234</td>
</tr>
<tr>
<td>5</td>
<td>.461</td>
<td>3,296</td>
</tr>
<tr>
<td>6</td>
<td>.055</td>
<td>.396</td>
</tr>
<tr>
<td>7</td>
<td>.032</td>
<td>.230</td>
</tr>
<tr>
<td>8</td>
<td>.007</td>
<td>.048</td>
</tr>
<tr>
<td>9</td>
<td>.007</td>
<td>.047</td>
</tr>
<tr>
<td>10</td>
<td>.002</td>
<td>.016</td>
</tr>
<tr>
<td>11</td>
<td>.002</td>
<td>.011</td>
</tr>
<tr>
<td>12</td>
<td>.001</td>
<td>.005</td>
</tr>
<tr>
<td>13</td>
<td>.000</td>
<td>.003</td>
</tr>
<tr>
<td>14</td>
<td>.000</td>
<td>.001</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de Componentes principales.

Se observa que las dos primeras componentes son las que más peso tienen en la explicación de los datos. La primera de ellas explica un 49,392% de la inercia total de la nube de puntos, y la segunda un 35,181%, con lo que aportan una explicación conjunta del 84,573%.

En la tabla de comunalidades, Tabla 4, puede apreciarse la varianza de cada variable del grupo de Capital, explicada por los componentes principales.
retenidos. En el método de componentes principales, cuando se retienen todas las variables resulta siempre una comunalidad inicial de 1. La comunalidad final es la que indica la variabilidad de cada variable, que es explicada por el conjunto de componentes principales retenidos.

**Tabla 4**: Tabla de Comunalidades asociada al análisis.

<table>
<thead>
<tr>
<th>Comunalidades</th>
<th>Inicial</th>
<th>Extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1RBR</td>
<td>1,000</td>
<td>,850</td>
</tr>
<tr>
<td>RBCAPR</td>
<td>1,000</td>
<td>,861</td>
</tr>
<tr>
<td>EQ_ASS</td>
<td>1,000</td>
<td>,927</td>
</tr>
<tr>
<td>TANGEQ_TANGASS</td>
<td>1,000</td>
<td>,938</td>
</tr>
<tr>
<td>INTEGRGROWTHREQ</td>
<td>1,000</td>
<td>,001</td>
</tr>
<tr>
<td>T1CC</td>
<td>1,000</td>
<td>,971</td>
</tr>
<tr>
<td>RWA</td>
<td>1,000</td>
<td>,995</td>
</tr>
<tr>
<td>LEVRAT</td>
<td>1,000</td>
<td>,862</td>
</tr>
<tr>
<td>LEVRAT_A</td>
<td>1,000</td>
<td>,617</td>
</tr>
<tr>
<td>EQCA_AVASS</td>
<td>1,000</td>
<td>,912</td>
</tr>
<tr>
<td>TANGCEQ_TANGASS</td>
<td>1,000</td>
<td>,937</td>
</tr>
<tr>
<td>T1C</td>
<td>1,000</td>
<td>,998</td>
</tr>
<tr>
<td>TRBC</td>
<td>1,000</td>
<td>,998</td>
</tr>
<tr>
<td>AVGADJASS</td>
<td>1,000</td>
<td>,974</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de Componentes principales.

Como queda reflejado en la tabla de comunalidades, prácticamente todas las variables del grupo Capital recogen más del 80% de sus varianzas en los dos componentes principales retenidos. Las dos únicas excepciones son las variables LEVRAT_A (61,7%) y INTEGRGROWTHREQ (1%).

**Criterio 2**: Consiste en representar un gráfico de sedimentación (scree plot) de los valores propios -como el de la figura 1- y considerar el número de componentes en el cual el descenso se estabiliza.
**Figura 1:** Gráfico de Sedimentación (scree plot) de los autovalores de las variables de Capital

En este caso nos hemos inclinado por retener los dos primeros componentes, ya que explican un 86,1% de la varianza y permiten una representación gráfica en dos dimensiones

Las variables que conformarán los dos componentes principales seleccionados vendrán determinados por la matriz de Componentes, Tabla 5.

**Tabla 5:** Matriz de componentes seleccionados.

<table>
<thead>
<tr>
<th>Matriz de componentes^a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Método de extracción: Análisis de componentes principales.
a. 2 componentes extraídos

En lo sucesivo los dos componentes seleccionados serán denominados FAC1_Capital y FAC2_Capital:

- **FAC1_Capital**: T1RBR, LEVRAT, RBCAPR, LEVRAT_A, EQ_ASS, EQCA_AVASS, TANGEQ_TANGASS, TANGCEQ_TANGASS,
- **FAC2_Capital**: INTGROWTHREQ, T1C, T1CC, TRBC, RWA, AVGADJASS

Teniendo presente que la Medida de adecuación muestral de Kaiser-Meyer-Olkin ha indicado que una relación débil entre las variables, intentamos realizar la creación de tres componentes principales cuyas medidas de adecuación muestral de Kaiser-Meyer-Olkin sean más elevados, de cara mejorar los resultados que serán obtenidos posteriormente al desarrollar el “modelo logit binomial de quiebras bancarias”. Los componentes principales obtenidos son:

1) Componente principal **FAC1_Kapital**: Este componente principal, formado por las variables T1RBR, LEVRAT, RBCAPR, tiene la siguiente matriz de correlaciones:

<table>
<thead>
<tr>
<th></th>
<th>FAC1_Capital</th>
<th>FAC2_Capital</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQCA_AVASS</td>
<td>.953</td>
<td>.064</td>
</tr>
<tr>
<td>TANGEQ_TANGASS</td>
<td>.967</td>
<td>.049</td>
</tr>
<tr>
<td>TANGCEQ_TANGASS</td>
<td>.967</td>
<td>.048</td>
</tr>
<tr>
<td>INTGROWTHREQ</td>
<td>.005</td>
<td>-.027</td>
</tr>
<tr>
<td>T1C</td>
<td>-.086</td>
<td>.995</td>
</tr>
<tr>
<td>T1CC</td>
<td>-.085</td>
<td>.982</td>
</tr>
<tr>
<td>TRBC</td>
<td>-.086</td>
<td>.995</td>
</tr>
<tr>
<td>RWA</td>
<td>-.089</td>
<td>.993</td>
</tr>
<tr>
<td>AVGADJASS</td>
<td>-.087</td>
<td>.983</td>
</tr>
</tbody>
</table>
Tabla 6: Matriz de correlaciones de las variables que conforman FAC1_Kapital

Matriz de correlaciones*

<table>
<thead>
<tr>
<th></th>
<th>T1RBR</th>
<th>LEVRAT</th>
<th>RBCAPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlación</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1RBR</td>
<td>1,000</td>
<td>.964</td>
<td>.965</td>
</tr>
<tr>
<td>LEVRAT</td>
<td>.964</td>
<td>1,000</td>
<td>.999</td>
</tr>
<tr>
<td>RBCAPR</td>
<td>.965</td>
<td>.999</td>
<td>1,000</td>
</tr>
<tr>
<td>Sig. (Unilateral)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1RBR</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEVRAT</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBCAPR</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Determinante = .000

El siguiente paso es el análisis de los índices de idoneidad del análisis factorial de las variables que conforman FAC1_Kapital, y que se obtiene de la matriz de correlaciones, que aparecen en la Tabla 7.

Tabla 7: Resultados de los Índices de idoneidad del análisis Factorial

Resultados Índices de idoneidad del Análisis Factorial

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Determinante de la Matriz de correlaciones</td>
<td>0,000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medida de adecuación muestral de Kaiser-Meyer-Olkin.</td>
<td>0,741</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prueba de esfericidad de Bartlett</td>
<td>Chi-cuadrado aproximado</td>
<td>70771,847</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>gl</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sig.</td>
<td></td>
<td>0,000</td>
<td></td>
</tr>
</tbody>
</table>

En este caso el valor del Determinante presenta un valor muy aceptable (0,000), y la prueba de esfericidad de Barlett rechaza la hipótesis H0 -lo que significa el ACP es factible-, al tiempo que se ha logrado una Medida de adecuación muestral de
Kaiser-Meyer-Olkin que muestra un valor de 0,741 lo que indica una relación intermedia entre las variables, mejorando el nivel resultante en el caso anterior.

La tabla 8 contiene los valores propios y las inercias explicadas por las componentes, observándose que la primera componente es la que más peso tiene a la hora de explicar los datos, mostrando un 98,4% de la inercia total de la nube de puntos.

**Tabla 8: Varianza Total explicada de los componentes de las variables de FAC1_Kapital.**

<table>
<thead>
<tr>
<th>Componente</th>
<th>Autovalores iniciales</th>
<th>Sumas de las saturaciones al cuadrado de la extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% de la varianza</td>
</tr>
<tr>
<td>1</td>
<td>2,952</td>
<td>98,402</td>
</tr>
<tr>
<td>2</td>
<td>.047</td>
<td>1,574</td>
</tr>
<tr>
<td>3</td>
<td>.001</td>
<td>.025</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de Componentes principales.

Según figura en la tabla de comunalidades, todas las variables del grupo Capital recogen más del 95% de sus varianzas en el componente principal retenido:

**Tabla 9: Tabla de Comunalidades asociada al análisis**

<table>
<thead>
<tr>
<th>Inicial</th>
<th>Extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1RBR</td>
<td>1,000</td>
</tr>
<tr>
<td>LEVRAT</td>
<td>1,000</td>
</tr>
<tr>
<td>RBCAPR</td>
<td>1,000</td>
</tr>
</tbody>
</table>
Método de extracción: Análisis de Componentes principales.

Siendo la Matriz de componentes del primer componente principal, FAC1_Kapital, la que se muestra en la Tabla 10:

**Tabla 10:** Tabla de Componente de las variables asociadas a FAC1_Kapital.

Matriz de componentes

<table>
<thead>
<tr>
<th>Variable</th>
<th>Componente 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1RBR</td>
<td>0.984</td>
</tr>
<tr>
<td>LEVRAT</td>
<td>0.996</td>
</tr>
<tr>
<td>RBCAPR</td>
<td>0.996</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de componentes principales.

a. 1 componentes extraídos

2) Componente principal **FAC2_Kapital:** formado por las variables EQ_ASS, EQCA_AVASS, TANGEQ_TANGASS, TANGCEQ_TANGASS, tiene la siguiente matriz de correlaciones:

**Tabla 11:** Matriz de correlaciones de las variables que conforman FAC2_Kapital

Matriz de correlaciones

<table>
<thead>
<tr>
<th></th>
<th>EQ AS S</th>
<th>EQCA_AVASS S</th>
<th>TANGEQ_TANGASS SS</th>
<th>TANGCEQ_TANGASS SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORRELACIÒN</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EQ_ASS</td>
<td>1.000</td>
<td>0.574</td>
<td>0.988</td>
<td>0.987</td>
</tr>
<tr>
<td>EQCA_AVASS</td>
<td>0.574</td>
<td>1.000</td>
<td>0.570</td>
<td>0.570</td>
</tr>
<tr>
<td>TANGEQ_TANGASS</td>
<td>0.988</td>
<td>0.570</td>
<td>1.000</td>
<td>0.999</td>
</tr>
<tr>
<td>TANGCEQ_TANGASS</td>
<td>0.987</td>
<td>0.570</td>
<td>0.999</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Si analizamos los índices de idoneidad del análisis factorial de las variables que conforman **FAC2_Kapital**, Tabla 12, todos son significativos, y en concreto el Indice KMO informa sobre la buena relación de las tres variables de capital de este componente principal.

**Tabla 12: Resultados de los Índices de idoneidad del análisis Factorial**

| Determinante de la Matriz de correlaciones | 2,80E-005 |
| Medida de adecuación muestral de Kaiser-Meyer-Olkin. | 0,796 |
| Prueba de esfericidad de Bartlett | Chi-cuadrado aproximado |
| | 80824,026 |
| | gl | 6 |
| | Sig. | 0,000 |

La Tabla 13 muestra el mayor peso de la primera componente, **FAC2_Kapital**, en la explicación de los datos, con un 84,8% de la inercia total de la nube de puntos.

**Tabla 13: Varianza Total explicada de los componentes de las variables de FAC2_Kapital.**

<table>
<thead>
<tr>
<th>Varianza total explicada</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

16
<table>
<thead>
<tr>
<th>Componente</th>
<th>Autovalores iniciales</th>
<th>Sumas de las saturaciones al cuadrado de la extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>1</td>
<td>3,392</td>
<td>84,805</td>
</tr>
<tr>
<td>2</td>
<td>.591</td>
<td>14,778</td>
</tr>
<tr>
<td>3</td>
<td>.016</td>
<td>.395</td>
</tr>
<tr>
<td>4</td>
<td>.001</td>
<td>.022</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de Componentes principales.

En la tabla de comunalidades del componente principal **FAC2_Kapital**, todas las variables del grupo Capital recogen más del 95% de sus varianzas, excepto la variable EQCA_AVASS (49,5%).

**Tabla 14:** Tabla de Comunalidades asociada al análisis

### Comunalidades

<table>
<thead>
<tr>
<th></th>
<th>Inicial</th>
<th>Extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ_ASS</td>
<td>1,000</td>
<td>.962</td>
</tr>
<tr>
<td>EQCA_AVASS</td>
<td>1,000</td>
<td>.495</td>
</tr>
<tr>
<td>TANGEQ_TANGASS</td>
<td>1,000</td>
<td>.968</td>
</tr>
<tr>
<td>TANGCEQ_TANGASS</td>
<td>1,000</td>
<td>.967</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de Componentes principales.

Por último, el componente principal **FAC2_Kapital** tiene asociada la siguiente matriz de componentes, Tabla 15.

**Tabla 15:** Tabla de Componente de las variables asociadas a **FAC2_Kapital**.
Matriz de componentes

<table>
<thead>
<tr>
<th>Componente</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQ_ASS</td>
<td>,981</td>
</tr>
<tr>
<td>EQCA_AVASS</td>
<td>,704</td>
</tr>
<tr>
<td>TANGEQ_TANGASS</td>
<td>,984</td>
</tr>
<tr>
<td>TANGCEQ_TANGASS</td>
<td>,983</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de componentes principales.

a. 1 componentes extraídos

3) Componente principal **FAC3_Kapital**: Este componente principal, formado por las variables T1C, T1CC, TRBC, RWA, AVGADJASS, tiene la siguiente matriz de correlaciones:

**Tabla 16: Matriz de correlaciones de las variables que conforman FAC3_Kapital**

<table>
<thead>
<tr>
<th></th>
<th>T1C</th>
<th>T1CC</th>
<th>TRBC</th>
<th>RWA</th>
<th>AVGADJASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlación</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1C</td>
<td>1,000</td>
<td>.990</td>
<td>,998</td>
<td>,994</td>
<td>,980</td>
</tr>
<tr>
<td>T1CC</td>
<td>.990</td>
<td>1,000</td>
<td>.983</td>
<td>,973</td>
<td>,949</td>
</tr>
<tr>
<td>TRBC</td>
<td>,998</td>
<td>.983</td>
<td>1,000</td>
<td>,996</td>
<td>,984</td>
</tr>
<tr>
<td>RWA</td>
<td>,994</td>
<td>.973</td>
<td>,996</td>
<td>1,000</td>
<td>,990</td>
</tr>
<tr>
<td>AVGADJASS</td>
<td>,980</td>
<td>,949</td>
<td>,984</td>
<td>,990</td>
<td>1,000</td>
</tr>
</tbody>
</table>

|               |     |      |      |     |           |
| Sig. (Unilateral) |     |      |      |     |           |
| T1C           | ,000 |     | ,000 | ,000 | ,000      |
| T1CC          | ,000 | ,000 |     | ,000 | ,000      |
| TRBC          | ,000 | ,000 | ,000 | ,000 | ,000      |
| RWA           | ,000 | ,000 | ,000 |     | ,000      |
| AVGADJASS     | ,000 | ,000 | ,000 | ,000 |     |

a. Determinante = 3,03E-009
Los índices de idoneidad del análisis factorial de las variables que conforman **FAC3_Kapital** son significativos, Tabla 17, existiendo una buena relación de las cinco variables de capital al tener el Índice KMO un valor de 0,813.

**Tabla 17: Resultados de los Índices de idoneidad del análisis Factorial**

**Resultados Índices de idoneidad del Análisis Factorial**

| Determinante de la Matriz de correlaciones | 3,03E-009 |
| Medida de adecuación muestral de Kaiser-Meyer-Olkin. | 0,813 |
| Prueba de esfericidad de Bartlett | Chi-cuadrado aproximado gl Sig. |
|  | 151303,968 10 0,000 |

**Comunalidades**

<table>
<thead>
<tr>
<th></th>
<th>Inicial</th>
<th>Extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1C</td>
<td>1,000</td>
<td>,998</td>
</tr>
<tr>
<td>T1CC</td>
<td>1,000</td>
<td>,971</td>
</tr>
<tr>
<td>TRBC</td>
<td>1,000</td>
<td>,998</td>
</tr>
<tr>
<td>RWA</td>
<td>1,000</td>
<td>,995</td>
</tr>
<tr>
<td>AVGADJASS</td>
<td>1,000</td>
<td>,974</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de Componentes principales.
Varianza total explicada

<table>
<thead>
<tr>
<th>Componente</th>
<th>Autovalores iniciales</th>
<th>Sumas de las saturaciones al cuadrado de la extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% de la varianza</td>
</tr>
<tr>
<td>1</td>
<td>4,935</td>
<td>98,703</td>
</tr>
<tr>
<td>2</td>
<td>0,055</td>
<td>1,105</td>
</tr>
<tr>
<td>3</td>
<td>0,007</td>
<td>0,131</td>
</tr>
<tr>
<td>4</td>
<td>0,002</td>
<td>0,046</td>
</tr>
<tr>
<td>5</td>
<td>0,001</td>
<td>0,015</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de Componentes principales.

Matriz de componentes*

<table>
<thead>
<tr>
<th>Componente</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1C</td>
<td>.999</td>
</tr>
<tr>
<td>T1CC</td>
<td>.985</td>
</tr>
<tr>
<td>TRBC</td>
<td>.999</td>
</tr>
<tr>
<td>RWA</td>
<td>.997</td>
</tr>
<tr>
<td>AVGADJASS</td>
<td>.987</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de componentes principales.
a. 1 componentes extraídos

En definitiva, los tres últimos factores obtenidos estarían compuestos por las siguientes variables de Capital:

- **FAC1_Kapital**: T1RBR, LEVRAT, RBCAPR
- **FAC2_Kapital**: EQ_ASS, EQCA_AVASS, TANGEQ_TANGASS, TANGCEO_TANGASS.
III.3.1.2 Análisis de componentes principales para las variables de Asset Quality Summary.

Las variables que forman parte del grupo Asset Quality Summary son las siguientes:

Tabla 21: Variables que conforman el grupo Asset Quality Summary.

<table>
<thead>
<tr>
<th>Variables del Grupo Asset Quality Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>NPL_L</td>
</tr>
<tr>
<td>ADJNPL_L</td>
</tr>
<tr>
<td>NPA_ASS</td>
</tr>
<tr>
<td>ADJNPA_ASS</td>
</tr>
<tr>
<td>TR</td>
</tr>
<tr>
<td>TRM</td>
</tr>
<tr>
<td>MTR</td>
</tr>
<tr>
<td>LTMLLP_T1C</td>
</tr>
<tr>
<td>RES_NPA</td>
</tr>
<tr>
<td>RES_ADJNPA</td>
</tr>
<tr>
<td>NCO_AVGL</td>
</tr>
<tr>
<td>NONAC&amp;90PD&amp;OREO</td>
</tr>
<tr>
<td>ADJNPL</td>
</tr>
<tr>
<td>ADJNPA</td>
</tr>
</tbody>
</table>

En el análisis exploratorio realizado en la Base de Datos UBPR de la FFIEC para estas variables en el año 2007 se ha detectado un subgrupo que presenta un número considerable de valores perdidos, como puede observarse en la Tabla 22.

Tabla 22: Número de valores válidos y perdidos de las variables del grupo Asset Quality Summary.
Las variables que seleccionaremos para realizar el análisis de componentes principales son aquellas que no tienen valores perdidos, cuya matriz de correlaciones se refleja en la Tabla 23.

**Tabla 23:** Matriz de correlaciones de las variables seleccionadas de Asset Quality Summary.

<table>
<thead>
<tr>
<th></th>
<th>NPL_L</th>
<th>ADJNPL_L</th>
<th>NPA_ASS</th>
<th>ADJNPA_ASS</th>
<th>TR</th>
<th>TRM</th>
<th>MTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Válidos</td>
<td>7621</td>
<td>7621</td>
<td>7717</td>
<td>7717</td>
<td>7717</td>
<td>7717</td>
<td>7717</td>
</tr>
<tr>
<td>Perdidos</td>
<td>96</td>
<td>96</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>LTMLLP_TIC</th>
<th>RES_NPA</th>
<th>RES_ADJ_NPA</th>
<th>NCO_AV</th>
<th>NONAC&amp;P0PD&amp;P0REO</th>
<th>ADJN_PL</th>
<th>ADJN_PA</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Válidos</td>
<td>0</td>
<td>593</td>
<td>5973</td>
<td>7622</td>
<td>7717</td>
<td>7717</td>
<td>7717</td>
</tr>
<tr>
<td>Perdidos</td>
<td>7717</td>
<td>1724</td>
<td>1744</td>
<td>95</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Matriz de correlaciones**
A partir de la Matriz de correlaciones de las variables Asset Quality Summary seleccionadas se han construido los índices de idoneidad del análisis factorial, Tabla 24, comprobándose que aunque el análisis de componentes principales es factible, debido a los valores del Determinante de la Matriz y de la prueba de esfericidad de Barlett, existe una relación débil entre las variables al alcanzarse un valor de 0,645 del índice KMO.

**Tabla 24: Resultados de los Índices de idoneidad del análisis Factorial**

| Determinante de la Matriz de correlaciones | 6,88E-013 |
| Medida de adecuación muestral de Kaiser-Meyer-Olkin. | 0,645 |
| Prueba de esfericidad de Bartlett | Chi-cuadrado aproximado 215990,857 |
| gl | 28 |
| Sig. | 0,000 |

La Varianza Total explicada de los componentes de las variables de **Asset Quality Summary**, se recoge en la Tabla 25, alcanzándose una explicación del 96,45% de la varianza total con los dos primeros componentes. Las variables representativas de cada componente se evidencian en la Matriz de Componentes, Tabla 26.

**Tabla 25: Varianza Total explicada de los componentes de las variables de Asset Quality Summary.**
### Varianza total explicada

<table>
<thead>
<tr>
<th>Componente</th>
<th>Autovalores iniciales</th>
<th>Sumas de las saturaciones al cuadrado de la extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% de la varianza</td>
</tr>
<tr>
<td>1</td>
<td>4,799</td>
<td>59,987</td>
</tr>
<tr>
<td>2</td>
<td>2,917</td>
<td>36,463</td>
</tr>
<tr>
<td>3</td>
<td>212</td>
<td>2,647</td>
</tr>
<tr>
<td>4</td>
<td>0,059</td>
<td>0,741</td>
</tr>
<tr>
<td>5</td>
<td>0,011</td>
<td>0,133</td>
</tr>
<tr>
<td>6</td>
<td>0,001</td>
<td>0,013</td>
</tr>
<tr>
<td>7</td>
<td>0,001</td>
<td>0,011</td>
</tr>
<tr>
<td>8</td>
<td>0,000</td>
<td>0,005</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de Componentes principales.

**Tabla 26:** Tabla de Componentes de las variables Asset Quality Summary asociadas.

### Matriz de componentes

<table>
<thead>
<tr>
<th></th>
<th>Componente</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPA_ASS</td>
<td>.961</td>
<td>-.101</td>
<td></td>
</tr>
<tr>
<td>ADJNPA_ASS</td>
<td>.960</td>
<td>-.101</td>
<td></td>
</tr>
<tr>
<td>TR</td>
<td>.983</td>
<td>-.073</td>
<td></td>
</tr>
<tr>
<td>TRM</td>
<td>.982</td>
<td>-.072</td>
<td></td>
</tr>
<tr>
<td>MTR</td>
<td>.981</td>
<td>-.079</td>
<td></td>
</tr>
<tr>
<td>NONAC&amp;90PD&amp;OREO</td>
<td>.129</td>
<td>.972</td>
<td></td>
</tr>
<tr>
<td>ADJNPL</td>
<td>.148</td>
<td>.981</td>
<td></td>
</tr>
<tr>
<td>ADJNPA</td>
<td>.146</td>
<td>.987</td>
<td></td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de componentes principales.

a. 2 componentes extraídos
Los dos componentes principales obtenidos los denominaremos con la siguiente nomenclatura, y contendrán las variables que figuran a continuación:

- **FAC1_AQS**: NPA_ASS, ADJNPA_ASS, TR, TRM, MTR
- **FAC2_AQS**: NONAC\&90PD\&OREO, ADJNPL, ADJNPA

3.3.1.3 **Análisis de componentes principales para las variables de EARNINGS Summary.**

En el caso de las variables del grupo *EARNINGS Summary*, nos encontramos que al realizar el análisis exploratorio en todas ellas hay valores perdidos, Tabla 27.

**Tabla 27**: Número de valores válidos y perdidos de las variables del grupo *EARNINGS Summary*.

<table>
<thead>
<tr>
<th>Estadísticos</th>
<th>COSTFU</th>
<th>YIELD1&amp;amp;pL</th>
<th>YIELD&amp;amp;E</th>
<th>YIELD_CO</th>
<th>ROAE</th>
<th>ROACE</th>
<th>ROA</th>
<th>ROAT</th>
<th>ROATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Válidos</td>
<td>7620</td>
<td>7601</td>
<td>7683</td>
<td>6889</td>
<td>7673</td>
<td>7667</td>
<td>7622</td>
<td>7660</td>
<td>7665</td>
</tr>
<tr>
<td>Perdidos</td>
<td>97</td>
<td>116</td>
<td>34</td>
<td>828</td>
<td>44</td>
<td>50</td>
<td>95</td>
<td>57</td>
<td>52</td>
</tr>
<tr>
<td>N LOR</td>
<td>7703</td>
<td>7683</td>
<td>7686</td>
<td>7652</td>
<td>7652</td>
<td>7690</td>
<td>7711</td>
<td>7711</td>
<td>7711</td>
</tr>
<tr>
<td>Perdidos</td>
<td>14</td>
<td>34</td>
<td>31</td>
<td>65</td>
<td>65</td>
<td>27</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Debido a que no existe ninguna variable con 0 valores perdidos, planteamos dos escenarios:

1. Hacer un factorial con NI, PPNETREV y OPREV, pues son las que muestran el menor número de valores perdidos (6 valores perdidos)
2. Hacer el factorial con todas las variables que no superen los 100 valores perdidos: COSTFUND, YIELD\&EA, ROAE, ROACE, ROATE, ROATA, ROATC, NI_OR, NIM NIM_FTE, OPEXP_OPREV, EFFICIENCY, NEXP/AVGASS, NI, PPNETREV y OPREV

Comenzaremos con el estudio de factibilidad del análisis factorial del segundo caso, a través de su matriz de correlaciones, Tabla 28:

**Tabla 28:** Matriz de correlaciones de las variables de EARNINGS Summary seleccionadas como segunda opción.

<table>
<thead>
<tr>
<th></th>
<th>COSTFUND</th>
<th>YIELD &amp; EA</th>
<th>ROAE</th>
<th>ROATE</th>
<th>ROACE</th>
<th>ROATE</th>
<th>NIORM</th>
<th>NIM</th>
<th>OPEXP_OPREV</th>
<th>EFFICIENCY</th>
<th>NEXP_AVGASS</th>
<th>N1</th>
<th>PPNETREV</th>
<th>OPREV</th>
</tr>
</thead>
<tbody>
<tr>
<td>COSTFUND</td>
<td>1</td>
<td>0.13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.26</td>
<td>0.0088</td>
<td>0.001</td>
<td>0</td>
<td>0.009</td>
<td>0.23</td>
</tr>
<tr>
<td>YIELD &amp; EA</td>
<td>0.13</td>
<td>1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.18</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0088</td>
<td>0.005</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
<td>0.23</td>
</tr>
<tr>
<td>ROAE</td>
<td>0.13</td>
<td>1</td>
<td>0.13</td>
<td>0.1</td>
<td>0.14</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.77</td>
<td>0.015</td>
<td>0.05</td>
<td>0</td>
<td>0</td>
<td>1.2</td>
</tr>
<tr>
<td>ROATE</td>
<td>0.186</td>
<td>0.13</td>
<td>0.13</td>
<td>0.1</td>
<td>0.14</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.24</td>
<td>0.01</td>
<td>0.008</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ROACE</td>
<td>0.186</td>
<td>0.13</td>
<td>0.13</td>
<td>0.1</td>
<td>0.14</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.24</td>
<td>0.01</td>
<td>0.008</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>ROATC</td>
<td>0.183</td>
<td>0.141</td>
<td>0.14</td>
<td>0.1</td>
<td>0.14</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.25</td>
<td>0.006</td>
<td>0.006</td>
<td>0</td>
<td>0.097</td>
<td>0.05</td>
</tr>
<tr>
<td>ROA</td>
<td>0.124</td>
<td>0.148</td>
<td>0.14</td>
<td>0.1</td>
<td>0.14</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.25</td>
<td>0.002</td>
<td>0.004</td>
<td>0</td>
<td>0.032</td>
<td>0.02</td>
</tr>
<tr>
<td>ROATCE</td>
<td>0.182</td>
<td>0.14</td>
<td>0.14</td>
<td>0.1</td>
<td>0.14</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.25</td>
<td>0.006</td>
<td>0.005</td>
<td>0</td>
<td>0.097</td>
<td>0.05</td>
</tr>
</tbody>
</table>
a. Determinante = 2,25E-015

Analizando los índices de idoneidad del análisis factorial con las variables que presentan menos de 100 valores perdidos, que conforman la segunda opción, Tabla 29, se comprueba la baja relación existente entre ellas al obtener un valor del KMO de 0,564, siendo el límite establecido en 0,5.

**Tabla 29: Resultados de los Índices de idoneidad del análisis Factorial**

Resultados Índices de idoneidad del Análisis Factorial

<table>
<thead>
<tr>
<th>Determinante de la Matriz de correlaciones</th>
<th>2,25E-015</th>
</tr>
</thead>
</table>

27
Medida de adecuación muestral de Kaiser-Meyer-Olkin. 0,564  
Prueba de esfericidad de Bartlett  
Chi-cuadrado aproximado 254145,581  
gl 120  
Sig. 0,000

Este reducido nivel de correlación entre las variables se corrobora con las varianzas de los componentes obtenidos, Tabla 30, en la matriz de Varianza total. Tal y como se observa son necesarios seis componentes para alcanzar el 87,77% de la varianza total, con varianzas muy bajas.

**Tabla 30:** Varianza Total explicada de los componentes de las variables de EARNINGS Summary.

**Varianza total explicada**

<table>
<thead>
<tr>
<th>Componente</th>
<th>Autovalores iniciales</th>
<th>Sumas de las saturaciones al cuadrado de la extracción</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% de la varianza</td>
</tr>
<tr>
<td>1</td>
<td>5,825</td>
<td>36,406</td>
</tr>
<tr>
<td>2</td>
<td>2,532</td>
<td>15,825</td>
</tr>
<tr>
<td>3</td>
<td>2,014</td>
<td>12,589</td>
</tr>
<tr>
<td>4</td>
<td>1,502</td>
<td>9,389</td>
</tr>
<tr>
<td>5</td>
<td>1,161</td>
<td>7,253</td>
</tr>
<tr>
<td>6</td>
<td>1,010</td>
<td>6,312</td>
</tr>
<tr>
<td>7</td>
<td>.946</td>
<td>5,910</td>
</tr>
<tr>
<td>8</td>
<td>.572</td>
<td>3,572</td>
</tr>
<tr>
<td>9</td>
<td>.294</td>
<td>1,839</td>
</tr>
<tr>
<td>10</td>
<td>.063</td>
<td>.391</td>
</tr>
<tr>
<td>11</td>
<td>.045</td>
<td>.280</td>
</tr>
<tr>
<td>12</td>
<td>.025</td>
<td>.158</td>
</tr>
<tr>
<td>13</td>
<td>.009</td>
<td>.054</td>
</tr>
<tr>
<td>14</td>
<td>.003</td>
<td>.018</td>
</tr>
<tr>
<td>15</td>
<td>.000</td>
<td>.003</td>
</tr>
</tbody>
</table>
Método de extracción: Análisis de Componentes principales.

La Matriz de coeficientes para el cálculo de las puntuaciones en los componentes se recoge en la Tabla 31.

**Tabla 31:** Tabla de Componente de las variables asociadas a EARNINGS

<table>
<thead>
<tr>
<th>Componente</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>COSTFUND</td>
<td>-0.039</td>
<td>-0.013</td>
<td>-0.002</td>
<td>0.020</td>
<td>0.723</td>
<td>0.369</td>
</tr>
<tr>
<td>YIELD&amp;EA</td>
<td>0.050</td>
<td>0.319</td>
<td>0.026</td>
<td>0.220</td>
<td>0.264</td>
<td>0.103</td>
</tr>
<tr>
<td>ROAE</td>
<td>0.160</td>
<td>-0.063</td>
<td>-0.040</td>
<td>0.033</td>
<td>0.059</td>
<td>0.236</td>
</tr>
<tr>
<td>ROACE</td>
<td>0.160</td>
<td>-0.063</td>
<td>-0.039</td>
<td>0.033</td>
<td>0.061</td>
<td>0.242</td>
</tr>
<tr>
<td>ROATE</td>
<td>0.162</td>
<td>-0.060</td>
<td>-0.023</td>
<td>0.048</td>
<td>0.034</td>
<td>0.208</td>
</tr>
<tr>
<td>ROATA</td>
<td>0.136</td>
<td>-0.015</td>
<td>-0.023</td>
<td>0.162</td>
<td>0.170</td>
<td>0.064</td>
</tr>
<tr>
<td>ROATCE</td>
<td>0.162</td>
<td>-0.060</td>
<td>-0.023</td>
<td>0.047</td>
<td>0.035</td>
<td>0.212</td>
</tr>
<tr>
<td>NI.OR</td>
<td>0.049</td>
<td>-0.016</td>
<td>0.140</td>
<td>0.404</td>
<td>0.204</td>
<td>-0.259</td>
</tr>
<tr>
<td>NIM</td>
<td>0.070</td>
<td>0.352</td>
<td>0.040</td>
<td>0.030</td>
<td>0.127</td>
<td>0.061</td>
</tr>
<tr>
<td>NIM_FTE</td>
<td>0.073</td>
<td>0.346</td>
<td>0.037</td>
<td>0.029</td>
<td>0.139</td>
<td>0.074</td>
</tr>
<tr>
<td>OPEXP_OPREV</td>
<td>-0.126</td>
<td>0.085</td>
<td>0.052</td>
<td>0.235</td>
<td>0.139</td>
<td>0.437</td>
</tr>
<tr>
<td>EFFICIENCY</td>
<td>-0.125</td>
<td>0.088</td>
<td>0.054</td>
<td>0.234</td>
<td>0.127</td>
<td>0.440</td>
</tr>
<tr>
<td>NEXP/AVGASS</td>
<td>0.008</td>
<td>0.107</td>
<td>0.082</td>
<td>0.494</td>
<td>0.284</td>
<td>-0.038</td>
</tr>
<tr>
<td>NI</td>
<td>0.026</td>
<td>-0.055</td>
<td>0.415</td>
<td>-0.105</td>
<td>-0.126</td>
<td>0.230</td>
</tr>
<tr>
<td>PPNETREV</td>
<td>0.023</td>
<td>-0.059</td>
<td>0.474</td>
<td>-0.113</td>
<td>-0.038</td>
<td>0.026</td>
</tr>
<tr>
<td>OPREV</td>
<td>0.001</td>
<td>-0.034</td>
<td>0.243</td>
<td>-0.034</td>
<td>0.144</td>
<td>-0.348</td>
</tr>
</tbody>
</table>

Método de extracción: Análisis de componentes principales.
Estos seis componentes principales serán denominados FAC1_ES, FAC2_ES, FAC3_ES, FAC4_ES, FAC5_ES, FAC6_ES.

Ahora nos centraremos en el caso reducido, opción 1, que solamente incluye las variables de EARNINGS Summary con un menor número de valores perdidos, siendo la matriz de correlaciones la mostrada en la Tabla 32.

**Tabla 32:** Matriz de correlaciones de las variables de EARNINGS Summary seleccionadas como primera opción

Matriz de correlaciones

<table>
<thead>
<tr>
<th></th>
<th>NI</th>
<th>PPNETREV</th>
<th>OPREV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlación</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NI</td>
<td>1,000</td>
<td>,868</td>
<td>,037</td>
</tr>
<tr>
<td>PPNETREV</td>
<td>,868</td>
<td>1,000</td>
<td>,461</td>
</tr>
<tr>
<td>OPREV</td>
<td>,037</td>
<td>,461</td>
<td>1,000</td>
</tr>
<tr>
<td>Sig. (Unilateral)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NI</td>
<td></td>
<td>,000</td>
<td>,001</td>
</tr>
<tr>
<td>PPNETREV</td>
<td></td>
<td>,000</td>
<td>,000</td>
</tr>
<tr>
<td>OPREV</td>
<td>,001</td>
<td>,000</td>
<td></td>
</tr>
</tbody>
</table>

a. Determinante = ,061

En este caso el análisis de los índices de idoneidad nos aconseja no utilizar el análisis de componentes principales entre estas tres variables, al no existir un nivel de correlación adecuado, pues el índice KMO es de 0,291.

**Tabla 33:** Resultados de los Índices de idoneidad del análisis Factorial

| Medida de adecuación muestral de Kaiser-Meyer-Olkin. | ,291 |
| Prueba de esfericidad de Bartlett | Chi-cuadrado aproximado 21509,940 |
| gl | 3 |
Resultados Índices de idoneidad del Análisis Factorial

| Medida de adecuación muestral de Kaiser-Meyer-Olkin. | .291 |
| Prueba de esfericidad de Bartlett | Chi-cuadrado aproximado | 21509.940 |
| gl | 3 |
| Sig. | .000 |

Estos resultados conducen a interpretar que las variables quizás deberían intervenir en el modelo logit de quiebras bancarias de forma individual, en lugar de utilizando un componente principal que las aglutine conjuntamente.

3.3.1.4 Análisis de componentes principales para las variables de LIQUILITY SUMMARY.

De forma análoga al proceso realizado en apartados anteriores, comenzaremos con el análisis exploratorio de las variables que conforman el grupo de variables de LIQUILITY SUMMARY, a fin de comprobar si es factible llevar a cabo el análisis de componentes principales.

**Tabla 34:** Número de valores válidos y perdidos de las variables del grupo LIQUIDITY SUMMARY

<table>
<thead>
<tr>
<th>Estadísticos</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIQASS ASS</td>
</tr>
<tr>
<td>LIAB</td>
</tr>
<tr>
<td>N Válidos</td>
</tr>
<tr>
<td>Perdidos</td>
</tr>
<tr>
<td>N Válidos</td>
</tr>
</tbody>
</table>
En la Tabla 34 se observa que de las once variables que conforman el grupo de LIQUILITY SUMMARY, únicamente tres variables tienen tres o menos valores perdidos, el resto de variables tienen un número considerable de valores perdidos. Como en los ejemplos anteriores, plantearemos nuevamente dos análisis factoriales:

- Haremos un análisis factorial con las tres variables que tienen 0 y 3 valores perdidos, que son las variables LIQASS_ASS, LIQRAT y LIQASS.

- Por otro lado, también realizaremos el análisis factorial para todas las variables que tengan un número de valores perdidos inferior a 100: LIQASS_ASS, LIQRAT, LOAN_DEP, GROSSL&amp;L_COREDEP L&amp;L/COREDEP y LIQASS

Los resultados de la matriz de correlaciones de las variables que conforman el segundo caso, Tabla 35, son los siguientes:

**Tabla 35**: Matriz de correlaciones de las variables de LIQUILITY SUMMARY seleccionadas como segunda opción

<table>
<thead>
<tr>
<th></th>
<th>LIQASS_ASS</th>
<th>LIQRAT</th>
<th>LOAN_DEP</th>
<th>GROSSL&amp;L_COREDEP</th>
<th>L&amp;L/COREDEP</th>
<th>LIQASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIQASS_ASS</td>
<td>1,000</td>
<td>,348</td>
<td>-.678</td>
<td>-.491</td>
<td>-.491</td>
<td>,047</td>
</tr>
<tr>
<td>LIQRAT</td>
<td>,348</td>
<td>1,00</td>
<td>,175</td>
<td>-.130</td>
<td>-.130</td>
<td>,005</td>
</tr>
<tr>
<td>LOAN_DEP</td>
<td>-.678</td>
<td>,175</td>
<td>1,000</td>
<td>,727</td>
<td>,727</td>
<td>,005</td>
</tr>
</tbody>
</table>
La lectura de esta matriz de correlaciones nos proporciona unos índices de idoneidad de análisis factorial posible, pero con una correlación entre las variables casi media, al aproximarse el índice MKO a un valor de 7, como se puede ver en la Tabla 36.

Tabla 36: Resultados de los Índices de idoneidad del análisis Factorial

<table>
<thead>
<tr>
<th>Índices de idoneidad del Análisis Factorial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determinante de la Matriz de correlaciones</td>
</tr>
<tr>
<td>Medida de adecuación muestral de Kaiser-Meyer-Olkin.</td>
</tr>
<tr>
<td>Prueba de esfericidad de Bartlett</td>
</tr>
<tr>
<td>gl</td>
</tr>
<tr>
<td>Sig.</td>
</tr>
</tbody>
</table>

Los componentes principales extraídos de este grupo de variables de LIQUILITY SUMMARY son los dos primeros con los que se alcanzan el 70,7% del total de la varianza, Tabla 37, y que denominaremos FAC1_LS y FAC2_LS.

Tabla 37: Varianza Total explicada de los componentes de las variables de EARNINGS Summary.

Varianza total explicada
Componente | Autovalores iniciales | Sumas de las saturaciones al cuadrado de la extracción
--- | --- | ---
| | Total | % de la varianza | % acumulado | Total | % de la varianza | % acumulado
1 | 3,147 | 52,445 | 52,445 | 3,147 | 52,445 | 52,445
2 | 1,096 | 18,265 | 70,710 | 1,096 | 18,265 | 70,710
3 | .968 | 16,127 | 86,837 |
4 | .554 | 9,234 | 96,071 |
5 | .236 | 3,929 | 100,000 |
6 | 1,641E-005 | .000 | 100,000 |

Método de extracción: Análisis de Componentes principales.

Las variables que conformarán estos dos componentes principales son los que vienen marcados por los resultados de la matriz de componentes, Tabla 38. Por un lado, el componente principal FAC1_LS estará compuesto por las variables LIQASS_ASS, LOAN_DEP, GROSSL&L_CORDEP y L&L/COREDEP. Por el otro, el componente principal FAC2_LS estará compuesto por LIQRAT y LIQASS.

Tabla 38: Tabla de Componente de las variables EARNINGS Summary, asociadas según la segunda opción metodológica

Matriz de componentes:

<table>
<thead>
<tr>
<th>Componente</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIQASS_ASS</td>
<td>-.751</td>
<td>.355</td>
</tr>
<tr>
<td>LIQRAT</td>
<td>-.305</td>
<td>.682</td>
</tr>
<tr>
<td>LOAN_DEP</td>
<td>.887</td>
<td>-.011</td>
</tr>
<tr>
<td>GROSSL&amp;L_CORDEP</td>
<td>.922</td>
<td>.245</td>
</tr>
<tr>
<td>L&amp;L/COREDEP</td>
<td>.922</td>
<td>.245</td>
</tr>
<tr>
<td>LIQASS</td>
<td>.054</td>
<td>.620</td>
</tr>
</tbody>
</table>
Método de extracción: Análisis de componentes principales.

a. 2 componentes extraídos

Una vez analizadas las variables del grupo EARNINGS Summary según la segunda opción metodológica, pasaremos hacer lo mismo con las variables de la primera opción. Los resultados del caso reducido a las variables con menor número de casos perdidos se analizan a partir de la matriz de correlaciones de las variables LIQASS_ASS, LIQRAT y LIQASS reflejadas en la Tabla 39.

Tabla 39: Matriz de correlaciones de las variables de LIQUILITY SUMMARY seleccionadas como primera opción

<table>
<thead>
<tr>
<th></th>
<th>LIQASS_ASS</th>
<th>LIQRAT</th>
<th>LIQASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlación</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIQASS_ASS</td>
<td>1,000</td>
<td>.300</td>
<td>.044</td>
</tr>
<tr>
<td>LIQRAT</td>
<td>.300</td>
<td>1,000</td>
<td>.001</td>
</tr>
<tr>
<td>LIQASS</td>
<td>.044</td>
<td>.001</td>
<td>1,000</td>
</tr>
<tr>
<td>Sig. (Unilateral)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIQASS_ASS</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>LIQRAT</td>
<td>.000</td>
<td>.477</td>
<td></td>
</tr>
<tr>
<td>LIQASS</td>
<td>.000</td>
<td>.477</td>
<td></td>
</tr>
</tbody>
</table>

a. Determinante = .908

Los resultados de los índices de idoneidad, Tabla 40, nos indican que no es factible la utilización del análisis de componentes principales entre estas tres variables al obtenerse un valor alto del Determinante de la Matriz de correlaciones, 0,908, lo que revela una baja correlación entre ellas. Este hecho viene corroborado por un índice MKO por debajo de 0,5, por lo que al rechazarse la creación de un componente principal que englobe la información de
ellas, se tendrá que probar su importancia a la hora de predecir la posible quiebra de una entidad financiera, dentro del modelo logit, de forma separada.

Tabla 40: Resultados de los Índices de idoneidad del análisis Factorial

Resultados Índices de idoneidad del Análisis Factorial

| Determinante de la Matriz de correlaciones | 0,908 |
| Medida de adecuación muestral de Kaiser-Meyer-Olkin. | .499 |
| Chi-cuadrado aproximado | 741,621 |
| Prueba de esfericidad de Bartlett | gl 3 |
| Sig. | 0,000 |

3.3.1.5 Análisis de componentes principales para las variables de SENSITIVITY SUMMARY

Por último, analizaremos las variables que conforma el grupo de SENSITIVITY SUMMARY dentro del modelo CAMEL, en la Tabla 41:

Tabla 41: Variables que conforman el grupo SENSITIVITY SUMMARY del Modelo CAMEL

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definición</th>
</tr>
</thead>
<tbody>
<tr>
<td>INVPRTDEP</td>
<td></td>
</tr>
<tr>
<td>LTA_ASS</td>
<td></td>
</tr>
<tr>
<td>NONMATDEP_LTA</td>
<td></td>
</tr>
<tr>
<td>RRE_TA</td>
<td></td>
</tr>
<tr>
<td>RATESENSASS</td>
<td></td>
</tr>
<tr>
<td>RATESENSLIAB</td>
<td></td>
</tr>
</tbody>
</table>
El problema que presentan estas variables cuando se realiza el análisis exploratorio, Tabla 42, es el número tan elevado de observaciones perdidas que presentan.

**Tabla 42:** Número de valores válidos y perdidos de las variables del grupo Asset Quality Summary.

<table>
<thead>
<tr>
<th>Estadísticos</th>
<th>INVPOR</th>
<th>LTA_ASS</th>
<th>NONMATD</th>
<th>RE_TTA</th>
<th>RATESE</th>
<th>RATESEN</th>
<th>1YGAP</th>
<th>1YCUMRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Válidos</td>
<td>7033</td>
<td>7033</td>
<td>6921</td>
<td>7033</td>
<td>7033</td>
<td>7033</td>
<td>7033</td>
<td>7033</td>
</tr>
<tr>
<td>Perdidos</td>
<td>684</td>
<td>684</td>
<td>796</td>
<td>684</td>
<td>684</td>
<td>684</td>
<td>684</td>
<td>684</td>
</tr>
</tbody>
</table>

A pesar de este inconveniente, y teniendo presente que el objetivo central de esta investigación es especificar un modelo logit de quiebras bancarias que contenga el mayor número de variables de todos los grupos del Modelo CAMEL, proseguiremos con el análisis de la matriz de correlaciones de estas variables, para comprobar el grado de factibilidad del método de componentes principales. Los resultados de la matriz de correlaciones se recogen en la Tabla 43.

**Tabla 39:** Matriz de correlaciones de las variables de Asset Quality Summary

<table>
<thead>
<tr>
<th>Matriz de correlaciones^ab</th>
</tr>
</thead>
<tbody>
<tr>
<td>INVPORT DEP</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
La matriz de correlaciones es una matriz no definida positivamente, al tiempo que muestra unos niveles de correlaciones entre las variables muy bajo, excepto la relación entre las variables RATESENSLIAB y RATESENSASS con una correlación del 96,8%. Esto nos lleva a rechazar el análisis de componentes principales como método para aglutinar toda la información de las variables del grupo Asset Quality Summary, teniendo que comprobar de forma individual su capacidad predictiva en el modelo logit de quiebras.

### 3.3.2. PLANTEAMIENTO DEL MODELO LOGIT DE QUIEBRAS BANCARIAS NORTEAMERICANAS.

#### 3.3.2.1. MODELO LOGIT DE QUIEBRAS BANCARIAS EN EL AÑO 2008.
Capítulo 4

RESUMEN DE HALLAZGOS Y CONCLUSIONES DE ‘POLICY’