

## ESTIMACIONES DE LA EAD PARA OPERACIONES CON LÍMITES DE CRÉDITO EXPLÍCITO

Gregorio Moral (\*)

(\*) Gregorio Moral pertenece a la Dirección General de Supervisión del Banco de España.



En este artículo se analizan varios métodos de estimación de la EAD en operaciones con límites de crédito explícitos y se estudia su optimalidad desde un punto de vista interno y regulatorio. Se parte de un conjunto de datos de referencia (RDS) obtenido de operaciones que han entrado en *default*. Primero, se definen los factores de conversión observados (CF observados), que son los estadísticos básicos en muchas estimaciones; se describen varios enfoques para calcular estos estadísticos; y se exploran temas previos a la estimación: estructura del RDS, limpieza de los datos y variables explicativas. Centrándose en la EAD, se establece la equivalencia entre los estimadores de la EAD y los de CF y, después, se presentan los métodos usados por los bancos como soluciones de problemas de regresión con una función de pérdida cuadrática y simétrica. A continuación se introduce una clase de funciones de pérdida diferentes, caracterizada por ser lineal y asimétrica, que está relacionada con los requisitos de capital bajo Basilea II y se obtienen estimadores, que podrían ser más apropiados en este ámbito. Finalmente, se presenta un ejemplo ilustrativo. En este caso, los nuevos estimadores propuestos resultan más conservadores que los actualmente empleados por los bancos.

### 1 Introducción

La estimación de la exposición en el momento del *default*, EAD, de una operación con riesgo de crédito ha recibido considerable atención, fundamentalmente, en el ámbito del riesgo de contrapartida y para contratos en los que la variabilidad de la exposición se debe: a la variabilidad del subyacente en el caso de un contrato derivado, a la utilización de una divisa diferente de la divisa de presentación al fijar el importe de la operación o, finalmente, a la existencia de un colateral cuyo valor (variable con el tiempo) reduce la exposición base. Menos atención ha recibido el caso de contratos de crédito con un límite explícito, en los que la fuente de variabilidad es la posibilidad de que el acreditado incremente la exposición mediante disposiciones adicionales amparadas en el límite existente. La implantación de Basilea II está forzando a las entidades de crédito a abordar de una manera rigurosa, transparente y objetiva este último problema. Además, Basilea II impone un conjunto mínimo de condiciones a las estimaciones internas de la EAD para que puedan utilizarse en el cálculo del capital mínimo y, actualmente, las entidades tienen problemas para cumplir con los requisitos referidos tanto a los datos utilizados como a las metodologías aplicadas.

Este artículo analiza varios métodos de estimación de la EAD en operaciones con límites de crédito explícitos y trata de valorar su optimalidad, tanto desde un punto de vista interno como regulatorio. Se centra en métodos objetivos, frecuentemente utilizados por los bancos en la práctica, que se basan en un conjunto de datos de referencia (RDS) obtenido a partir de observaciones de operaciones que han entrado en *default*. La sección 2 presenta la definición de factores de conversión observados (*realised conversion factors*, CF) que son los estadísticos básicos en muchos de los procedimientos de estimación. La sección 3 describe varios enfoques para calcular estos CF: horizonte fijo, cohortes y horizonte variable, y resume sus ventajas e inconvenientes. La sección 4 explora diversos temas que deben abordarse antes de estimar las EAD, como, por ejemplo, la estructura y el alcance del conjunto de datos de referencia, la limpieza de los datos, el tratamiento de las observaciones con CF negativos o mayores que uno y las variables explicativas (RD, *risk drivers*). La sección 5 se

---

1. Este artículo es una versión en castellano del original inglés publicado por B. Engelmann y R. Rauhmeier (eds.) (2006), *The Basel II Risk Parameters: Estimation, Validation, and Stress Testing*, Springer Berlin Heidelberg, Nueva York.

centra en la estimación de la EAD. En primer lugar, se establece la equivalencia entre los estimadores de las EAD y de los CF bajo ciertas condiciones. En segundo lugar, se presentan los métodos más usados por los bancos en la práctica como casos particulares de soluciones a ciertos problemas de optimización. Se concluye que estos métodos son soluciones de problemas de regresión con una función de pérdida cuadrática y simétrica. La sección 6 discute temas relacionados con la optimalidad de las estimaciones e introduce una clase de funciones de pérdida diferentes, caracterizada por ser lineal y asimétrica. Estas funciones de pérdida están relacionadas de forma natural con los requisitos de capital bajo Basilea II y se utilizan para obtener estimadores óptimos que, en consecuencia, podrían ser más apropiados cuando las estimaciones van a ser usadas en el cálculo de los requisitos de capital bajo enfoques avanzados basados en *ratings* internos (AIRB). La sección 7 ilustra aspectos discutidos en las secciones anteriores y las consecuencias de utilizar diferentes métodos de estimación de la EAD con un ejemplo simplificado, pero realista. Finalmente, la sección 8 resume la práctica actual relativa a estimación de EAD y CF, destaca algunos aspectos problemáticos, sugiere posibles mejoras y concluye que los métodos tradicionales, basados en promedios, son menos conservadores que los métodos basados en cuantiles a la derecha de la mediana.

## 2 Definición de los factores LEQ observados

En la práctica, a la hora de estimar la EAD de una operación viva  $f$  (que no ha entrado en *default*) con un límite de crédito explícito<sup>2</sup>, hay dos clases de métodos principales atendiendo a la ecuación básica que se utiliza para relacionar la estimación con el límite:

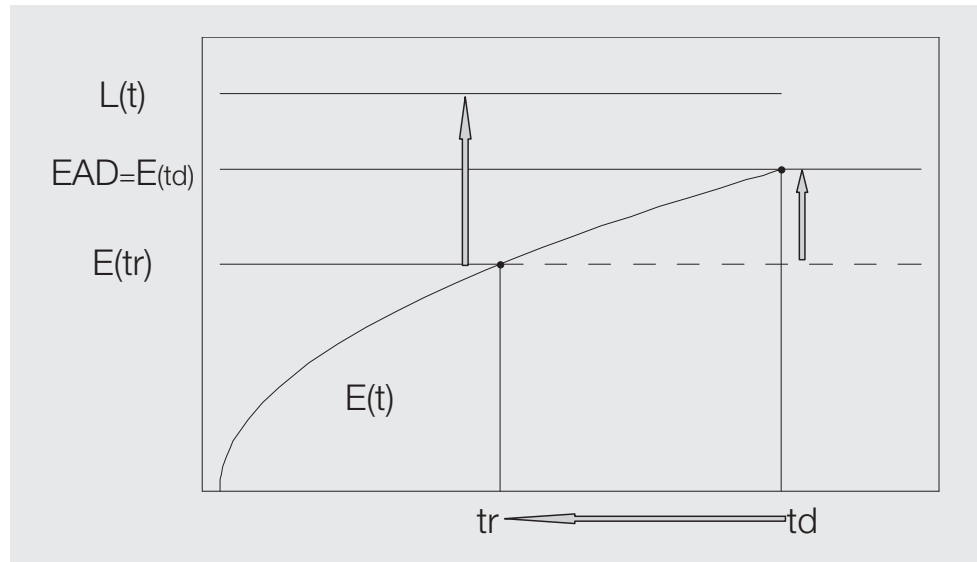
- En la primera clase, las estimaciones de la EAD se basan en un factor de conversión de crédito adecuado para el total del límite de la operación,  $EAD(f) = CCF(f) * Límite(f)$ .
- En la segunda clase, las estimaciones de la EAD se basan en otro factor<sup>3</sup> que se aplica a la parte del límite no utilizada,  $EAD(f) = E(f) + LEQ(f) * Límite no utilizado(f)$ <sup>4</sup>.

Como se verá en la sección 5, los dos enfoques son esencialmente equivalentes, y el problema de estimación de la  $EAD(f)$  puede reducirse a la estimación de factores de conversión de crédito adecuados  $CF(f)$  ( $CCF(f)$  o  $LEQ(f)$ ).

Para obtener las estimaciones de los CF se utiliza, como datos básicos, un conjunto de observaciones, en ciertas fechas anteriores a la de *default*, de operaciones que han entrado en *default*. La mayor parte de los métodos utilizados por los bancos para estimar estos CF se basa en ciertos estadísticos relacionados con el incremento en el uso<sup>5</sup> del límite de crédito entre una fecha de referencia y la fecha de *default* calculados a partir de las observaciones anteriores. Uno de estos estadísticos se denomina factor LEQ observado y su definición se presenta a continuación.

---

2. Por ejemplo, líneas de crédito comprometidas, en las que el acreditado puede disponer de fondos hasta que se alcanza un límite  $L(t)$ . 3. En el Nuevo Marco y en la Directiva de Capital tales factores se denominan factores de conversión de crédito (*Credit Conversion Factors*, CCFs) y factores de conversión (*Conversion Factors*, CFs), respectivamente. En los borradores de las Reglas de Implementación de Basilea II en Estados Unidos, el factor utilizado se denomina LEQ y el documento de Guías del CEBS usa el término *Conversion Factors* (CFs). En este artículo, por claridad, los factores de conversión que se aplican a la parte no dispuesta del límite se denominan factores LEQ (*Loan Equivalent*, LEQ factors) y el término CCF se reserva para los factores que se aplican al límite total. 4. Este es el enfoque exigido para este tipo de operaciones en el Nuevo Marco, en la Directiva de Capital, en los borradores de las Reglas de Implementación de Basilea II en Estados Unidos y en el documento de Guías del CEBS. 5. En lo sucesivo, el término uso o utilización del límite se refiere al dispuesto en euros (algunas veces se utilizan los términos exposición, importe dispuesto o importe utilizado con el mismo significado).



Considérese una operación  $g$ , con una exposición variable en función del tiempo, representada por  $E(t)$  y un límite de crédito dado por  $L(t)$ , que ha entrado en *default*. La figura 1 presenta la evolución de la exposición.

Si la operación entró en *default* en la fecha  $td$ , dada una fecha de referencia  $tr < td$ , se denomina índice de la observación a la pareja  $\{g, tr\} = i$ . Si  $EAD_i$  denota a la exposición observada<sup>6</sup> en el momento del *default*, esta se puede expresar en función de la exposición y del límite de la operación observados a la fecha de referencia, suponiendo que  $L(tr) \neq E(tr)$ , como:

$$EAD_i = E(tr) + LEQ_i * (L(tr) - E(tr)) \tag{1}$$

donde,  $LEQ_i$  viene dado por:

$$LEQ_i = \frac{E(td) - E(tr)}{L(tr) - E(tr)} \tag{2}$$

o:

$$LEQ_i = \frac{\frac{E(td)}{L(tr)} - \frac{E(tr)}{L(tr)}}{1 - \frac{E(tr)}{L(tr)}} = \frac{ead_i - e(tr)}{1 - e(tr)} \tag{3}$$

Por lo tanto, dada una observación, caracterizada por un par  $i = \{g, tr\}$ , con  $L(tr) \neq E(tr)$ , las fórmulas anteriores se pueden usar para obtener un factor LEQ observado. En general, el factor LEQ asociado a una observación se denota por  $LEQ_i$ . Por comodidad, cuando se quiera destacar la fecha de referencia  $tr$ , el  $LEQ_i$  se denota por  $LEQ(tr)$ .

6. En este artículo se supone que existe previamente una definición precisa de la EAD observada de las operaciones que entran en *default* y que dicha definición se aplica de manera uniforme entre las diferentes operaciones y a lo largo del tiempo para los diferentes usos internos. Para entender por qué es necesaria una definición explícita de la EAD observada, véase Araten (2001), p. 37, donde se citan dos situaciones en las que la definición simple («importe final observado en el momento del *default*») no es adecuada debido a la existencia de reducciones de deuda, previas al *default*, por reconocimientos de pérdida o por adjudicaciones de colateral.

Hay tres limitaciones al usar este estadístico como el dato básico para los procedimientos de estimación:

- No está definido cuando  $L(tr) = E(tr)$ . Esto implica que no es posible estimar directamente la  $EAD(f)$  basándose en el valor de este estadístico para operaciones que en el momento actual presentan una utilización porcentual,  $e(tr)$ , igual a uno<sup>7</sup>.
- No es estable cuando  $L(tr) \cong E(tr)$ . Esto significa que los factores LEQ observados no contienen mucha información útil cuando la utilización porcentual es casi uno. Como se muestra en la sección 4.2.2, el diferente comportamiento de los factores LEQ observados dependiendo del nivel de utilización porcentual del crédito en la fecha de referencia tiene importantes consecuencias prácticas.
- No tiene en cuenta los cambios en el límite de crédito a lo largo del tiempo. En las fórmulas [2] y [3], los factores LEQ observados se han definido sin tener en cuenta posibles cambios en el límite de la operación entre la fecha de referencia y la fecha de *default*<sup>8</sup>. Como se muestra en detalle en la sección 4.2.3, esta es solo una de las causas que justifican la existencia de factores LEQ<sub>i</sub> mayores que uno.

Debido a estas razones, los bancos a veces utilizan otros estadísticos como punto de partida para la estimación de la EAD. Por ejemplo, una posibilidad evidente es definir un factor CCF observado,  $CCF_i$ , de forma similar a como se ha definido el factor  $LEQ_i$ . Utilizando una ecuación análoga a [1], la expresión para este estadístico viene dada por la exposición porcentual en el momento del *default*:

$$CCF_i = \frac{EAD_i}{L(tr)} = ead_i \quad [4]$$

Aunque, en la práctica, este estadístico es menos usado que el factor  $LEQ_i$  para este tipo de operaciones, presenta dos ventajas:

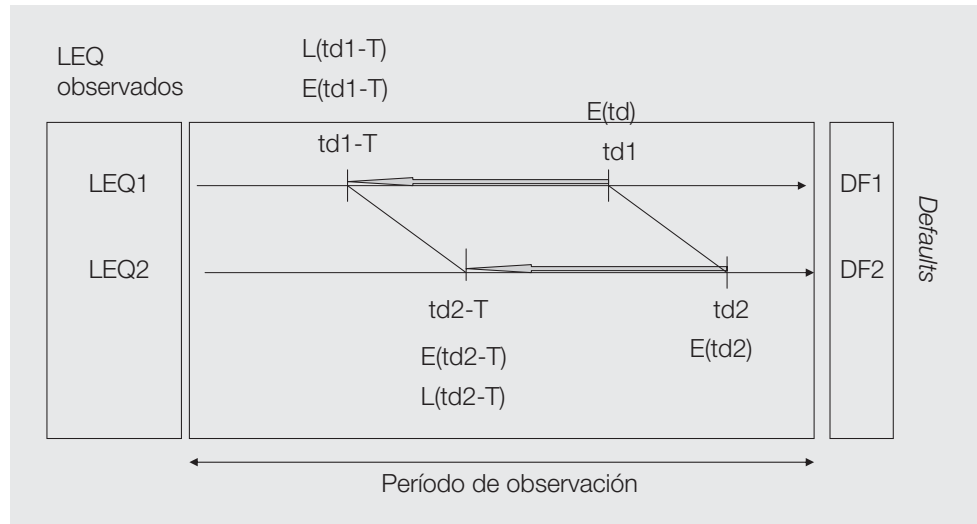
- El factor CCF observado está bien definido incluso si  $L(tr) = E(tr)$ .
- Es estable, aunque ocurra que  $L(tr) \cong E(tr)$ .

Algunas veces se dice que con ese estadístico, si la operación  $g$  tiene un límite constante  $L(g)$ , no es necesario especificar una fecha de referencia. Sin embargo, como se muestra en la sección 4, los conjuntos de datos para los procedimientos de estimación necesitan incluir los valores de ciertas *variables explicativas* cuyo valor varía con el tiempo y, por lo tanto, es necesario considerar una fecha de referencia explícita.

Otros estadísticos útiles se presentarán en la sección 5; hasta entonces se supone que los factores LEQ observados se utilizan como el punto de partida para los procesos de estimación.

---

7. Esto es una restricción cuando las estimaciones se van a usar para usos internos, ya que, en principio, para estos usos internos no es necesario suponer que la estimación del factor LEQ para una operación,  $LEQ(f)$ , tenga que ser positivo o nulo, o, de modo equivalente, que la estimación de  $EAD(f)$  tenga que ser mayor o igual que su exposición actual,  $E(f)$ . 8. Algunos bancos definen los factores LEQ observados utilizando  $E(td)/L(td)$  en vez de  $ead_i = E(td)/L(tr)$  en la fórmula [3]. El propósito de esa definición es tener en cuenta los cambios en el límite de crédito habidos después de la fecha de referencia y evitar la obtención de factores LEQ observados mayores que uno. No presenta ninguna complicación ver que la definición anterior es compatible con la ecuación [1] si la EAD se multiplica por el factor  $L(tr)/L(td)$ .



### 3 Cómo obtener un conjunto de factores LEQ observados

Dado un conjunto de operaciones que han entrado en *default*, hay varios enfoques empleados normalmente por los bancos para obtener los factores  $LEQ_i$  u otros estadísticos<sup>9</sup> que puedan ser usados, junto con otra información, para obtener estimaciones de la EAD de operaciones vivas. Todos estos enfoques se basan en observaciones, en determinadas fechas de referencia, de operaciones que han entrado en *default*. Dependiendo de la regla utilizada para seleccionar esas fechas de referencia, estos enfoques se denominan: horizonte fijo, cohortes u horizonte variable.

#### 3.1 ENFOQUE DE HORIZONTE FIJO

En este enfoque, primero se selecciona un horizonte,  $T$ , y después, para cada operación que ha entrado en *default* que verifica  $L(td - T) \neq E(td - T)$ , se calcula un factor  $LEQ_i$  usando como fecha de referencia  $td - T$ :

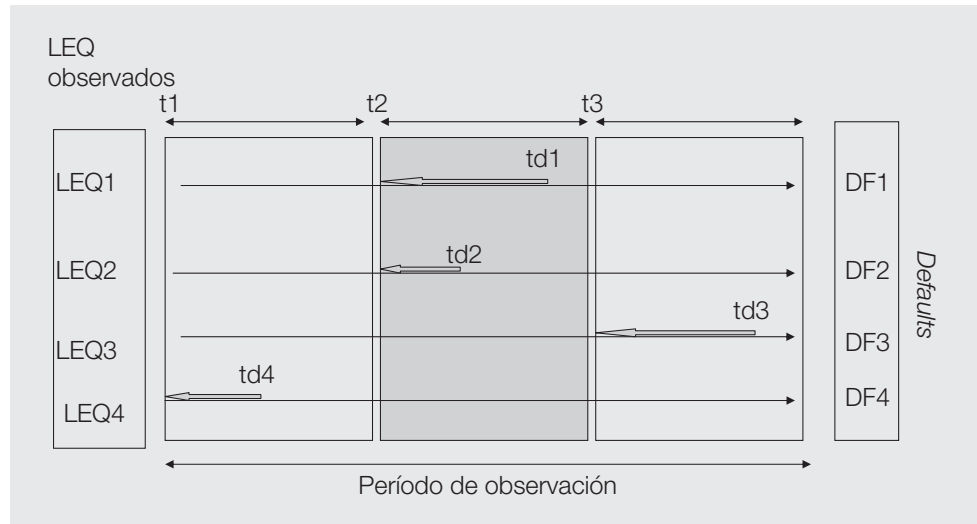
$$LEQ(td - T) = \frac{E(td) - E(td - T)}{L(td - T) - E(td - T)} \quad [5]$$

En la práctica,  $T$  se fija frecuentemente a un año.

Inconvenientes:

- El horizonte fijo,  $T$ , es convencional.
- No es posible incluir directamente operaciones que han entrado en *default* cuando la edad de la operación en la fecha del *default* es menor que  $T$ .
- No se tiene en cuenta toda la información relevante, ya que, para cada operación  $g$  que ha entrado en *default* durante el período de observación, solo se utiliza una observación dada por  $\{g, td - T\}$ .
- No se tiene en cuenta la posibilidad de que las exposiciones vivas entren en *default* en cualquier momento durante el año siguiente. Implícitamente, las estimaciones

<sup>9</sup> Como se muestra en la sección 5, además de los factores LEQ observados, el incremento porcentual de la utilización del crédito entre la fecha de referencia y la fecha de *default* o el incremento en la exposición entre esas fechas son estadísticos que se pueden usar para estimar los CF y las EAD.



basadas en este enfoque asumen que la fecha de *default* para aquellas operaciones que entren en *default* durante los próximos doce meses será el final de ese intervalo. Este supuesto puede introducir sesgos en las estimaciones.

Ventajas:

- Dispersión de las fechas de referencia.
- El uso de un horizonte común,  $T = td - tr$ , contribuye a la homogeneidad de los factores  $LEQ_i$  seleccionados.

### 3.2 ENFOQUE DE COHORTES

En primer lugar, el período de observación<sup>10</sup> se divide en intervalos de duración fija (cohortes); por ejemplo, intervalos de un año. En segundo lugar, las operaciones se agrupan en cohortes atendiendo al intervalo que incluye la fecha de *default*. En tercer lugar, para calcular un factor  $LEQ_i$  observado asociado a cada operación, se utiliza como fecha de referencia el inicio del intervalo que contiene la fecha de *default*,  $\{t_1, t_2, \dots, t_i, \dots, t_n\}$ :

$$LEQ(t_i) = \frac{E(td) - E(t_i)}{L(t_i) - E(t_i)} \quad [6]$$

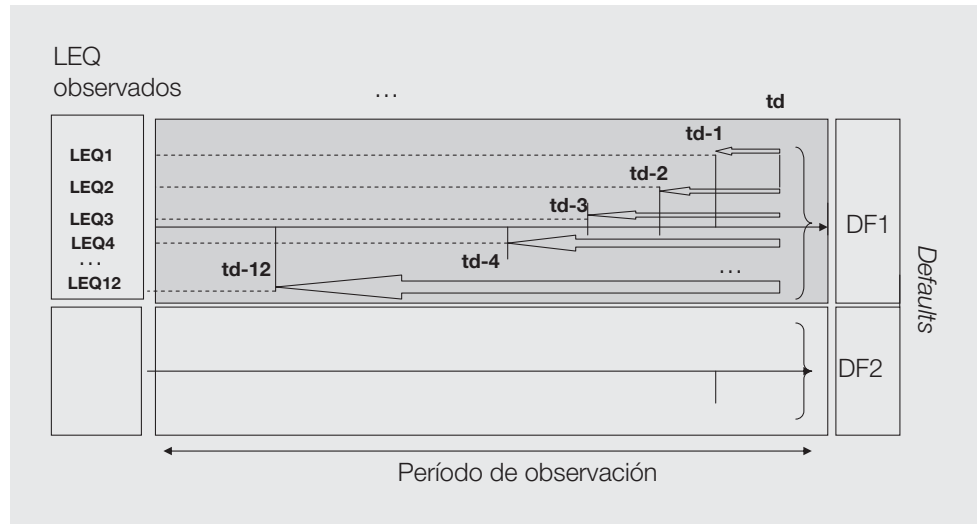
Esto se ilustra en la figura 3.

Inconvenientes:

- La duración de las cohortes es convencional.
- Las fechas de referencia son convencionales.
- No se utiliza toda la información relevante, ya que para cada operación  $g$ , que ha entrado en *default* durante el período de observación, solo se utiliza una observación dada por  $\{g, t_j\}$ , donde  $j$  depende de la cohorte que contiene a la  $td$ .

<sup>10</sup> El período de tiempo que cubre los datos es el período de observación.





– Las fechas de referencia están concentradas.

– Los factores  $LEQ_i$  observados son menos homogéneos que los que se obtienen utilizando el mismo horizonte. La razón fundamental es que se han calculado factores  $LEQ_i$  con horizontes (valores de  $td - tr$ ) muy diferentes.

Ventajas:

– Tiene en cuenta la posibilidad de que las exposiciones vivas puedan entrar en *default* en cualquier momento durante el siguiente año.

### 3.3 ENFOQUE DE HORIZONTE VARIABLE

En primer lugar, se fija un rango para el horizonte al que se quiere calcular factores  $LEQ_i$  (normalmente, un año). En segundo lugar, para cada operación que ha entrado en *default* se calculan factores  $LEQ_i$  observados asociados con un conjunto de fechas de referencia (por ejemplo<sup>11</sup>, un mes, dos meses, ..., doce meses antes de la fecha de *default*), que recorren el rango de horizontes.

La lógica de este método es tener en cuenta un conjunto más amplio que en los otros enfoques de posibles fechas de *default* cuando se vaya a estimar un factor LEQ para una operación viva condicionado a que el *default* tenga lugar durante el próximo año.

$$\{LEQ(td - j) = \frac{E(td) - E(td - j)}{L(td - j) - E(td - j)}, \quad j = 1, \dots, 12 \text{ meses}\} \quad [7]$$

En principio<sup>12</sup>, se pueden asociar doce factores  $LEQ_i$  observados a cada operación que haya entrado en *default*. Sin embargo, estos factores  $LEQ_i$  son claramente no homogéneos, ya que algunos de estos valores se han calculado utilizando observaciones muy cercanas a la fecha de *default* ( $i = \{g, td - 1\}$ ) y otras están basadas en observaciones un año antes del *default*

<sup>11</sup> Aun cuando en este enfoque, en teoría, no es necesario utilizar observaciones mensuales, de ahora en adelante se supone que las fechas de referencia son los finales de mes desde el primer final de mes anterior a la fecha de *default* ( $td - tr = 1$ ) hasta doce meses antes ( $td - tr = 12$ ). Esta elección puede ser adecuada para la mayor parte de los productos y, en muchos casos, compatible con la información que los bancos ya almacenan actualmente. <sup>12</sup> Por ejemplo, si una operación tiene solo cuatro meses de vida cuando pasa a *default*, entonces, como mucho, se podrán calcular cuatro factores LEQ observados asociados a ella directamente.

( $j = \{g, td - 12\}$ ). Esto significa que es necesario reconocer estas diferencias vía *variables explicativas*. Como se muestra en la sección 4.3, el punto clave es tener en cuenta en qué momento el banco identificó dichas observaciones como asociadas a operaciones en una situación no normal. La principal razón es que, cerca de la fecha de *default*, los acreditados están en general clasificados internamente en un categoría (que en adelante denominaremos estatus) no normal. Esto significa que esas operaciones están sujetas a una vigilancia especial y, en general, el acreditado no puede hacer disposiciones adicionales en las mismas condiciones que anteriormente. Por ejemplo, en carteras minoristas, durante los tres meses anteriores al *default*, desde el momento en el que se produce el primer impago, es muy difícil para el acreditado hacer disposiciones adicionales y, en general, solo se pueden cargar en la cuenta liquidaciones de intereses y otros cargos internos. Por lo tanto, es necesario identificar cuándo fue marcada como no normal una operación que ha entrado en *default* y utilizar solamente los factores LEQ observados asociados a fechas anteriores, cuando se estimen factores LEQ para operaciones vivas normales. En la práctica, para carteras minoristas, frecuentemente se pueden utilizar como mínimo seis fechas y como máximo nueve. Por otra parte, en carteras de empresas, el estatus de las operaciones está muy ligado al *rating* del deudor y, por lo tanto, puede haber casos en los que la operación está en un estatus normal hasta que se conoce que el deudor ha entrado en *default*.

En general, es conveniente tener en cuenta de forma separada los doce factores  $LEQ_i$  asociados a la misma operación, ya que los valores de las *variables explicativas* pueden ser diferentes en cada una de las fechas de referencia.

Ventajas:

- Tiene en cuenta más observaciones que los métodos anteriores.
  - Las operaciones con  $L(tr) = E(tr)$ , que en los métodos anteriores no se podían tener en cuenta, ahora se pueden usar para aquellas fechas de referencia en las que  $L(td - i) \neq E(td - i)$  para  $i = 1, \dots, 12$ .
  - Cada operación puede proporcionar hasta doce factores  $LEQ_i$  observados asociados con doce observaciones diferentes.
- En principio, los procedimientos de estimación basados en este enfoque deberían producir estimaciones más estables (se usan más observaciones) y más precisas (se utiliza más información).

Inconvenientes:

- Los bancos tienen que almacenar más datos para cada operación que ha entrado en *default* (hasta doce observaciones).
- Es necesario utilizar una variable (estatus) que contribuya a identificar factores  $LEQ_i$  homogéneos.

#### **4 Conjuntos de datos (RDS) para los procedimientos de estimación**

En esta sección se discuten los requisitos ideales para el conjunto de datos de referencia (RDS) que incluye la información disponible que se puede usar en los procedimientos de estimación. Se centra en los RDS construidos a partir de información histórica obtenida de un conjunto de operaciones que han entrado en *default* durante un período de observación. En primer lugar, se presenta una estructura general para este RDS que facilita la implantación de

procedimientos de estimación y después se enumeran algunos campos que deberían incluirse en el RDS. En segundo lugar, se listan ciertos requisitos de alcance. Finalmente, se comentan ciertos ajustes y decisiones que deben tomarse antes de la fase de estimación.

#### 4.1 ESTRUCTURA Y ALCANCE DEL CONJUNTO DE DATOS DE REFERENCIA

##### 4.1.1 Estructura

Dado el interés en métodos de estimación basados en observaciones en ciertas fechas de referencia, de operaciones que han entrado en *default*, es útil tener una estructura para el conjunto de datos de referencia adaptada a este enfoque. Consecuentemente, la estructura de datos debería contener la información relevante a partir de observaciones,  $O_i$ , que tienen asociado un único par formado por una operación,  $g$ , que ha entrado en *default* y una fecha de referencia válida,  $tr < td$  [más específicamente, el par anteriormente mencionado  $i = (g, tr)$  debería ser la clave primaria del conjunto de datos de referencia]. Cada una de estas observaciones,  $O_i$ , incluye:

- Los valores de ciertas características estáticas de  $g$ ,  $I(g)$ .
- Los valores de un conjunto de variables observables relacionadas con  $g$  a la fecha de referencia  $tr$ , que van a ser usadas como variables explicativas o *risk drivers*,  $RD(tr)$ ,
- Los valores observados de la EAD y de la fecha de *default* asociadas a la operación:  $EAD_i$ ,  $td$ .

En resumen, una estructura muy general para el RDS viene dada por:

$$RDS = \{O_{i=(g,tr)}\}; O_{i=(g,tr)} = \{(g, tr), I(g), RD(tr), EAD_i = E(td)\} \quad [8]$$

Respecto a los campos que contienen la información asociada a cada una de las observaciones, en la práctica, como mínimo, se requieren los siguientes datos:

- Características estáticas,  $I(g)$ : identificador de la operación, NF; tipo de operación, TF; identificador de cartera, TP; e identificador de acreditado, NB.
- Variables explicativas,  $RD$ : fecha de referencia,  $tr$ ; fecha de *default*,  $td$ ; exposición en la fecha de referencia,  $E(tr)$ ; límite en la fecha de referencia,  $L(tr)$ ; estatus de la operación,  $S(tr)$ ; y clase de *rating* o *pool*,  $R(tr)$ .

Si hubiera identificadas otras variables explicativas de la EAD, el RDS debería contener campos para esos valores a la fecha de referencia  $tr$ . Por ejemplo, vale la pena incluir indicadores macroeconómicos, MI, que se puedan utilizar para aumentar el carácter *forward looking* de las estimaciones y la capacidad predictiva de los estimadores. En símbolos:

$$RD(tr) = \{E, L, S, R, td, MI, Otras\}; I(g) = \{NF, NB, TF, TP, Otras\} \quad [9]$$

Las variables explicativas se discuten con más detalle en la sección 4.3.

##### 4.1.2 Alcance y otros requisitos del RDS

Además de una estructura del RDS adecuada para los procedimientos de estimación, el RDS tiene que cumplir con ciertos requisitos internos y externos relacionados con el alcance del RDS.

- El alcance del RDS tiene que estar definido sin ambigüedad. Como mínimo, es necesario:

- Definir el tipo de operaciones, el tipo de acreditados y el tipo de carteras.
  - Explicitar la definición de *default* utilizada y el período de observación cubierto.
  - Identificar y describir la fuente (o fuentes) de los datos.
- El RDS debería incluir observaciones de todas las operaciones que han entrado en *default* durante el período de observación y que cumplan con el resto de requisitos de alcance (tipo de operación, cartera, etc). Todas las exclusiones deberían estar identificadas y justificadas.
- La definición de *default* utilizada debería ser coherente con las definiciones utilizadas para las estimaciones de las PD y las LGD.
- El período de observación debería ser suficientemente largo para incluir observaciones de operaciones que han entrado en *default* bajo circunstancias macroeconómicas muy diferentes, idealmente cubriendo un ciclo económico completo.
- Además, para utilizar las estimaciones en el cálculo de los requisitos de capital bajo enfoques AIRB:
- La definición de *default* utilizada debería ser compatible con la definición de *default* para enfoques IRB.
  - El período de observación debe cubrir, como mínimo, siete años para carteras de empresas y cinco años para carteras minoristas.
  - Cuando sea necesario, el período de observación debería contener un período bajo condiciones *downturn*.

## 4.2 LIMPIEZA DE DATOS

Además de otros temas más generales relacionados con las tareas de limpieza de datos (identificación y tratamiento de *outliers*, eliminación de datos de mala calidad, etc.), antes de la fase de estimación es necesario tomar ciertas decisiones que pueden afectar a las observaciones incluidas en el RDS. Algunos de estos temas se analizan en las próximas secciones.

### 4.2.1 Tratamiento de operaciones diferentes con el mismo acreditado

Aunque es claro que los factores LEQ<sub>i</sub> y el resto de la información relevante contenida en el RDS se calculan o se observan a nivel de operación, bajo ciertas circunstancias, para obtener estimaciones razonables, puede ser necesario o conveniente agrupar en la misma observación información de diferentes operaciones con el mismo acreditado. Como mínimo, hay dos situaciones que cabe considerar:

- Si hay dos o más observaciones de operaciones similares con el mismo acreditado y con los mismos valores de las variables explicativas, excluido el valor de la utilización del crédito y otros valores que sean función de L(t) y de E(t), entonces puede ser conveniente agrupar esas observaciones en una nueva, de la forma siguiente<sup>13</sup>:

$$\begin{aligned} & \left\{ (h, tr), E(h, tr), L(h, tr), B(h), RD(h, tr) \right\} \\ & \left\{ (g, tr), E(g, tr), L(g, tr), B(g), RD(g, tr) \right\} \Rightarrow \{ (h + g, tr) \} \\ & \{ (h + g) \} = \{ (h + g, tr), E(h, tr) + E(g, tr), L(h, tr) + L(g, tr), B, RD(tr) \} \end{aligned} \quad [10]$$

13. Este procedimiento es descrito en Araten (2001), p. 36.

- Para ciertas carteras y ciertos tipos de operaciones es frecuente que la duración de las operaciones sea de un año. No obstante, en muchos de los casos el banco aprueba una nueva operación (quizá con un límite diferente) cuando vence la operación previa. En estas circunstancias, las operaciones siempre entran en *default* con menos de doce meses y, por lo tanto, no es posible obtener doce observaciones para el RDS. Sin embargo, si esa operación se aprobó al vencimiento de una operación del mismo tipo (que no entró en *default*), con el mismo acreditado, entonces podría ser útil encadenar esas operaciones. Mediante ese procedimiento se pueden incluir más observaciones en el RDS.

Dependiendo de las características de la cartera, estas decisiones podrían tomarse mediante un análisis caso a caso o siguiendo una regla.

4.2.2 Tratamiento de las observaciones con factores LEQ observados negativos<sup>14</sup>

Como muestra la figura 5, es posible obtener factores LEQ observados negativos asociados con operaciones que han entrado en *default*.

Aritméticamente, los factores  $LEQ_i$  negativos aparecen cuando  $EAD = E(td) < E(tr)$ , y esta situación es especialmente frecuente cuando  $td - tr$  es grande y cuando la utilización porcentual del crédito a la fecha de referencia  $e(tr)$  es próxima a uno; además, hay factores  $LEQ_i$  con valores absolutos muy grandes. Es muy importante notar que:

- Las distribuciones empíricas de los factores  $LEQ_i$  condicionadas a la utilización porcentual en la fecha de referencia  $e(tr)$  son muy diferentes.
- Estas distribuciones empíricas son muy asimétricas, especialmente para valores de la utilización porcentual del crédito próximos a uno.

Para ilustrar estos puntos de la definición [3] se puede ver que un pequeño incremento de  $e(tr)$  afecta al factor  $LEQ_i$  proporcionalmente a:

$$\frac{\partial LEQ_i}{\partial e(tr)} = -\frac{1 - e(td)}{(1 - e(tr))^2} \quad [11]$$

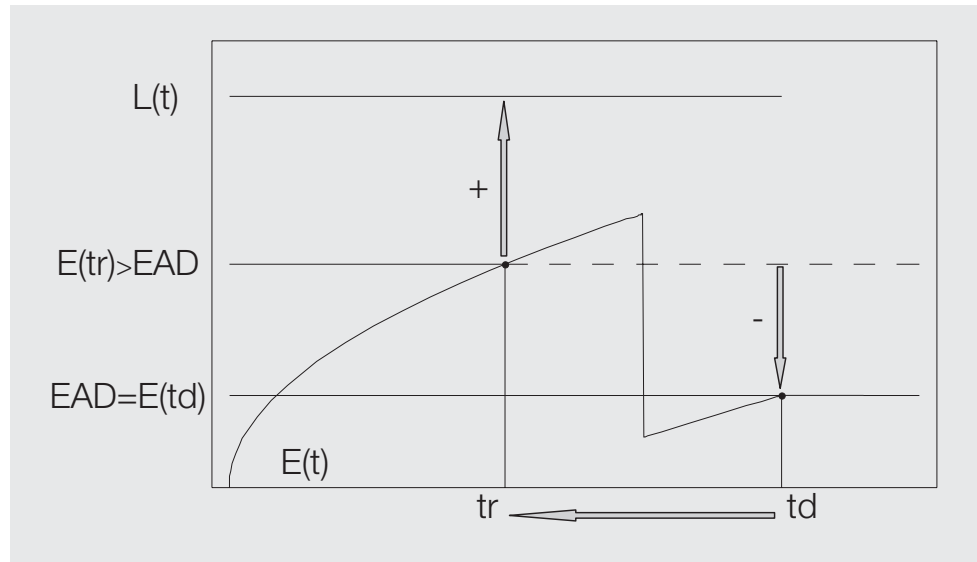
Por lo tanto, la sensibilidad del factor  $LEQ_i$  frente a pequeños cambios en la utilización porcentual del crédito en la fecha de referencia depende crucialmente del nivel de  $e(tr)$ . Cuanto más pequeño es  $(1 - e(tr))^2$ , mayor tiende a ser la variabilidad de los factores  $LEQ_i$  condicionada a  $e(tr)$ .

Más aún, si el factor  $LEQ_i$  se expresa en términos de la utilización porcentual en el momento del *default* en proporción a la utilización en la fecha de referencia, de la definición [3] se obtiene:

$$LEQ_i(\Delta) = \frac{e(tr) \cdot (1 + \Delta) - e(tr)}{(1 - e(tr))} = \Delta \frac{e(tr)}{1 - e(tr)} \quad [12]$$

y para valores grandes de  $e(tr)$  no hay posibilidad de valores grandes de  $\Delta$ ; por el contrario, sí es posible encontrar valores negativos de  $\Delta$  grandes en valor absoluto.

<sup>14</sup> Desde un punto de vista formal, esta discusión es similar a la discusión relativa a las observaciones de LGD negativas. Sin embargo, existen diferencias sustanciales en las razones que justifican la existencia de valores observados negativos en uno y otro caso.



Las anteriores asimetrías entre factores  $LEQ_i$  (para valores bajos y altos de la utilización porcentual) y la existencia de más observaciones con factores  $LEQ_i$  negativos (y grandes en valor absoluto) que observaciones con factores  $LEQ_i$  positivos grandes tienen importancia práctica. La principal razón es que, como se muestra en la sección 5, los bancos usan con frecuencia promedios de los factores  $LEQ_i$  como estimadores de  $LEQ(f)$  y esas medias muestrales están muy afectadas por ambas circunstancias. Los puntos anteriores sugieren que, como mínimo, estos promedios deberían estar restringidos a aquellas observaciones con niveles semejantes de utilización porcentual del crédito o, en otras palabras, que el nivel de utilización porcentual del crédito debería ser una variable explicativa de  $LEQ(f)$ .

Como consecuencia, es importante aclarar el tratamiento de aquellas observaciones con valores negativos de los factores  $LEQ_i$  observados. En la práctica, hay varias posibilidades:

- Censura<sup>15</sup> de los datos (los factores  $LEQ_i$ ) para imponer ciertas restricciones:
  - Algunos bancos cambian la definición de los factores  $LEQ_i$  observados para forzar la no negatividad:  $LEQ_i^+ = \max[0, LEQ_i]$
  - En otros casos, los bancos cambian directamente la definición de la EAD observada utilizada en el cálculo de los factores  $LEQ_i$ :  $EAD_i^+ = \max[EAD_i, E(tr)]$

Como se ha discutido previamente, observaciones válidas de operaciones que han entrado en *default* pueden tener asociados valores negativos de los factores  $LEQ_i$ . Para justificar esta práctica (la censura), algunos bancos argumentan que, ceteris paribus, este ajuste introduce un sesgo conservador en las estimaciones.

- Truncamiento: esta práctica consiste en la eliminación de las observaciones asociadas con factores  $LEQ_i$  negativos. Es difícil encontrar una justificación para el

<sup>15</sup> Es necesario tener cuidado con el uso de esta nomenclatura (censura y truncamiento), ya que estas palabras no se usan de manera uniforme en la literatura. Por ejemplo, Araten (2001), p. 36, usa el término truncamiento para describir lo que en este artículo se denomina censura. La terminología empleada en el texto sigue la utilizada en el Working Paper No.14 BCBS (2005), p. 66.

truncamiento de dichas observaciones. En principio, este truncamiento podría ser un método práctico para generar una distribución estresada de factores LEQ<sub>i</sub>. Sin embargo, este procedimiento presenta, como mínimo, dos importantes inconvenientes:

- La eliminación de las observaciones con  $LEQ_i \leq 0$  podría introducir incoherencias con el RDS utilizado para obtener las estimaciones de las LGD, debido a que algunas de esas observaciones podrían estar asociadas con operaciones con pérdidas observadas altas.
  - Cuando el método estimación usa medias muestrales, las estimaciones de los factores LEQ(f) basadas en un RDS truncado pueden ser muy inestables frente a cambios en RDS, dependiendo del número de observaciones con factores LEQ<sub>i</sub> próximos a cero.
- No hacer nada con los factores LEQ<sub>i</sub> (pero fijar un suelo para las estimaciones<sup>16</sup>,  $LEQ(f) \geq 0$ ). Esta es la decisión más natural.

Como se prueba en la sección 6.3.1, si se impone la restricción  $LEQ(f) \geq 0$  a los estimadores y se ajusta un modelo determinado basado en la minimización de los errores de estimación (medidos en términos de una función de pérdida especial), entonces se obtienen las mismas estimaciones cuando se utilizan los datos originales que cuando se usan los datos censurados.

#### 4.2.3 Tratamiento de observaciones con factores LEQ<sub>i</sub> observados mayores que uno

En principio, dada la definición de los factores LEQ observados [2], sería natural esperar que dichos factores LEQ<sub>i</sub> fueran menores o iguales que uno, en un banco con entorno de control adecuado. Sin embargo, la existencia de factores LEQ<sub>i</sub> mayores que uno no es en todos los casos un indicador de fallos en los controles establecidos por el banco para asegurar que los límites de crédito establecidos son efectivos. Hay situaciones en las que factores LEQ<sub>i</sub> mayores que uno aparecen de forma natural. Por ejemplo:

- En algunos casos, los bancos usan límites internos (no comunicados a los clientes, *unadvised limits*)<sup>17</sup> en vez de los límites nominales de las operaciones para gestionar internamente el riesgo. La posibilidad de disposiciones adicionales del crédito solo cesa cuando la exposición es mayor que el límite interno.
- En algunos productos (por ejemplo, tarjetas de crédito o descubiertos en cuentas corrientes), esos problemas son difíciles de evitar, ya que, típicamente hay un retraso entre la exposición real y la usada por el banco para establecer controles.
- Algunas veces, la exposición en el momento del *default* incluye la última liquidación de intereses (y comisiones), y esta cantidad se carga en la cuenta, incluso cuando el límite ha sido alcanzado previamente.

Los anteriores excesos sobre los límites nominales son, en general, pequeños. En estas situaciones, puede ser apropiado tratar estas observaciones como cualquier otro caso. Sin embargo, en otras circunstancias hay observaciones con factores LEQ<sub>i</sub> grandes que son el resultado de otras causas totalmente diferentes, como, por ejemplo:

---

<sup>16</sup>. Como mínimo, este suelo es un requisito cuando las estimaciones se van a usar para fines regulatorios. <sup>17</sup>. Frecuentemente, estos límites *unadvised* se calculan como un porcentaje o una cantidad fija por encima de los límites comunicados.

- Cambios en los límites después de la fecha de referencia y antes del conocimiento de dificultades en la operación.
- Cambios explícitos o implícitos en los límites en el momento del *default* o cuando las dificultades en la operación ya han aparecido.
- Inadecuado entorno de control y existencia de errores o fraudes que podrían ser tratados como eventos de riesgo operacional.

A pesar de la diversidad de las circunstancias anteriores, algunos bancos acotan superiormente todos los factores  $LEQ_i$  a uno. En general, esta regla no es adecuada ni para fines internos ni para los regulatorios, y, por el contrario, se necesita un análisis detallado de las causas que dan lugar a esas observaciones antes de tomar una decisión aceptable. En cualquier caso, un requisito es la coherencia con los procedimientos utilizados para estimar las LGD.

#### 4.3 VARIABLES EXPLICATIVAS DE LA EAD

En la práctica, las variables explicativas (RD) afectan a las estimaciones por dos vías diferentes. En primer lugar, la cartera se segmenta en clases homogéneas usando ciertas características cualitativas y cuantitativas. Entre las variables explicativas utilizadas para esta segmentación, diferentes estudios citan, como mínimo:

- Tipo de operación: la importancia de esta característica se debe, básicamente, a que existe todo un espectro de operaciones con límites explícitos y diferentes condiciones para las disposiciones, que van desde operaciones con límites incondicionales a operaciones en las que cada disposición requiere algún tipo de aprobación previa.
- *Covenants*: es frecuente que el banco tenga la posibilidad de denegar disposiciones adicionales cuando se dan determinadas circunstancias. Las cláusulas que detallan estas circunstancias se denominan *covenants*<sup>18</sup>. Típicamente, estas cláusulas tienen que ver con condiciones objetivas que son indicadores de deterioro de la calidad crediticia del acreditado, tales como: bajadas de *rating*, caídas de rentabilidad o cambios en ciertas ratios financieras clave por debajo de algunos umbrales explícitos<sup>19</sup>.

En segundo lugar, una vez que se ha identificado una clase incluyendo operaciones que, en principio, son suficientemente homogéneas para diseñar un modelo explicativo común para las EAD, es necesario seleccionar un conjunto apropiado de variables explicativas (*risk drivers* cuantitativos). Entre esas variables cuantitativas, diferentes estudios realizados a partir de bases de datos privadas sugieren que es conveniente considerar, como mínimo:

- El tamaño del límite de crédito  $L(tr)$ .
- El dispuesto y el disponible a la fecha de referencia,  $E(tr)$  y  $L(tr) - E(tr)$ .
- La utilización porcentual del límite a la fecha de referencia  $e(tr)$ . Como se discutió en la sección 4.2.2, este valor de utilización porcentual tiene poder discriminante respecto a los factores  $LEQ_i$ .

<sup>18</sup>. Algunas veces estas cláusulas se denominan cláusulas de cambios adversos significativos (*Material Adverse Changes, MAC clauses*). Véase, por ejemplo, Lev (2004), p. 14. <sup>19</sup>. Para más detalles sobre los *covenants*, véase, por ejemplo, Sufi (2005), p. 5.



- El tiempo que falta para la entrada en *default*  $t_d - t_r$ : hay evidencia empírica de que esta variable tiene capacidad explicativa significativa, al menos, cerca de la fecha de *default*.
- La clase de *rating* en la fecha de referencia  $R(t_r)$ : esta variable es, en general, relevante, pero diferentes estudios han encontrado una correlación positiva significativa entre calidad crediticia y los CF, en algunos casos, y correlación significativa negativa en otros. Parece que el papel del *rating* como variable explicativa relevante está relacionado con el tipo de cartera, las propiedades dinámicas de cada sistema de *rating* y los usos del *rating* para aplicaciones internas.
- Estatus de la operación en la fecha de referencia  $S(t_r)$ : muchos bancos, además de sistemas de *rating* o *scoring*, tienen en funcionamiento sistemas de alerta que están enfocados a la identificación rápida de problemas de liquidez y de otras dificultades a corto plazo de los acreditados. La diferencia básica con los sistemas de *rating* es que estos sistemas de alerta son más dinámicos e identifican problemas antes de que el *rating* lo refleje<sup>20</sup>. Como resultado de estos sistemas, ciertas operaciones se clasifican en clases muy amplias, típicamente: estatus normal y unos pocos grados sujetos a vigilancia especial. Esto significa que, una vez que una operación ha sido identificada como relacionada con un acreditado problemático, el nivel de vigilancia y, en algunos casos, las condiciones prácticas para disposiciones de fondos adicionales cambian<sup>21</sup>. Por lo tanto, el estatus es una variable explicativa crucial para estimar la EAD.
- Indicadores macroeconómicos.

Para las observaciones en el RDS, los valores de las variables anteriores son, en general, conocidos. Para una operación que no ha entrado en *default*, estas variables se calculan utilizando la fecha actual,  $t$ , como fecha de referencia,  $t_r$ . Respecto al tiempo hasta el momento de entrada en *default* existe un problema, ya que, para una operación que no ha entrado en *default*, este tiempo es desconocido. En el contexto de Basilea II se está interesado en estimaciones de la EAD sujetas a la condición de que la operación entra en *default* en un período de un año. Por lo tanto, en este contexto, lo relevante es la influencia de esta variable cuando el rango de valores varía desde uno a doce meses.

## 5 Estimaciones de la EAD

### 5.1 RELACIÓN ENTRE OBSERVACIONES DEL RDS Y LA CARTERA ACTUAL

Esta sección presenta diferentes métodos para asignar una estimación de la EAD a una operación,  $f$ , que no ha entrado en *default* en la fecha  $t$ , perteneciente a la cartera actual, basándose en un subconjunto del RDS que agrupa observaciones (de operaciones que han entrado en *default*) similares a la observación asociada a  $f$  en la fecha en la que desea estimar su EAD,  $EAD(f)$ . Este subconjunto se denota por  $RDS(f)$ .

El proceso que asigna un subconjunto del RDS a cada operación  $f$  en la cartera actual se denomina *mapping* y permite clasificar la cartera agrupando operaciones con el mismo o si-

<sup>20</sup> La relación más común entre estos sistemas de alerta temprana y los *ratings* es que ciertos cambios en el estatus desencadenan los procesos para una nueva evaluación del *rating* del acreditado. **21.** 477. «*Due consideration must be paid by the bank to its specific policies and strategies adopted in respect of account monitoring and payment processing. The bank must also consider its ability and willingness to prevent further drawings in circumstances short of payment default, such as covenant violations or other technical default events. Banks must also have adequate systems and procedures in place to monitor facility amounts, current outstandings against committed lines and changes in outstandings per borrower and per grade. The bank must be able to monitor outstanding balances on a daily basis.*», BCBS (2004).

milar RDS(f). Recíprocamente, algunos bancos usan un procedimiento que segmenta la cartera de exposiciones actuales en clases que agrupan operaciones similares. Este enfoque se puede reducir al anterior, ya que, después de esta clasificación de exposiciones, a cada clase se le asocia un subconjunto del RDS denotado por RDS(C), que se utiliza para la estimación de la EAD(f) para las operaciones incluidas en dicha clase C.

5.2 EQUIVALENCIA  
ENTRE ESTIMACIONES DE EAD(F)  
Y ESTIMACIONES DE CF(F)

Considérese una operación f de la cartera actual que no ha entrado en *default*, dado un estimador de EAD(f), si  $L(f) \neq E(f)$ , la estimación anterior se puede expresar en términos de un factor LEQ(f), de acuerdo con la siguiente ecuación:

$$EAD(f) = E(f) + LEQ(f) \cdot (L(f) - E(f)) \quad [13]$$

donde LEQ(f) viene dado por:

$$LEQ(f) = \frac{EAD(f) - E(f)}{L(f) - E(f)} = \frac{ead(f) - e(f)}{1 - e(f)} \quad [14]$$

Adicionalmente, si se está interesado en estimaciones EAD(f) que verifiquen  $EAD(f) \geq E(f)$ , entonces de la ecuación [13]:

- Si  $L(f) > E(f)$ , se tiene que  $EAD(f) \geq E(f)$  si y solo si  $LEQ(f) \geq 0$ .
- Si  $L(f) < E(f)$ , se tiene que  $EAD(f) \geq E(f)$  si y solo si  $LEQ(f) \leq 0$ .

Por lo tanto, sin hipótesis adicionales, para operaciones que verifican  $L(f) \neq E(f)$ , se ha probado que para estimar la EAD(f) es suficiente centrarse en métodos que estimen factores de conversión adecuados LEQ(f) basados en las observaciones incluidas en el RFD(f) y, después de eso, emplear la ecuación [13] para asignar una estimación individual de la EAD.

Finalmente, el método más simple para estimar una EAD para una clase es sumar las EAD individuales para todas las operaciones incluidas en la clase.

Por ejemplo, para ciertos tipos de operaciones, algunos bancos asignan una EAD a f, mediante la fórmula [15] y utilizando un CCF(f) estimado previamente::

$$EAD(f) = CCF(f) \cdot L(f) \quad [15]$$

Este método se llama, a veces, método del coeficiente de utilización del límite<sup>22</sup> (*Usage at Default Method*); este caso se puede reducir al método general de la ecuación [13] si  $e(f) \neq 1$ , asignando un factor LEQ(f) dado por:

$$LEQ(f) = \frac{CCF(f) \cdot L(f) - E(f)}{L(f) - E(f)} = \frac{CCF(f) - e(f)}{1 - e(f)} \quad [16]$$

Recíprocamente, si tenemos una estimación del LEQ(f), despejando en la ecuación [16] el CCF(f) en función del LEQ(f), obtenemos:

$$CCF(f) = LEQ(f) \cdot (1 - e(f)) + e(f) \quad [17]$$

22. Este método se denomina *Momentum Method* en CEBS *Guidelines* (2006), párrafos 253 y 254.

Por lo tanto, los métodos de estimación de la EAD(f) basados en factores LEQ(f) y los basados en factores CCF(f) son equivalentes, con la excepción de aquellas operaciones en las que  $e(f) = 1$ .

En las próximas secciones se presentan los métodos normalmente empleados en la práctica por los bancos para estimar los factores LEQ(f) desde una perspectiva unificada. Esto se utiliza después para analizar la optimalidad de los diferentes enfoques. Además, se derivan las fórmulas más empleadas en la práctica como casos especiales de los métodos previamente presentados [cuando se asume una determinada forma funcional para el factor LEQ(f)].

5.3 MODELIZACIÓN DE LOS FACTORES DE CONVERSIÓN, CF, A PARTIR DEL RDS

Esta sección presenta varios métodos para estimar factores de conversión a partir de problemas de regresión basados en la siguiente ecuación básica:

$$EAD(f) - E(f) = LEQ(f) \cdot (L(f) - E(f)) \quad [18]$$

Todos estos métodos intentan explicar los incrementos observados en la exposición de las operaciones entre la fecha de referencia y la fecha de *default*, y se pueden agrupar en tres modelos atendiendo a cómo se miden dichos incrementos: como porcentaje del disponible (énfasis en los factores LEQ observados), como porcentaje del límite observado (énfasis en la utilización del límite) o, finalmente, en valor absoluto (énfasis en la exposición).

a. Modelo I. Énfasis en los factores LEQ<sub>i</sub> observados

Dividiendo [18] por  $L(f) - E(f)$ , se obtiene:

$$\frac{ead(f) - e(f)}{1 - e(f)} = \frac{EAD(f) - E(f)}{L(f) - E(f)} = LEQ(f) \quad [19]$$

Por lo tanto, la lógica en este enfoque es determinar una función de las variables explicativas LEQ(RD) que «explique» los factores LEQ<sub>i</sub> observados asociados con RDS(f), dados por  $LEQ_i = (EAD_i - E_i) / (L_i - E_i)$ , en función de LEQ(RD<sub>i</sub>). Esto se puede hacer comenzando por fijar una expresión para el error asociado a la diferencia  $LEQ_i - LEQ(RD_i)$  y resolviendo un problema de minimización del error. Actualmente, en la práctica se usa una función cuadrática y simétrica para el error en casi todos los casos. Como consecuencia de esa elección, el problema de minimización que se ha de resolver viene dado por (problema P.I):

$$\text{Min}_{LEQ} \left\{ \sum_i (LEQ_i - LEQ(RD_i))^2 \right\} = \text{Min}_{LEQ} \left\{ \sum_i \left( \frac{EAD_i - E_i}{L_i - E_i} - LEQ(RD_i) \right)^2 \right\} \quad [20]$$

O, equivalentemente:

$$\overline{LEQ(f)} = \text{Min}_{LEQ} \left\{ \sum_i \frac{1}{(L_i - E_i)^2} \cdot (EAD_i - E_i - LEQ(RD_i) \cdot (L_i - E_i))^2 \right\} \quad [21]$$

b. Modelo II. Énfasis en el incremento de la exposición medido como porcentaje del límite (incremento porcentual en utilización)

Dividiendo la ecuación básica [18] por  $L(f)$ , se obtiene:

$$\frac{EAD(f) - E(f)}{L(f)} = LEQ(f) \cdot \frac{L(f) - E(f)}{L(f)} \quad [22]$$

Por lo tanto, en este enfoque las cantidades observables que se han de explicar vienen dadas por  $(EAD_i - E_i) / L_i$  y los valores explicativos son  $LEQ(RD_i) \cdot (L_i - E_i) / L_i$ . Siguiendo el mismo razonamiento que en el enfoque anterior, el problema de minimización que se debe resolver es (problema P.II):

$$\text{Min}_{\text{LEQ}} \left\{ \sum_i \left( \frac{\text{EAD}_i - E_i}{L_i} - \text{LEQ}(\text{RD}_i) \cdot \frac{(L_i - E_i)}{L_i} \right)^2 \right\} \quad [23]$$

O, de forma equivalente:

$$\overline{\text{LEQ}}(f) = \text{Min}_{\text{LEQ}} \left\{ \sum_i \frac{1}{L_i^2} * (\text{EAD}_i - E_i - \text{LEQ}(\text{RD}_i) \cdot (L_i - E_i))^2 \right\} \quad [24]$$

c. Modelo III. Énfasis en el incremento de la exposición

Directamente de la ecuación básica se tiene:

$$\text{EAD}(f) - E(f) = \text{LEQ}(f) \cdot (L(f) - E(f)) \quad [25]$$

En este caso, las cantidades que se han de explicar son las  $\text{EAD}_i - E_i$ , y la variable explicativa es  $\text{LEQ}(\text{RD}_i) \cdot (L_i - E_i)$ . Como en los otros casos, el problema de minimización asociado viene dado por (problema P.III):

$$\overline{\text{LEQ}}(f) = \text{Min}_{\text{LEQ}} \left\{ \sum_i (\text{EAD}_i - E_i - \text{LEQ}(\text{RD}_i) \cdot (L_i - E_i))^2 \right\} \quad [26]$$

De las ecuaciones [21], [24] y [26], se deduce que todos estos problemas se pueden reducir a uno más general (problema P.IV):

$$\text{Min}_{\text{LEQ}} \left\{ \sum_i \left( \frac{\text{EAD}_i - E_i}{\omega_i} - \text{LEQ}(\text{RD}_i) \cdot \frac{(L_i - E_i)}{\omega_i} \right)^2 \right\} \quad [27]$$

donde  $\omega_i$  representa a  $L_i - E_i$  en el modelo I, a  $L_i$  en el modelo II y a 1 en el modelo III. Si  $F^*$  denota a la función de distribución empírica de  $(\text{EAD} - E)/\omega$  asociada con las observaciones incluidas en el RDS(f), el problema P.IV se puede expresar como:

$$\overline{\text{LEQ}}(\text{RD}) = \text{Min}_{\text{LEQ}} \left\{ E_{F^*} \left\langle \left( \frac{\text{EAD} - E}{\omega} - \text{LEQ}(\text{RD}) \cdot \frac{(L - E)}{\omega} \right)^2 \right\rangle \right\} \quad [28]$$

En el caso más general, suponiendo que  $(L - E)/\omega$  es constante para observaciones en RDS(f), la solución de [28] viene dada por<sup>23</sup>:

$$\overline{\text{LEQ}}(f) = E_{F^*} \left\langle \left( \frac{\text{EAD} - E}{\omega} \right) \middle| \text{RD}(f) \right\rangle \cdot \frac{\omega(f)}{L(f) - E(f)} \quad [29]$$

En consecuencia, en la práctica, el problema se reduce a encontrar métodos para aproximar estas esperanzas condicionadas.

Si se asume una forma paramétrica para el factor LEQ, el problema se convierte en:

$$\begin{aligned} \overline{\text{LEQ}}(f) &= \text{LEQ}(\hat{a}, \hat{b}, \dots), \\ \{\hat{a}, \hat{b}, \dots\} &= \text{Min}_{\{a, b, \dots\}} \left\{ E_{F^*} \left\langle \left( \frac{\text{EAD} - E}{\omega} - \text{LEQ}(a, b, \dots) \cdot \frac{(L - E)}{\omega} \right)^2 \right\rangle \right\} \end{aligned} \quad [30]$$

23. Véase el anejo 2.

Si, además, la forma funcional es lineal en los parámetros, el problema se reduce a un problema de regresión lineal.

En resumen, los métodos tradicionales se pueden clasificar como modelos de regresión que se centran en la minimización del error cuadrático en las predicciones de: los factores  $LEQ_i$ , las EAD en porcentaje del límite o las EAD. Estos métodos producen diferentes estimaciones de la EAD(f) que se basan en estimaciones de los factores  $LEQ(f)$  proporcionales a ciertas esperanzas condicionadas. A primera vista, el método que se fija directamente en los factores  $LEQ$  (modelo I) parece el más natural, el que se fija en los incrementos porcentuales en la utilización del límite (modelo II) parece más estable que el método anterior y, como se muestra en la sección 6, el método basado en los incrementos de la EAD (modelo III) podría presentar ventajas cuando las estimaciones se van a usar en los cálculos de capital regulatorio, debido a la relación entre los requisitos de capital y la EAD.

#### 5.4 LEQ = CONSTANTE

##### 5.4.1 P.I. La media muestral

En la práctica<sup>24</sup>, los bancos usan con frecuencia, como estimador de  $LEQ(f)$  en  $t$ , la media muestral de los factores  $LEQ$  observados,  $LEQ_i$ , restringida a aquellas observaciones que son similares a  $\{f,t,RD\}$ . Suponiendo que el factor de conversión es constante para observaciones similares a  $\{f,t\}$ ,  $LEQ(f) = LEQ$ , y resolviendo el problema P.I se obtiene:

$$\overline{LEQ} = \underset{LEQ \in R}{\text{Min}} \left\{ \sum_i \left( \frac{EAD_i - E_i}{L_i - E_i} - LEQ \right)^2 \right\} = \frac{1}{n} \sum \frac{EAD_i - E_i}{L_i - E_i} = \frac{1}{n} \sum LEQ_i \quad [31]$$

En otros casos, los bancos utilizan una media muestral ponderada que trata de tener en cuenta una posible relación entre el tamaño de las exposiciones (o de los límites) y los factores  $LEQ$ . Si en el problema P.I se introduce una ponderación  $w_i$ , y se supone que el factor  $LEQ$  es constante para observaciones similares a  $\{f,t\}$ , entonces:

$$\overline{LEQ} = \underset{LEQ \in R}{\text{Min}} \left\{ \sum_i w_i \left( \frac{EAD_i - E_i}{L_i - E_i} - LEQ \right)^2 \right\} = \frac{\sum w_i \cdot LEQ_i}{\sum w_i} \quad [32]$$

Cuando la justificación para incluir las ponderaciones es tener en cuenta una variable explicativa para el  $LEQ$ , este tratamiento es incoherente. La razón es que la media ponderada es una solución óptima solamente cuando suponemos que  $LEQ = \text{constante}$ , y esto significa que no se están considerando variables explicativas adicionales.

#### 5.4.2 P.II. La regresión

##### sin constante

Otro método, ampliamente utilizado por los bancos es usar el estimador de la pendiente de una regresión lineal basada en el modelo II y asumir que el factor  $LEQ$  es constante. Bajo esas condiciones, la expresión para el estimador de regresión viene dada por:

$$\begin{aligned} \overline{LEQ} &= \underset{LEQ \in R}{\text{Min}} \left\{ \sum_i \left( \frac{EAD_i - E_i}{L_i} - LEQ \cdot \left( \frac{L_i - E_i}{L_i} \right) \right)^2 \right\} \\ &= \frac{\sum \frac{(EAD_i - E_i)(L_i - E_i)}{L_i^2}}{\sum \left( \frac{L_i - E_i}{L_i} \right)^2} = \frac{\sum (ead_i - e_i)(1 - e_i)}{\sum (1 - e_i)^2} \end{aligned} \quad [33]$$

24. Al menos, este es el caso de los modelos aplicados por algunos bancos españoles en la actualidad (2006).

Si en el problema P.III se supone que  $LEQ = \text{constante}$ , este se puede expresar como:

$$\overline{LEQ}(f) = \text{Min}_{LEQ \in R} \left\{ \sum_i (L_i - E_i)^2 \left( \frac{EAD_i - E_i}{(L_i - E_i)} - LEQ \right)^2 \right\} \quad [34]$$

Y el óptimo viene dado por:

$$\overline{LEQ} = \frac{\sum w_i \cdot LEQ_i}{\sum w_i}, \text{ con } w_i = (L_i - E_i)^2 \quad [35]$$

Por lo tanto, bajo este planteamiento, aparece una media ponderada muestral de forma natural. Sin embargo, merece la pena notar que los pesos  $(L_i - E_i)^2$  son diferentes de los actualmente propuestos por algunos bancos (basados en  $L_i$  o en  $E_i$ ).

Este método es utilizado algunas veces por bancos que tratan de evitar el uso explícito de factores  $LEQ$  observados (o estimados) negativos o para tipos de operaciones en los que la utilización actual del límite no tiene capacidad predictiva para la EAD. La EAD de una operación que no ha entrado en *default*  $EAD(f)$  se estima utilizando la ecuación [15] directamente y una estimación sencilla del CCF (por ejemplo, la media muestral de los factores CCF observados asociados a un conjunto de operaciones,  $C$ , que han entrado en *default*).

$$\overline{EAD}(f) = \overline{CCF}(C) \cdot L(f) \quad [36]$$

De la ecuación [16], y suponiendo que  $CCF = \text{constante}$ , se obtiene una forma funcional específica para el factor  $LEQ(e(f))$  dada por:

$$\overline{LEQ}(f) = \frac{\overline{CCF} \cdot L(f) - E(f)}{L(f) - E(f)} = \frac{\overline{CCF} - e(f)}{1 - e(f)} \quad [37]$$

En general, dos operaciones con el mismo CCF estimado y con diferentes valores de su utilización porcentual actual,  $e(t)$ , tendrán asociados diferentes factores  $LEQ$  según la ecuación anterior.

El principal problema con el método de la ecuación [36] es que la experiencia muestra que, en general, el dispuesto y el disponible tienen fuerte capacidad explicativa para la EAD. Por esta razón, este método (con  $CCF = \text{constante}$ ) no parece satisfacer el requisito de utilizar toda la información relevante<sup>25</sup> (ya que no tiene en cuenta el dispuesto y el disponible como variables explicativas de la EAD) para muchos de los tipos de operaciones que aparecen en la práctica.

## 6 Cómo valorar la optimalidad de las estimaciones

Para valorar la optimalidad de las diferentes estimaciones de los CF que se pueden asociar a un conjunto de datos de referencia y a una cartera, es necesario precisar algunos elementos en el problema básico. El primer elemento que se ha de clarificar es el tipo de estimaciones requerido atendiendo al papel de las variables macroeconómicas explicativas en el método de estimación. El segundo elemento que se debe precisar es cómo se miden los errores asociados con las estimaciones y motivar dicha elección. Esto se puede hacer introduciendo una función de pérdida que especifica cómo se penalizan las diferencias existentes entre los valores estimados de la EAD y los valores reales.

25. 476. «The criteria by which estimates of EAD are derived must be plausible and intuitive, and represent what the bank believes to be the material drivers of EAD. The choices must be supported by credible internal analysis by the bank [...] A bank must use all relevant and material information in its derivation of EAD estimates...», BCBS (2004).

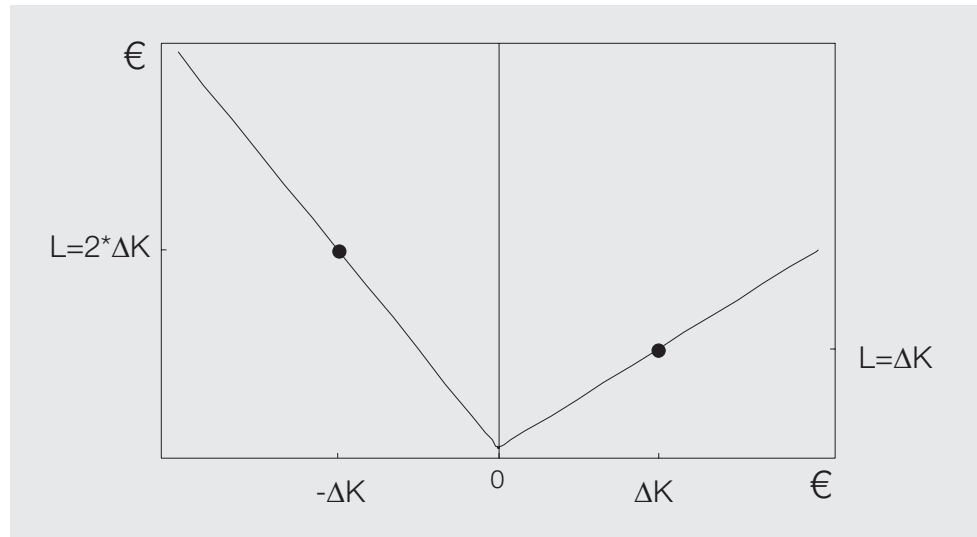
Atendiendo al uso de las variables macroeconómicas, se puede distinguir entre los siguientes tipos de estimaciones:

- *Estimaciones PIT (Point in time estimates)*: estas estimaciones están condicionadas a ciertos valores de las variables macroeconómicas explicativas; por ejemplo, valores próximos a los valores actuales. Lo anterior permite que las estimaciones estén afectadas por las condiciones económicas actuales y que varíen siguiendo el ciclo económico. Esto, en teoría, es una buena propiedad para las estimaciones internas que los bancos necesitan para la fijación de precios de referencia y otras finalidades internas. El principal problema con las estimaciones PIT es que están basadas en menos datos que las estimaciones promedio a largo plazo (estimaciones LR, que se definen más abajo) y, por lo tanto, en la práctica son menos estables y más difíciles de estimar que las LR.
- *Estimaciones promedio a largo plazo LR (Long-run estimates)*: se trata de estimaciones incondicionales respecto a las variables explicativas macroeconómicas, es decir, se ignoran los valores de las variables explicativas macroeconómicas en las estimaciones. La principal ventaja es que son más robustas y más estables que las estimaciones PIT. Estas estimaciones LR son las requeridas para los enfoques IRB<sup>26</sup>, con la excepción de aquellas carteras en las cuales exista evidencia de una dependencia adversa entre las tasas de *default* y los factores LEQ. Actualmente, este tipo de estimaciones LR también son las que los bancos usan actualmente para sus fines internos.
- *Estimaciones DT (Downturn estimates)*: se trata de estimaciones PIT basadas en escenarios macroeconómicos (*downturn conditions*) en los que las tasas de *default* para la cartera considerada se supone que serían especialmente altas. Cuando existe evidencia de que hay una dependencia adversa entre las tasas de *default* y los factores LEQ, este podría ser el tipo de estimaciones que, en teoría, deberían usarse para los enfoques IRB<sup>27</sup>. En la práctica, el uso de estimaciones DT es difícil, ya que, además de las dificultades asociadas con las estimaciones PIT, es necesario identificar los escenarios macro (*downturn conditions*) y disponer de un número suficiente de observaciones bajo esas circunstancias.

En lo que sigue, se supone que se está interesado en estimaciones promedios a largo plazo.

El objetivo de esta sección es determinar un tipo de función de pérdida que verifique los requisitos básicos del problema de estimación de la EAD cuando es necesario obtener estimaciones de la EAD adecuadas para los enfoques IRB. Por lo tanto, tiene sentido especificar la pérdida asociada con la diferencia entre el valor estimado y el valor real en términos del error en el capital mínimo regulatorio (calculado como la diferencia entre los requisitos de capital asociados a las dos cifras). Utilizando la fórmula regulatoria a nivel de

26. 475. «Advanced approach banks must assign an estimate of EAD for each facility. It must be an estimate of the long-run default – weighted average EAD for similar facilities and borrowers over a sufficiently long period of time, [...] If a positive correlation can reasonably be expected between the default frequency and the magnitude of EAD, the EAD estimate must incorporate a larger margin of conservatism. Moreover, for exposures for which EAD estimates are volatile over the economic cycle, the bank must use EAD estimates that are appropriate for an economic downturn, if these are more conservative than the long-run average.», BCBS (2004). 27. Esto se puede interpretar a la luz de las clarificaciones a los requisitos sobre las estimaciones de la LGD contenidos en el párrafo 468 de Marco Revisado. Véase BCBS (2005).



operación, la pérdida asociada con la diferencia entre el requisito de capital bajo el valor estimado de la exposición  $K(EAD(f))$  y bajo el valor real de la exposición  $K(EAD)$  se puede expresar como sigue<sup>28</sup>:

$$\begin{aligned}
 L(\Delta K(f)) &= L(K(EAD) - K(EAD(f))) = \\
 &= L(\phi(PD) \cdot LGD \cdot (EAD - EAD(f))) = L(\phi(PD) \cdot LGD \cdot \Delta(EAD(f)))
 \end{aligned}
 \tag{38}$$

Además, al menos desde un punto de vista regulatorio, subestimar el nivel de capital requerido es peor que sobreestimar dicho importe. Por esta razón, es conveniente utilizar una función de pérdida asimétrica que penaliza más una subestimación del requisito de capital que una sobreestimación del mismo importe. La familia más simple de tales funciones viene dada por [39], donde  $b > a$ :

$$L(\Delta K) = \begin{cases} a \cdot \Delta K & \text{sii } \Delta K \geq 0 \\ b \cdot \Delta K & \text{sii } \Delta K < 0 \end{cases}
 \tag{39}$$

Estas funciones de pérdida permiten cuantificar el grado de prudencia. Cuanto mayor sea  $b/a$ , mayor es la penalización asociada a una subestimación del capital con respecto a una sobreestimación. Por ejemplo, si  $a = 1$  y  $b = 2$ , la pérdida asociada con una subestimación del capital requerido ( $\Delta K < 0$ ) es dos veces la pérdida asociada con una sobreestimación de la misma magnitud<sup>29</sup>. El gráfico de la función de pérdida se presenta en la figura 6.

Utilizando este tipo específico de función de pérdida [39], y suponiendo que  $LGD \geq 0$ , se obtiene una expresión más simple para el error en el capital  $K$  en términos del error en la EAD:

$$L(\Delta K(f)) = \phi(PD) \cdot LGD \cdot L(\Delta(EAD(f)))
 \tag{40}$$

<sup>28</sup>. En lo que sigue, se supone que la operación  $f$  tiene una  $PD$  y una  $LGD$  adecuadas, estimadas previamente. <sup>29</sup>. Que yo tenga noticia, la primera aplicación de una función de pérdida de ese tipo en el contexto del riesgo de crédito fue propuesta en Moral (1996). En ese artículo la función de pérdida se usó para determinar el nivel óptimo de provisiones como un cuantil de la distribución de pérdida de la cartera.



Se demuestra que la pérdida asociada con un error en el requisito de capital es proporcional a la pérdida asociada con el error en términos de la exposición y que las unidades en que está expresada la pérdida son las mismas que las de la exposición (€).

### 6.3 LA FUNCIÓN OBJETIVO

Una vez que se ha escogido la función de pérdida, hay que encontrar la función de objetivo más natural para el problema de estimación.

6.3.1 Minimización de la esperanza matemática del error en el capital mínimo a nivel de operación

Si se utiliza como función objetivo el error esperado del requisito de capital mínimo a nivel de operación, utilizando la ecuación [40] se obtiene el siguiente resultado:

$$\text{Min}_{\text{LEQ}} \{E[L(\Delta K(f))]\} = \phi(\text{PD}) \cdot \text{LGD} \cdot \text{Min}_{\text{LEQ}} \{E[L(\Delta(\text{EAD}(f)))]\} \quad [41]$$

Esto significa que aparece el problema P.III de la sección 5.3 con una función de pérdida diferente:

$$\text{Min}_{\text{LEQ}} \left\{ E_{F^*} \left( L(\text{EAD} - E - \text{LEQ}(\text{RD})) \cdot (L - E) \right) \right\} \quad [42]$$

O, en términos de la muestra:

$$\text{Min}_{\text{LEQ}} \left\{ \sum_i L(\text{EAD}_i - E_i - \text{LEQ}(\text{RD}_i)) \cdot (L_i - E_i) \right\} \quad [43]$$

Y una solución<sup>30</sup> viene dada por:

$$\overline{\text{LEQ}}(f) = \frac{Q}{F^*} \left( \text{EAD} - E, \frac{b}{a+b} | \text{RD}(f) \right) \cdot \frac{1}{L(f) - E(f)} \quad [44]$$

donde  $Q(x, b/(a+b))$  representa un cuantil de la distribución  $F(x)$ , tal que<sup>31</sup>  $F(Q) = b/(a+b)$ . Si  $a = b$ , la función de pérdida dada por [39] es simétrica y el cuantil resultante es la mediana. Para valores de  $b/a > 1$  se obtienen cuantiles situados a la derecha de la mediana y, por lo tanto, estimaciones más conservadoras del LEQ. Es interesante notar que [44] penaliza la incertidumbre cuando  $b > a$ <sup>32</sup>.

Una consecuencia importante de utilizar la anterior función de pérdida es que los problemas M.I y M.II descritos en las ecuaciones [45] y [46] son equivalentes<sup>33</sup>.

Problema M.I:

$$\text{Min}_{\text{LEQ}} \left\{ \sum_i L(\text{EAD}_i - E_i - \text{LEQ}(\text{RD}_i)) \cdot (L_i - E_i) \right\} \quad [45]$$

Sujeto a :  $0 \leq \text{LEQ}(f) \leq 1$

Problema M.II:

$$\text{Min}_{\text{LEQ}} \left\{ \sum_i L(\text{Min}[\text{Max}[\text{EAD}_i, E_i], L_i] - E_i - \text{LEQ}(\text{RD}_i)) \cdot (L_i - E_i) \right\} \quad [46]$$

Sujeto a :  $0 \leq \text{LEQ}(f) \leq 1$

<sup>30</sup>. Véase anejo 2. <sup>31</sup>. En la práctica, es necesario ser más preciso al definir un  $q$  - cuantil, ya que la distribución  $F(x)$  es discreta. Una definición común es: un  $q$  - cuantil de  $F(x)$  es un número real,  $Q(q, x)$ , que satisface  $P[X \leq Q(q, x)] \geq q$  y  $P[X \geq Q(q, x)] \geq 1 - q$ . En general, con esta definición hay más de un  $q$  - cuantil. <sup>32</sup>. 475. «Advanced approach banks must assign an estimate of EAD for each facility. It must be an estimate [...] with a margin of conservatism appropriate to the likely range of errors in the estimate.», BCBS (2004). <sup>33</sup>. La prueba se sigue de la proposición contenida en el anejo 1.

Esto significa que un estimador que satisfaga la restricción  $0 \leq \text{LEQ}(f) \leq 1$ , si es óptimo utilizando los datos originales, también lo será utilizando los datos censurados para mostrar factores LEQ observados entre cero y uno.

6.3.2 Minimización del error en el requisito de capital a nivel de operación para clases regulatorias

Algunas veces, a pesar de la existencia de estimaciones internas de los factores LEQ a nivel de operación, puede ser necesario asociar una estimación común del LEQ para todas las operaciones incluidas en una clase que agrupa operaciones con diferentes valores para las variables explicativas internas. Esto puede ocurrir, por ejemplo, debido a la existencia de dificultades para demostrar con los datos disponibles que la discriminación al nivel de granularidad interno está justificada. En este caso, para uso regulatorio es necesario adoptar una estructura menos granular para las variables explicativas que la utilizada para los fines internos. Por lo tanto, el problema de encontrar un estimador óptimo para uso regulatorio se puede resolver utilizando la estructura de las variables explicativas válida para fines regulatorios. En otras palabras, el procedimiento propuesto consiste en calcular unas nuevas estimaciones, utilizando el mismo método, con una estructura menos granular para las variables explicativas. En general, el nuevo estimador no será una media simple o ponderada de los estimadores anteriores.

## 7 Ejemplo 1

Este ejemplo<sup>34</sup> ilustra las ventajas e inconvenientes básicos de utilizar los métodos explicados en las secciones anteriores para estimar los factores LEQ y las EAD. Se centra en estimaciones promedios a largo plazo de la EAD de una operación  $f$ , que se encuentre en un estatus normal, utilizando como variables explicativas básicas el límite  $L(t)$  y la exposición  $E(t)$  actuales.

7.1 RDS

7.1.1 Características

Las principales características del RDS utilizadas en este ejemplo se describen seguidamente:

- Fuente del RDS: las observaciones se obtienen de un conjunto de operaciones que han entrado en *default*, durante un determinado período de observación, pertenecientes a una cartera de medianas empresas (SME).
- Período de observación: cinco años.
- Tipos de producto: cuentas de crédito con un límite de crédito comprometido, conocido por el cliente, dado por  $L(t)$ .
- Exclusiones: No se incluyen todos los *defaults* internos que ocurrieron durante el período de observación, ya que se han aplicado previamente varios filtros. Como mínimo, se han eliminado las siguientes operaciones del conjunto de datos de referencia:
  - operaciones que han entrado en *default* con  $L(td - 12) < E(td - 12)$ , y
  - aquellas con menos de doce observaciones mensuales antes de la fecha de *default*.
- Número de observaciones,  $O_i$ :  $\#RDS = 417 * 12 = 5.004$  observaciones, que están asociadas con 417 operaciones que han entrado en *default* y las fechas de referencia 1, 2, ..., 12 meses anteriores a la entrada en *default*.

<sup>34</sup>. Aunque este ejemplo podría considerarse representativo para las cuentas de crédito de ciertas carteras de medianas empresas, no se corresponde con una cartera real extraída de un banco.

- Estructura del conjunto de datos de referencia, RDS: la estructura propuesta en [8], pero, por simplicidad, solo considera un conjunto básico de variables explicativas:

$$O_i = \{i = (f, tr), RD_i = \{L(tr), E(tr), S(tr), td\}, EAD_i\} \quad [47]$$

- Estatus de una operación a la fecha de referencia: no hay información sobre el estatus de las operaciones. El banco tiene implementado un sistema de alertas que clasifica las exposiciones en cuatro clases muy amplias: N = controles y nivel de vigilancia normales; V = bajo estrecha vigilancia para las disposiciones; I = excedidos, es decir, la exposición actual es mayor que el límite, y esto implica que existen controles férreos que impiden disposiciones adicionales sin autorización expresa; y, por último, D = operación que ha entrado en *default*, donde no son posibles nuevas disposiciones, si bien, en algunos casos, la exposición se incrementa debido a cargos por liquidaciones de intereses y comisiones. En este ejemplo, a los efectos de tener en cuenta la variable estatus como variable explicativa, las observaciones con estatus  $S(tr) = N$  se han aproximado mediante el siguiente procedimiento:

- En primer lugar, todas las observaciones con  $E(tr) > L(tr)$  se marcan con un estatus no normal.
- En segundo lugar, después de analizar las distribuciones empíricas de los factores LEQ observados (y otra información), se decidió considerar todas las observaciones con  $td - tr$  menor que cinco meses como si estuvieran en un estatus no normal y eliminar todas las observaciones con  $td - tr =$  siete meses (véase la sección siguiente).

En la práctica, el uso de los valores reales de la variable estatus es necesario, ya que la identificación temprana de acreditados problemáticos y los cambios subsiguientes en la disponibilidad de acceso al límite nominal tienen importantes consecuencias en las EAD observadas. Por esta razón, las observaciones hasta cinco meses antes de la fecha de *default* para las que  $E(tr) \leq L(t)$  se consideraron con un estatus normal. En este caso, el número de observaciones con  $S(tr) = N$  es:  $\#RDS(N) = 2.919$ .

### 7.1.2 Distribuciones empíricas de ciertos estadísticos

El histograma de la figura 7 resume la variabilidad empírica de los factores LEQ asociados con las 2.442 observaciones para las que es posible calcular dicho estadístico<sup>35</sup>.

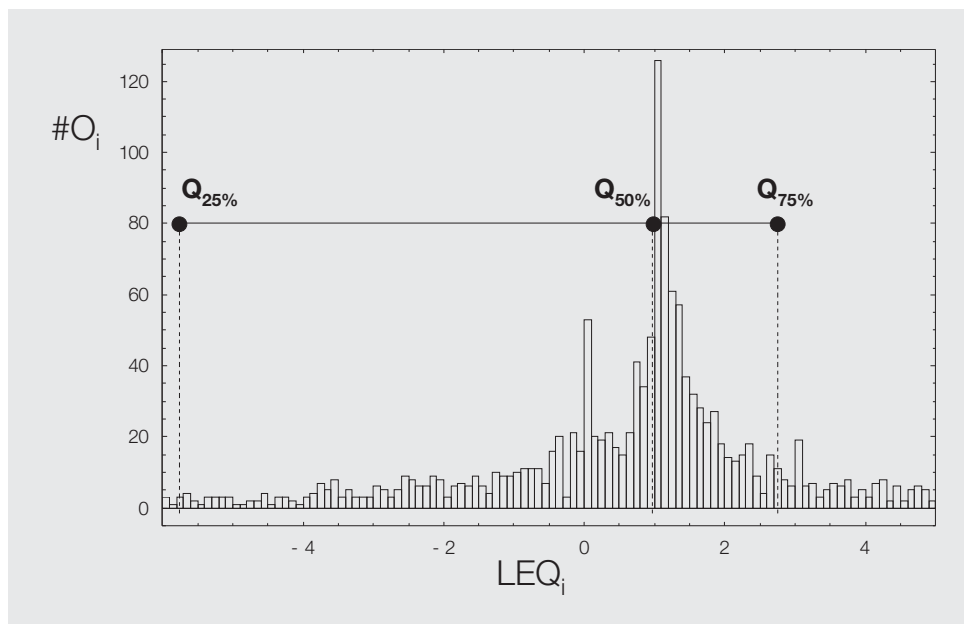
#### a. Distribución de los factores LEQ observados

La citada figura muestra que la distribución es muy asimétrica, con un número muy alto de observaciones fuera del intervalo  $[0, 1]$ , que es el rango natural para los factores LEQ observados. La media muestral es aproximadamente  $-525$ , debido a la existencia de muchas observaciones con valores negativos muy grandes en valor absoluto, y resalta uno de los principales problemas que aparecen cuando se utiliza la media muestral como estimador. La mediana es  $0,97$  y ese valor, comparado con el de la media, pone de manifiesto las ventajas de utilizar estadísticos menos dependientes de los valores extremos de la distribución para la estimación.

#### b. Distribución conjunta de los factores LEQ observados y de la utilización porcentual a la fecha de referencia

Para reducir la variabilidad de los factores LEQ observados es necesario considerar alguna variable que tenga capacidad explicativa, al menos, para el rango de valores de los factores

<sup>35</sup>. Se han eliminado las observaciones asociadas con el horizonte  $td - tr = 7$ , por las razones que se explican más adelante.

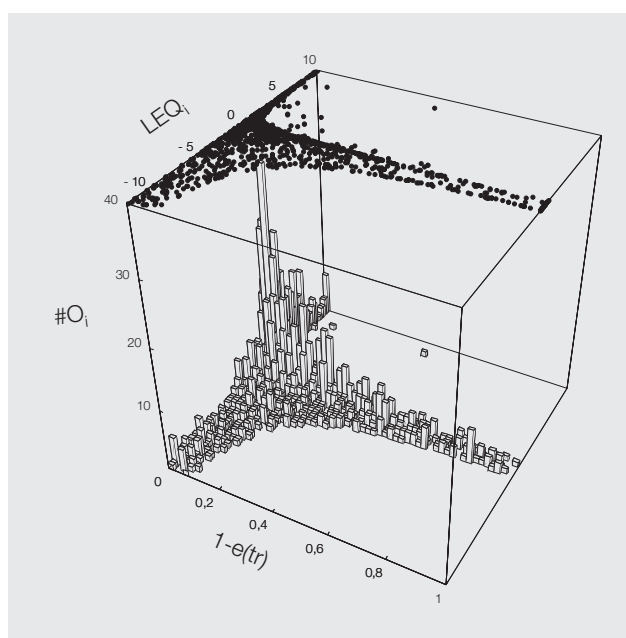


LEQ observados. Por ejemplo, la distribución empírica conjunta presentada en la figura 8 muestra que la variable utilización porcentual a la fecha de referencia es importante para limitar la variabilidad de los factores LEQ observados. Los puntos negros, en la parte superior de la figura, representan las observaciones en el espacio  $\{1 - e(tr), LEQ_i\}$ .

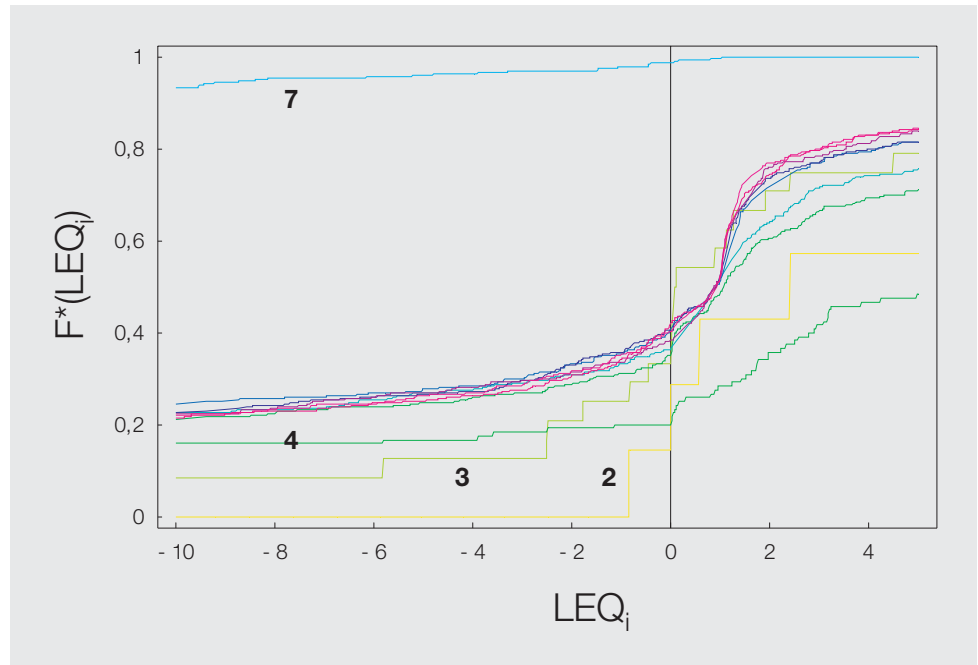
c. Influencia de  $td - tr$  en los estadísticos básicos

La figura 9 presenta las distribuciones empíricas de los factores LEQ observados asociadas con un distancia fija en meses entre la fechas de *default* y de referencia para valores los valores  $td - tr = 1, 2, \dots, 12$ .

DISTRIBUCIÓN CONJUNTA DE LOS LEQ<sub>i</sub> Y DE LA UTILIZACIÓN PORCENTUAL EN LA FECHA DE REFERENCIA  $tr$



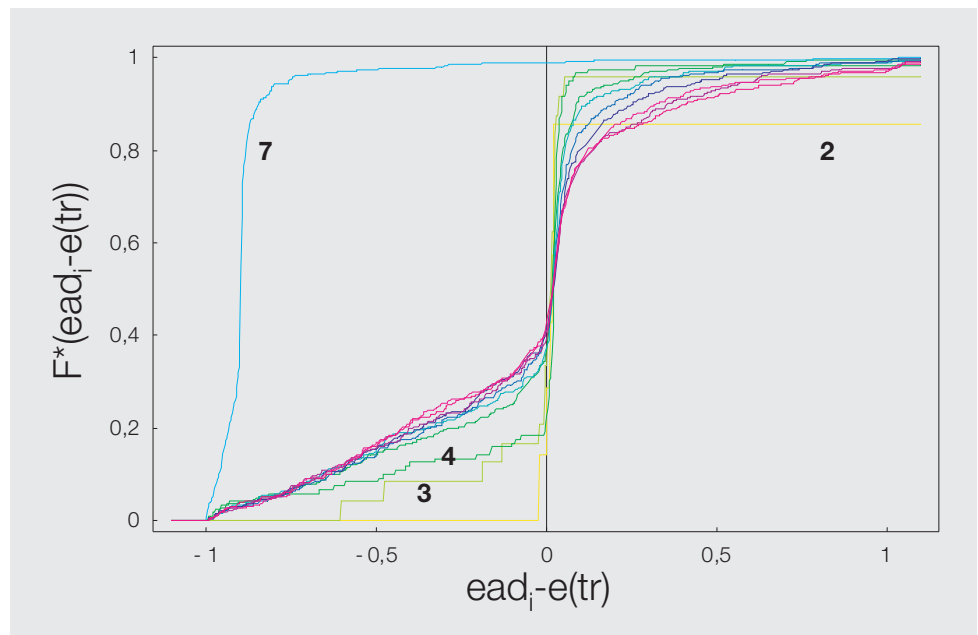
DE  $td - tr$

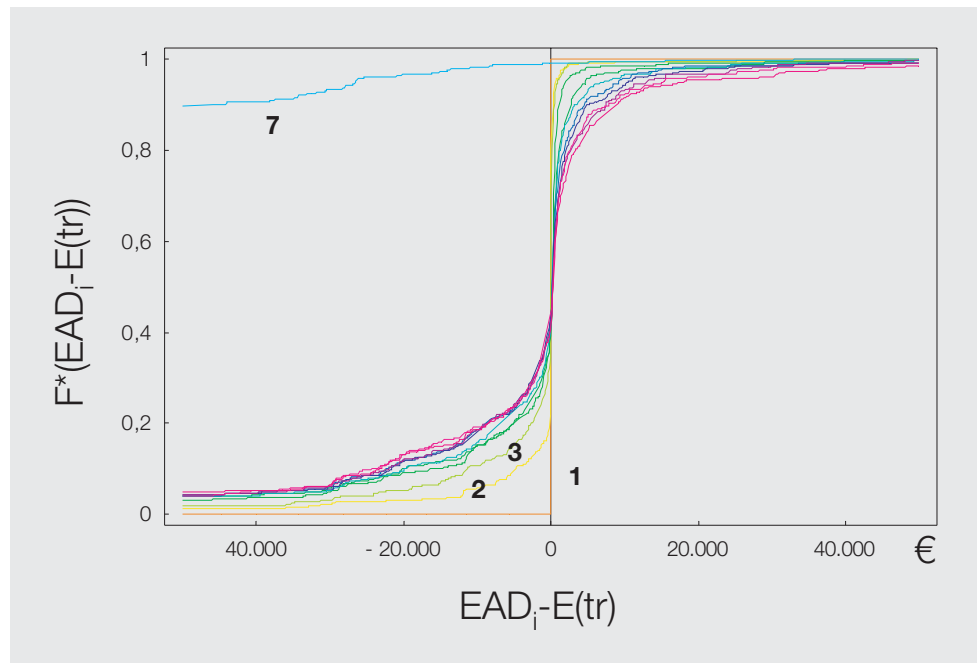


Las distribuciones asociadas con valores de  $td - tr = 1, 2, 3, 4$  son muy diferentes de las otras. La distribución condicional a  $td - tr = 7$  es totalmente anómala, y la razón para este comportamiento es un error existente en el proceso que generó estos datos.

La figura 10 representa las distribuciones empíricas del incremento en la utilización porcentual entre la fecha de referencia y la fecha de *default*,  $ead_i - e(tr)$ , asociadas con una distancia fija en meses entre las fechas de referencia y de *default* para valores de  $td - tr = 1, \dots, 12$ . Nuevamente, las diferencias entre las distribuciones condicionadas a observaciones con fechas

EN LA UTILIZACIÓN DESDE  $tr$





de referencia cercanas al *default* y lejanas al *default* son obvias, y la existencia de valores anómalos para el valor  $t_d - t_r = 7$  es evidente.

Finalmente, la figura 11 muestra las distribuciones empíricas del incremento de la exposición,  $EAD_i - E(tr)$ , entre las fechas de referencia y de *default*.

7.2 PROCEDIMIENTOS  
DE ESTIMACIÓN  
7.2.1 Modelo II

Algunos bancos plantean el modelo II suponiendo un factor LEQ constante y, normalmente, utilizan un enfoque de horizonte fijo a doce meses,  $T = 12$ . Esto significa que ajustan un modelo de regresión lineal sin término independiente, dado por:

a. Datos originales y enfoque de horizonte fijo

$$\frac{EAD_i}{L(td - 12)} - \frac{E(td - 12)}{L(td - 12)} = k \cdot \left( 1 - \frac{E(td - 12)}{L(td - 12)} \right) \quad [48]$$

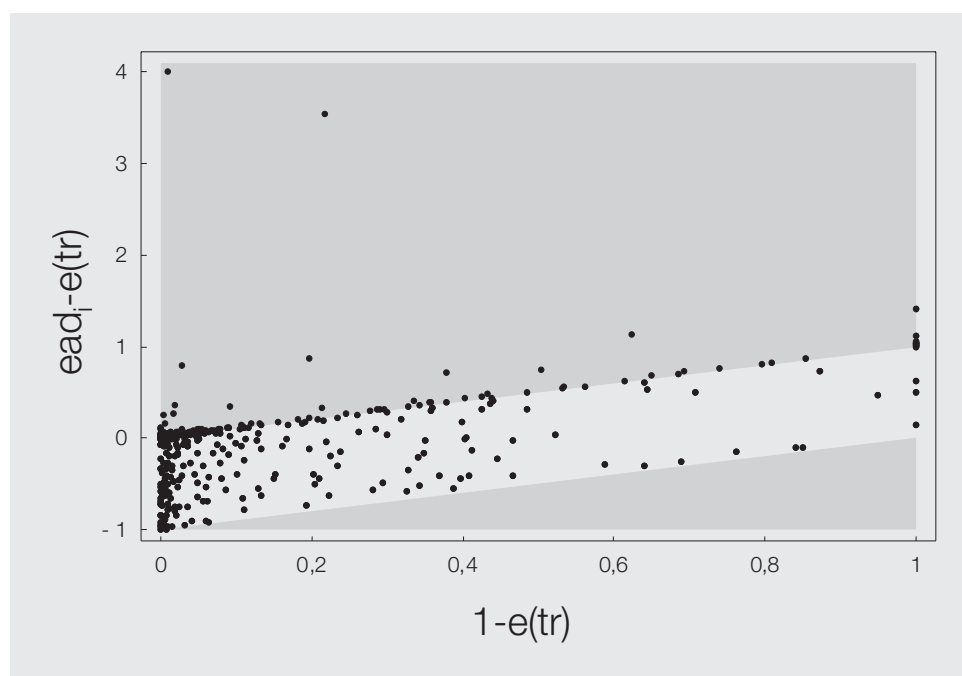
Por lo tanto, en estos casos, el planteamiento del banco se centra en la minimización del error cuadrático del incremento de la exposición expresada como porcentaje del límite. Los resultados con este método se resumen a continuación.

Utilizando los *datos originales*, la estimación del factor LEQ es  $LEQ = 0,637$  y el  $R^2$  ajustado es 0,13. Por lo tanto, la estimación final para la LGD de una operación,  $f$ , que se encuentra en un estatus normal viene dada por la fórmula:

$$EAD(f) = E(t) + 0,637 \cdot (L(t) - E(t)) \quad [49]$$

La figura 12 presenta, para cada observación en  $RDS(td - 12)$ , los valores de las parejas  $\{1 - e(td - 12), ead_i - e(td - 12)\}$ . La zona sombreada en la parte superior de las figuras 12, 13 y 14 se corresponde con puntos en los que  $LEQ_i > 1$ .

Del análisis de la distribución anterior y de los resultados de la regresión es claro que, como mínimo:

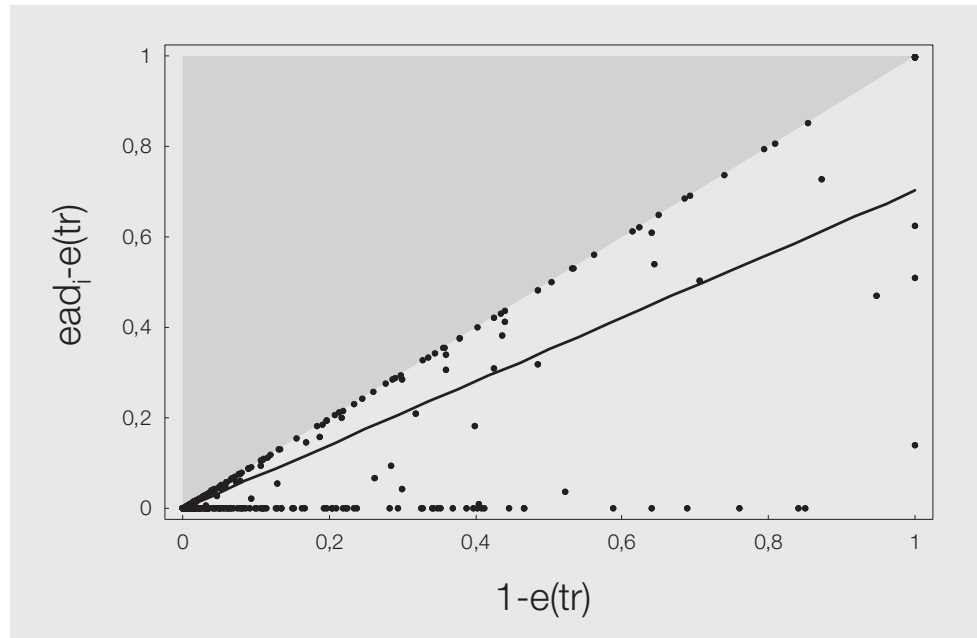


- 1 Es necesario llevar a cabo un procedimiento explícito de limpieza de los datos originales antes de la fase de estimación. Por ejemplo, es necesario analizar las observaciones asociadas a los puntos situados encima de la línea  $y = x$ , y, después de eso, decidir qué observaciones deben ser eliminadas del RDS.
- 2 El grado de ajuste es muy bajo. Muchos de los puntos (aquellos con disponibilidad,  $1 - e(tr)$ , próxima a cero) tienen poca influencia en el resultado de la regresión, debido a la restricción de que no hay término independiente.
- 3 Para valorar la fiabilidad de la estimación del LEQ es necesario identificar los *outliers* y las observaciones influyentes y efectuar análisis de estabilidad. En este caso, dada la forma funcional del modelo,  $y = k \cdot x$ , y el bajo número de puntos asociados con valores grandes de  $1 - e(tr)$ , estas observaciones son puntos influyentes<sup>36</sup>. Es fácil comprender que los cambios en esos puntos afectan al resultado de la regresión y, por lo tanto, a la estimación del LEQ.
- 4 Para obtener resultados más estables es necesario contar con más observaciones (por ejemplo, utilizando el método de horizonte variable).

b. Datos censurados y enfoque de horizonte fijo

Algunas veces, los bancos censuran los datos para conseguir que los factores LEQ observados satisfagan la restricción  $0 \leq LEQ_i \leq 1$ . Utilizando datos censurados, la estimación del factor LEQ es 0,7 y el  $R^2$  se incrementa hasta llegar a 0,75. En este caso, todos los puntos se encuentran en la región triangular blanca de la figura 13, y es claro que la existencia de puntos muy influyentes (los asociados con valores grandes de  $1 - e(tr)$ ) da lugar a inestabilidad. La figura 13 presenta los puntos asociados a las observaciones censuradas y la recta de regresión.

<sup>36</sup>. Los denominados puntos influyentes tienen un impacto significativo en la pendiente de la recta de regresión, que, en el modelo II, es precisamente la estimación del factor LEQ.



El estimador de la EAD en este caso es:

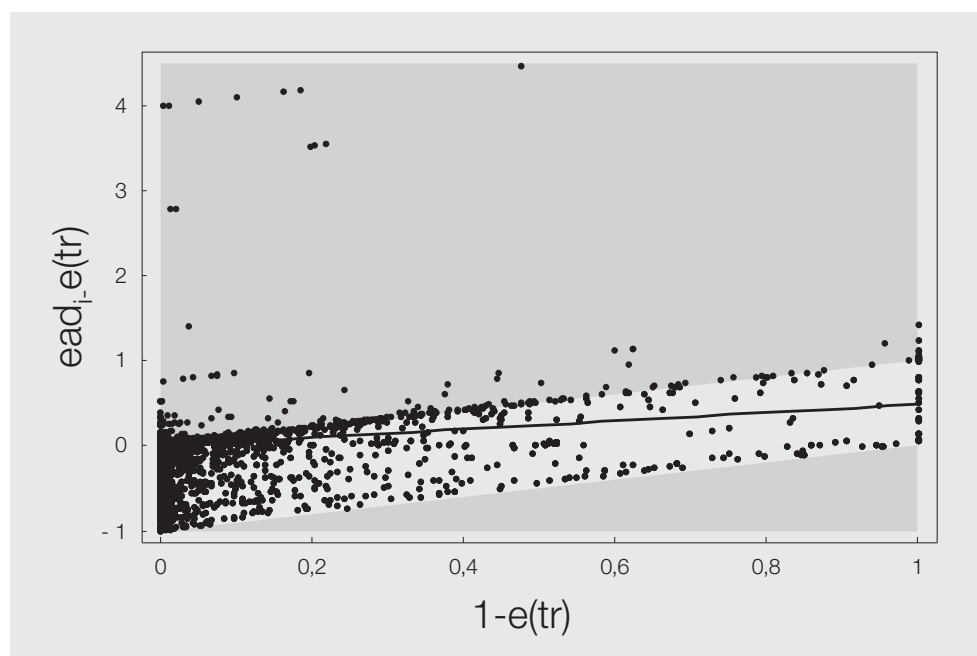
$$EAD(f) = E(t) + 0,7 \cdot (L(t) - E(t)) \quad [50]$$

c. Datos originales y enfoque de horizonte variable

Utilizando un enfoque basado en el uso de un *horizonte variable* con observaciones en las fechas  $tr = td - \{12, 11, 10, 9, 8\}$ , el factor LEQ estimado es  $LEQ = 0,49$  y el  $R^2$  es  $0,06$ . La figura 14 representa, para cada observación en el RDS, las parejas  $\{1 - e(tr), ead_i - e(tr)\}$  y la recta de regresión asociada con este conjunto de datos ampliado aplicando el modelo II.

REGRESIÓN LINEAL EN EL MODELO II CON UN HORIZONTE VARIABLE

FIGURA 14





En el modelo II, el uso de un horizonte variable no incrementa el grado de ajuste (que es muy bajo, debido a la forma funcional incluida en el modelo), pero sí incrementa la estabilidad de los resultados.

El estimador de la EAD, en este caso, es:

$$EAD(f) = E(t) + 0,49 \cdot (L(t) - E(t)) \quad [51]$$

7.2.2 La media muestral y la media muestral condicionada

Si se utiliza el modelo I y se asume que para operaciones similares a  $f$  el factor LEQ es constante, se obtiene una estimación para  $EAD(f)$  calculando la media muestral de los factores LEQ observados para observaciones incluidas en  $RDS(f)$  como estimador de  $LEQ(f)$  y, después, aplicando la ecuación [13]. Respecto a cómo definir el  $RDS(f)$ , en este ejemplo se analizan dos posibilidades:

- $RDS(f) = RDS$  o, de forma equivalente, utilizar la media muestral global como estimador.
- $RDS(f) = \{O_i \text{ tales que } e_i \text{ sea «similar» a } e(f)\}$  o, equivalentemente, utilizar como estimador una función basada en medias locales diferentes, dependiendo del valor de  $e(f)$ .

a. Caso  $RDS(f) = RDS$ , uso de la media muestral global

Si se calcula la media muestral de todos los factores LEQ observados asociados con las observaciones en el RDS, se obtiene una cifra absurda:

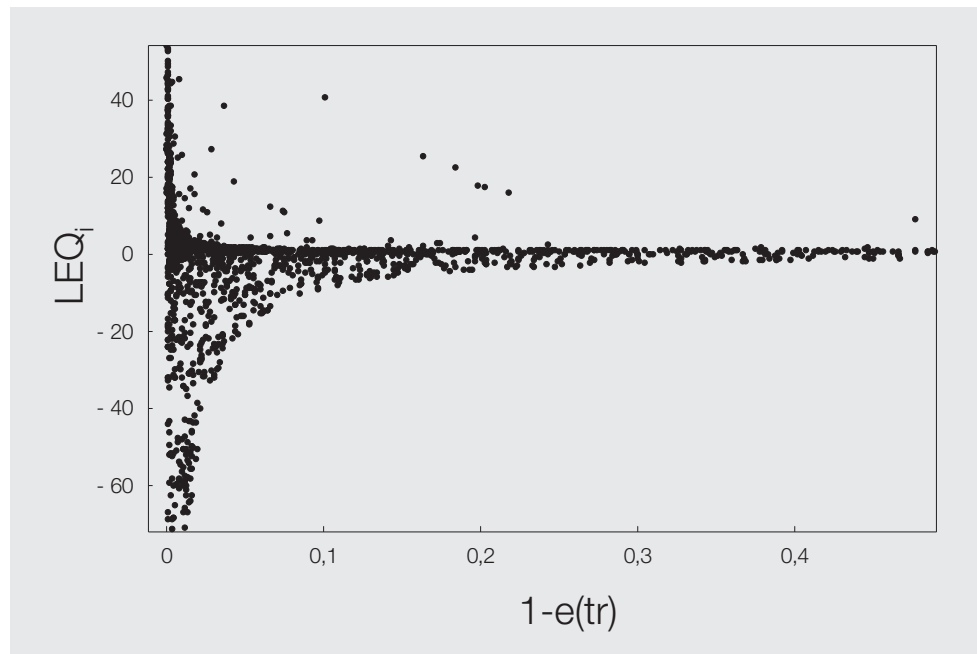
$$LEQ(f) = LEQ = \frac{1}{n} \sum_i LEQ_i = -578 \quad [52]$$

Los problemas que aparecen al utilizar esta media global son debidos a:

- 1 La inestabilidad de ciertos factores LEQ observados: cuando  $1 - E(f)/L(f)$  es pequeño, los valores de los factores LEQ observados no son informativos.
- 2 La existencia de valores muy altos para ciertas observaciones, en algunos casos varias veces  $L(tr) - E(tr)$ . El tratamiento de estas observaciones requiere un análisis caso por caso.
- 3 Asimetrías en el comportamiento de los factores LEQ observados (entre los valores negativos y los positivos).
- 4 Existe evidencia empírica de que la media muestral de los  $LEQ_i$  no es constante, dependiendo de los valores de  $1 - E(f)/L(f)$ .

La figura 15 representa la distribución de los factores LEQ observados y los disponibles en porcentaje del límite. Esta figura puede ayudar a entender los principales problemas asociados con este método.

La figura 16 se centra en las observaciones asociadas con valores de los factores LEQ observados menores que dos. Es claro que hay observaciones mayores que uno (zona sombreada superior en las figuras 16 y 17) a lo largo de todo el rango de valores de utilización, aunque tales observaciones son mucho más frecuentes cuando los valores de la utilización porcentual del crédito en la fecha de referencia son muy altos (valores pequeños de  $1 - e(tr)$ ).



Por estas razones, después de analizar las observaciones en el RDS es necesario tomar ciertas decisiones; por ejemplo:

- Eliminar del RDS las observaciones anómalas asociadas con valores de los factores LEQ observados muy grandes.
- Censurar los valores de otras observaciones asociadas con factores  $LEQ_i$  mayores que uno.
- Eliminar del RDS las observaciones asociadas con valores muy bajos de  $E(f) - L(f)$ , ya que los factores LEQ observados no son en esos casos informativos.

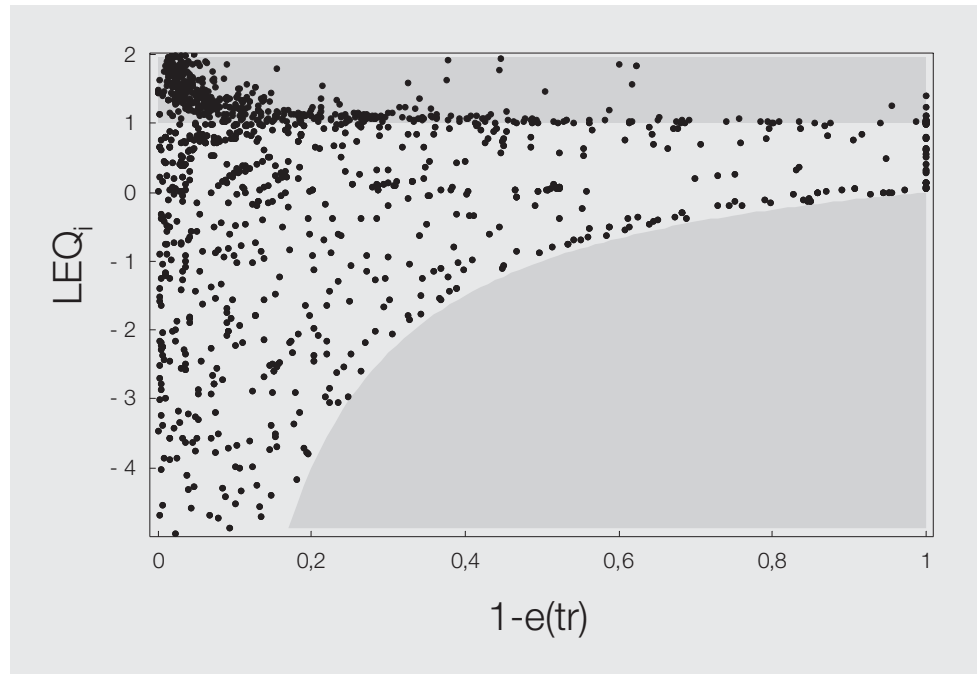
En este ejemplo, se han eliminado del RDS las observaciones con  $1 - E(tr)/L(tr) \leq 0,1$  y aquellas asociadas a factores  $LEQ_i \geq 2$ . Después de estas modificaciones del RDS, el nuevo valor para la media muestra de los factores LEQ observados es:

$$LEQ(f) = LEQ = \frac{1}{m} \sum_i LEQ_i = 0,08 \quad [53]$$

Es evidente que esta estimación global del 8% para el factor LEQ es muy baja para muchas de las operaciones de la cartera, debido al peso que tienen en la media global los factores LEQ negativos asociados con observaciones con valores muy bajos de  $1 - e(f)$ .

Una mejora frente a la anterior estimación es eliminar *outliers*, es decir, observaciones asociadas con valores extremos de los factores LEQ observados. Si las observaciones asociadas a los factores LEQ observados por debajo del percentil 10% y por encima del percentil del 90% se consideran *outliers*, la media restringida al RDS sin *outliers* es de un 33% y, este valor es estable frente a cambios en los percentiles anteriores.

$$LEQ(f) = LEQ = \frac{1}{r} \sum_i LEQ_i = 0,33 \quad [54]$$



No obstante, las medias muestrales locales (calculadas para un rango de valores de  $1 - E(f)/L(f)$ ) son muy diferentes y, por lo tanto, esta estimación global del 33% para el factor LEQ no es adecuada para todas las operaciones existentes en la cartera actual. Por esta razón es necesario considerar diferentes estimaciones para el factor LEQ en función de los diferentes valores de la disponibilidad,  $1 - E(f)/L(f)$ .

b. Caso  $RDS(f) = \{O_i \text{ tales que } e_i \text{ sea «similar» a } e(f)\}$

En este caso, el  $RDS(f)$  agrupa todas las observaciones  $O_i$  para las que se verifica que  $1 - e(tr) \in [1 - e(f) - 0,2, 1 - e(f) + 0,2]$  y después se utiliza como estimador del factor LEQ para la operación  $f$ ,  $LEQ(f)$ , la media muestral de los factores LEQ observados restringida a las observaciones que pertenecen a  $RDS(f)$ . Para seleccionar una forma funcional para  $LEQ(f)$ , primero se computan los valores estimados para las diferentes observaciones en  $RDS(f)$  y a continuación se ajusta un modelo de regresión que utiliza como variable explicativa  $1 - e(tr)$  y la correspondiente media local como variable dependiente. Después de rechazar varios modelos y utilizando intervalos de anchura 0,4, se obtuvo una expresión para la media muestral local<sup>37</sup> de los factores LEQ observados, basada en  $a + b \cdot \sqrt{1 - e(tr)}$ , dada por:

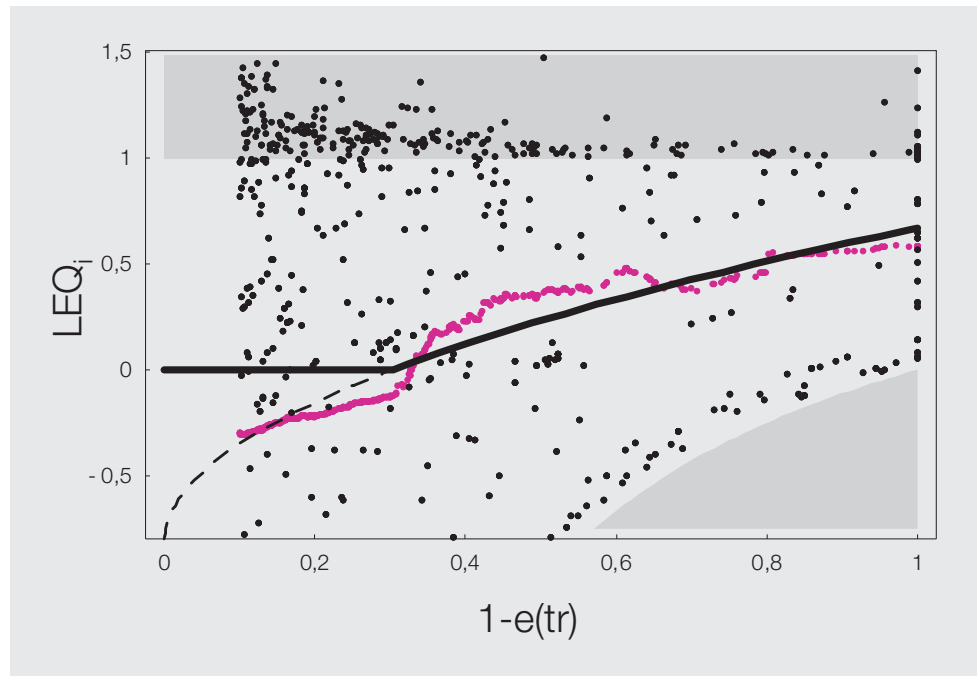
$$LEQ(f) = -0,82 + 1,49 \cdot \sqrt{1 - E(f)/L(f)} \quad [55]$$

Con un  $R^2$  ajustado de 0,94. La figura 17 representa los factores LEQ observados, las medias locales y la función ajustada con la restricción  $LEQ(f) > 0$ .

Por lo tanto, un estimador para la EAD de la operación  $f$ ,  $EAD(f)$ , que se encuentra en estatus normal, viene dado por:

$$EAD(f) = E(f) + \text{Max} \left[ 0, \left( -0,82 + 1,49 \cdot \sqrt{1 - E(f)/L(f)} \right) \cdot (L(f) - E(f)) \right] \quad [56]$$

37. La condición «local» es tener en cuenta solo aquellas observaciones comprendidas en un intervalo centrado en  $1 - e(f)$  y con longitud 0,4.



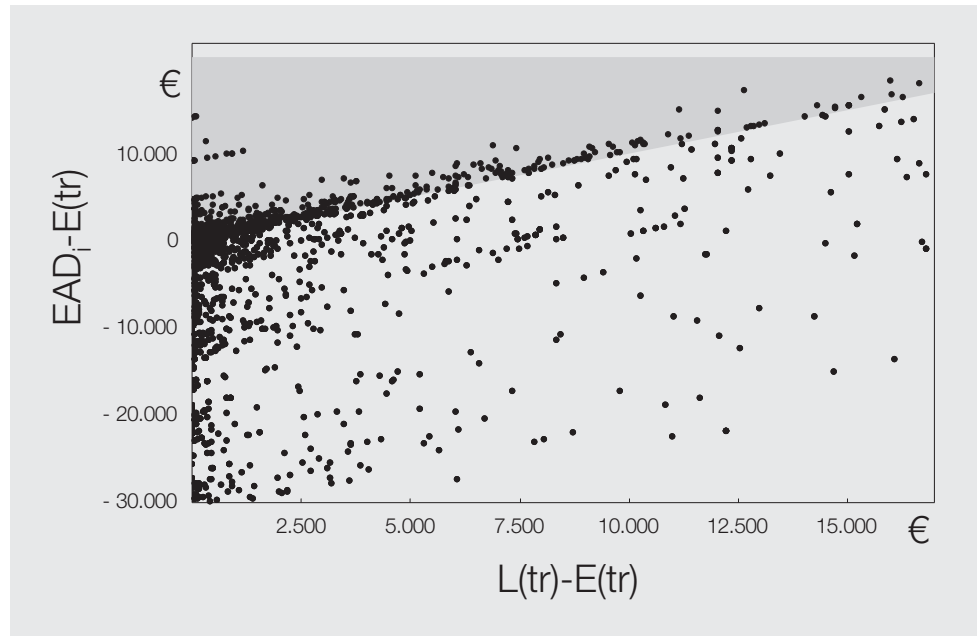
7.2.3 La mediana y los cuantiles condicionados

La lógica subyacente al modelo III es explicar directamente el incremento en la exposición desde la fecha de referencia hasta la fecha de *default*. Por lo tanto, se trata de explicar  $EAD_i - E(tr)$  en términos de  $LEQ(RD_i) * (L(tr) - E(tr))$ . Por simplicidad, se asume en este ejercicio que  $RD_i = \{S(tr), L(tr) - E(tr)\}$ , y se centra el interés en observaciones que se encuentran en estatus normal  $S(tr) = N$ , y la única variable cuantitativa que se tiene en cuenta es el disponible actual en euros  $L(f) - E(f)$ . Además, la función de pérdida propuesta en [39] se utiliza para determinar las estimaciones óptimas y, por lo tanto, como se demostró en la sección 6.3.1, la solución es aproximar el cuantil  $Q[b/(a+b)]$  de la distribución de  $EAD_i - E(tr)$  condicionada a aquellas observaciones que cumplen  $L(tr) - E(tr) = L(f) - E(f)$ . Para aproximar ese cuantil para cada valor de  $EAD(f) - E(f)$ , el proceso es similar al anteriormente explicado en la sección previa para aproximar las esperanzas condicionadas. En primer lugar,  $RDS(f)$  se define como el subconjunto de RDS que contiene todas las observaciones que verifican  $(L(tr) - E(tr)) \in [(L(f) - E(f)) * 0,8, (L(f) - E(f)) * 1,2]$ . En segundo lugar, para cada valor de  $(L(tr) - E(tr))$  se calcula el cuantil óptimo. En tercer lugar, se ajusta un modelo de regresión que utiliza  $(L(tr) - E(tr))$  como variable explicativa y el valor del correspondiente cuantil óptimo como variable dependiente y, finalmente, el estimador  $LEQ(f)$  se obtiene aplicando la fórmula [44].

La figura 18 representa, para cada observación en el RDS, con  $tr = td - \{12, 11, 10, 9, 8\}$ , las parejas  $\{L(tr) - E(tr), EAD_i - E(tr)\}$  en el rango de valores de  $L(tr) - E(tr)$  dado por  $[0, 17.000]€$ , para el que se considera que se dispone de un número suficiente de observaciones. La zona sombreada en las figuras 18 y 19 está definida por la desigualdad  $EAD_i \geq L(tr)$

Los resultados de la regresión para las medianas locales (caso  $a = b$ ) y para el percentil 66,66% (caso  $2*a = b$ ) dan lugar a los siguientes resultados:

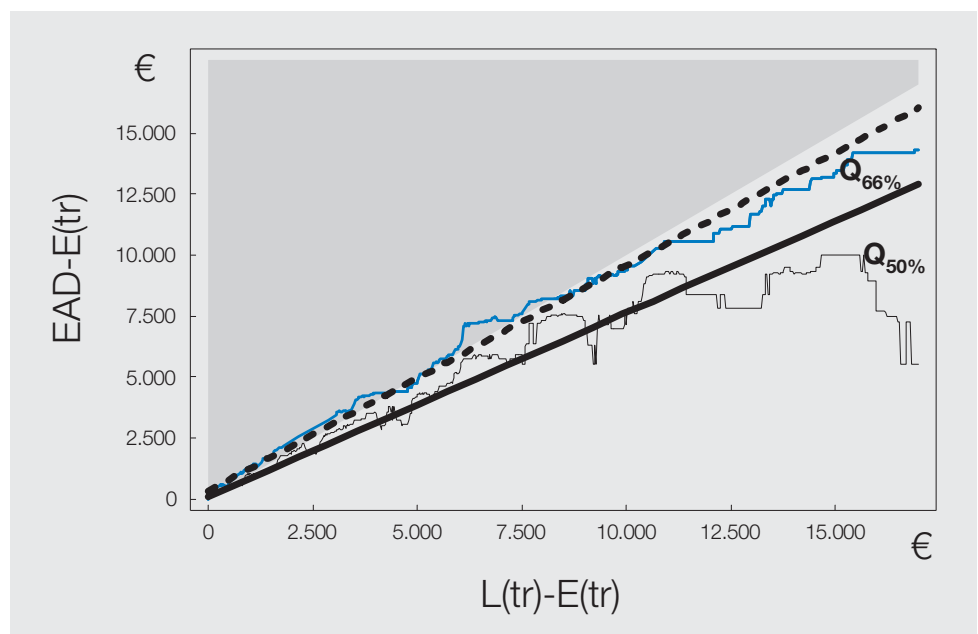
$$\begin{aligned} \text{Mediana}[EAD(f) - E(f)] &= 86,8 + 0,76 * (L(f) - E(f)) \\ \text{Cuantil}[EAD(f) - E(f), 0,666] &= 337,8 + 0,92 * (L(f) - E(f)) \end{aligned} \quad [57]$$



con  $R^2$  ajustados iguales a 0,95 y 0,99, respectivamente. Por lo tanto, las correspondientes estimaciones de los factores LEQ, obtenidos dividiendo [57] entre  $(L(f) - E(f))$ , son prácticamente constantes (factores LEQ próximos a 0,76 y a 0,92, respectivamente) y tienen valores mayores que los de las estimaciones previas.

La figura 19 representa las medianas locales (línea  $Q_{50\%}$ ) y los percentiles (66,6%) locales (línea  $Q_{66\%}$ ) obtenidos a partir de las observaciones originales y las rectas de regresión asociadas a [57] [línea discontinua para los percentiles (66,6%) ajustados y línea gruesa para las medianas ajustadas].

#### CUANTILES CONDICIONADOS AL DISPONIBLE Y VALORES AJUSTADOS DE EAD - E(tr)



## 8 Resumen y conclusión

- Los puntos siguientes resumen la práctica actual relativa a estimaciones de los CF y la EAD, y destacan algunos aspectos problemáticos:
  - Los estimadores aplicados por los bancos se pueden obtener como soluciones a casos particulares de problemas de regresión y, por lo tanto, esos estimadores se basan en esperanzas condicionadas.
  - Implícitamente, el uso de esos estimadores asume que el objetivo es la minimización de los errores de predicción utilizando una función de pérdida cuadrática y simétrica que ni está directamente correlacionada con los errores en términos de requisitos de capital, ni penaliza la incertidumbre. Estos errores son grandes y, por lo tanto, las decisiones sobre cómo medirlos son cruciales.
  - En muchos de los casos, la estimación de la EAD se fundamenta en la hipótesis no realista de que los factores LEQ promedio son constantes.
  - Frecuentemente, los estadísticos básicos para el proceso de estimación se censuran para obtener factores LEQ observados entre cero y uno.
  - Los bancos utilizan los procedimientos denominados «enfoque de cohortes» y «enfoque de horizonte fijo» para seleccionar las observaciones incluidas en el proceso de estimación. Estos enfoques no tienen en cuenta toda la información relevante, ya que se centran solamente en una fecha de referencia convencional para cada operación que ha entrado en *default*.
  - Variables explicativas: se centran en el *rating* en la fecha de referencia.
- Otras posibilidades y algunos comentarios sobre diferentes aspectos:
  - Para uso regulatorio, parece apropiado que las estimaciones sean soluciones de problemas de optimización que utilizan una función de pérdida directamente relacionada con los errores en términos de requisitos de capital.
  - Por ejemplo, una elección lógica es utilizar una función de pérdida lineal y asimétrica aplicada a nivel de operación. Esta función de pérdida permite cuantificar el nivel de prudencia implícito en las estimaciones.
  - Utilizando esta función de pérdida, los estimadores óptimos están basados en cuantiles condicionados (por ejemplo, la mediana para usos internos y un cuantil más conservador para uso regulatorio).
  - Si las estimaciones se basan en medias muestrales, como mínimo, los factores LEQ estimados deberían depender del valor de la disponibilidad:  $LEQ(1 - e(tr))$ .
  - La práctica común de censurar los factores LEQ observados a  $[0,1]$  no está justificada y, en general, no es posible concluir ex ante si las correspondientes estimaciones de los LEQ están sesgadas de forma conservadora.
  - No obstante, bajo ciertas hipótesis el uso de datos censurados no cambia los estimadores óptimos del LEQ.

- Las estimaciones deberían basarse en observaciones de las operaciones que han entrado en *default* para todas las fechas de referencia relevantes (método de horizonte variable).
- Variables explicativas: si existen sistemas de alerta en la cartera, es importante fijarse en el estatus de la operación en la fecha de referencia más que en el *rating*.
- El ejemplo presentado sugiere que:
  - Las estimaciones basadas en medias muestrales son mucho menos conservadoras que las que utilizan cuantiles condicionados (por encima de la mediana) y, por lo tanto, el impacto práctico del nuevo método es grande.
  - Las estimaciones de los CF obtenidas utilizando cuantiles condicionados son suficientemente grandes y, en consecuencia, de utilizarse este método, no parece que el uso de estimaciones *downturn* sea una prioridad para este parámetro.

## BIBLIOGRAFÍA

- ARATEN, M., y M. JACOBS (2001). «Loan Equivalents for Revolving Credits and Advised Lines», *The RMA Journal*, pp. 34-39.
- BASEL COMMITTEE ON BANKING SUPERVISION (2005). *Guidance on Paragraph 468 of the Framework Document*. – (2004). *International Convergence of Capital Measurement and Capital Standards. A Revised Framework*. – (2005). *Studies on the Validation of Internal Rating Systems*, Working Paper No. 14.
- CEBS (2006). *Guidelines on the implementation, validation and assessment of Advanced Measurement (AMA) and Internal Ratings Based (IRB) Approaches*, CP 10 revised.
- LEV, B., y S. RAYAN (2004). *Accounting for commercial loan commitments*.
- MORAL, G. (1996). «Pérdida latente, incertidumbre y provisión óptima», *Boletín de la Inspección de ECA*, Banco de España.
- Office of the Comptroller of the Currency (OCC), Treasury; Board of Governors of the Federal Reserve System (Board); Federal Deposit Insurance Corporation (FDIC); and Office of Thrift Supervision (OTS), Treasury. Federal (2003), *Draft Supervisory Guidance on Internal Ratings – Based Systems for Corporate Credit (corporate IRB guidance)*, Register / vol. 68, n.º 149.
- Office of the Comptroller of the Currency (OCC), Treasury; Board of Governors of the Federal Reserve System (Board); Federal Deposit Insurance Corporation (FDIC); and Office of Thrift Supervision (OTS) (2004). *Proposed supervisory guidance for banks, savings associations, and bank holding companies (banking organizations) that would use the internal – ratings – based (IRB) approach to determine their regulatory capital requirements for retail credit exposures*, Treasury, Federal Register / vol. 69, n.º 207.
- PRATT, J., H. RAIFFA y R. SCHLAIFER (1995). *Introduction to Statistical Decision Theory*, The MIT Press.
- SUFI, A. (2005). *Bank Lines of Credit in Corporate Finance: An Empirical Analysis*, University of Chicago Graduate School of Business.

### ANEJO 1

#### **Equivalencia entre dos problemas de minimización**

Véase *Appendix A* en el original inglés publicado por B. Engelmann y R. Rauhmeier (eds.) (2006), *The Basel II Risk Parameters: Estimation, Validation, and Stress Testing*, Springer Berlin Heidelberg, Nueva York.

### ANEJO 2

#### **Soluciones óptimas de algunos problemas de regresión y optimización**

Véase *Appendix B* en el original inglés.

**Diagnóstico de los modelos de regresión utilizados en el ejemplo**

– Utilizando los datos originales:

$$\frac{EAD_i}{L(td-12)} - \frac{E(td-12)}{L(td-12)} = 0,64 * \left(1 - \frac{E(td-12)}{L(td-12)}\right) \quad [69]$$

		ESTIMACIÓN	SE	TSTAT	PVALUE
Tabla de parámetros	x	0,64	0,08	8,03	0
		R CUADRADO	R CUAD. AJUSTADO	VARIANZA ESTIMADA	
Grado de ajuste		0,13	0,13	0,21	
TABLA ANOVA					
	DF	SUMOF SQ	MEANSQ	FRATIO	PVALUE
Modelo	1	13,5	13,54	64,45	0
Error	416	87,41	0,21		
Total	417	100,95			

– Utilizando datos censurados:

$$\frac{EAD_i}{L(td-12)} - \frac{E(td-12)}{L(td-12)} = 0,7 * \left(1 - \frac{E(td-12)}{L(td-12)}\right) \quad [70]$$

		ESTIMACIÓN	SE	TSTAT	PVALUE
Tabla de parámetros	x	0,7	0,02	35,07	0
		R CUADRADO	R CUAD. AJUSTADO	VARIANZA ESTIMADA	
Grado de ajuste		0,75	0,75	0,013	
TABLA ANOVA					
	DF	SUMOF SQ	MEANSQ	FRATIO	PVALUE
Modelo	1	16,49	16,49	1.229,67	0
Error	416	5,58	0,013		
Total	417	22,05			

– Utilizando un enfoque de horizonte variable:

$$\frac{EAD_i}{L(tr)} - \frac{E(tr)}{L(tr)} = 0,49 * \left(1 - \frac{E(tr)}{L(tr)}\right) \quad [71]$$



		ESTIMACIÓN	SE	TSTAT	PVALUE
Tabla de parámetros	x	0,49	0,04	13,86	0
		R CUADRADO	R CUAD. AJUSTADO	VARIANZA ESTIMADA	
Grado de ajuste		0,06	0,06	0,19	
TABLA ANOVA					
	DF	SUMOF SQ	MEANSQ	FRATIO	PVALUE
Modelo	1	35,92	16,49	1.229,67	0
Error	2.918	545,3	0,19		
Total	2.919	581,2			

### Modelo I (sección 7.2.2)

– Utilizando el modelo I y un enfoque de horizonte variable:

$$LEQ(f) = -0,82 + 1,49 \cdot \sqrt{1 - E(f)/L(f)} \quad [72]$$

Los diagnósticos para el modelo de regresión son:

		ESTIMACIÓN	SE	TSTAT	PVALUE
Tabla de parámetros	1	-0,82	0,009	-93,09	0
	x^0,5	1,49	0,014	104,57	0
GRADO DE AJUSTE		R CUADRADO	R CUAD. AJUSTADO	VARIANZA ESTIMADA	
		0,94	0,94	0,006	
TABLA ANOVA					
	DF	SUMOF SQ	MEANSQ	FRATIO	PVALUE
Modelo	1	66,13	66,13	10.934,1	0
Error	663	4,01	0,006		
Total	664	70,14			

### Modelo III (sección 7.2.3)

– Utilizando un enfoque de horizonte variable:

$$\begin{aligned} \text{Mediana}[EAD(f) - E(f)] &= 86,8 + 0,76 * (L(f) - E(f)) \\ \text{Cuantil}[EAD(f) - E(f), 0,666] &= 337,8 + 0,92 * (L(f) - E(f)) \end{aligned} \quad [73]$$

Con los diagnósticos, para la mediana, dados por:

		ESTIMACIÓN	SE	TSTAT	PVALUE
Tabla de parámetros	1	86,8	11,23	7,73	0
	x	0,76	0,003	222,74	0
GRADO DE AJUSTE	R CUADRADO		R CUAD. AJUSTADO	VARIANZA ESTIMADA	
		0,95	0,95	227.741	
TABLA ANOVA					
	DF	SUMOF SQ	MEANSQ	FRATIO	PVALUE
Modelo	1	1,13*10 <sup>10</sup>	1,13*10 <sup>10</sup>	49.611	0
Error	2.370	5,40*10 <sup>8</sup>	227.741		
Total	2.371	1,18*10 <sup>10</sup>			

Y, para el cuantil, por:

		ESTIMACIÓN	SE	TSTAT	PVALUE
Tabla de parámetros	1	337,8	5,14	65,7	0
	x	0,92	0,002	594,6	0
GRADO DE AJUSTE	R CUADRADO		R CUAD. AJUSTADO	VARIANZA ESTIMADA	
		0,99	0,99	47.774,6	
TABLA ANOVA					
	DF	SUMOF SQ	MEANSQ	FRATIO	PVALUE
Modelo	1	1,69*10 <sup>10</sup>	1,7*10 <sup>10</sup>	353.621	0
Error	2.370	1,13*10 <sup>8</sup>	47.774,6		
Total	2.371	1,7*10 <sup>10</sup>			

#### ANEJO 4

#### Abreviaturas

AIRB	Enfoque avanzado basado en <i>ratings</i> internos.
CCF	Factor de conversión del crédito. <i>Credit conversion factor</i> .
CF	Factor de conversión. <i>Conversion factor</i> .
EAD	Exposición en el momento del <i>default</i> . <i>Exposure at default</i> .
$EAD_i = E(td)$	Exposición observada en el momento del <i>default</i> asociada a la observación $O_i$ .
EAD(f)	Estimación de la EAD para la operación f.
$ead_i$	Exposición porcentual observada en el momento del <i>default</i> asociada a $O_i$ .
E(t)	Exposición de una operación en la fecha t.
e(t)	Utilización porcentual de una operación en la fecha t.
$e_i = e(tr)$	Utilización porcentual asociada a la observación $O_i$ .
f	Operación que no ha entrado en <i>default</i> .
g	Operación que ha entrado en <i>default</i> .
$i = \{g, tr\}$	Índice asociado a la observación de la operación g en la fecha tr.
IRB	Enfoque basado en <i>ratings</i> internos.
LEQ	Factor LEQ. <i>Loan equivalent exposure</i> .
LEQ(f)	Estimación del factor LEQ para la operación f.

LEQ <sub>i</sub>	Factor LEQ observado asociado a la observación O <sub>i</sub> .
LGD	Pérdida en caso de incumplimiento. <i>Loss given default</i> .
L(t)	Límite explícito de crédito de una operación en el momento t.
O <sub>i</sub>	Observación asociada con la pareja i = {g, tr}.
PD	Probabilidad de <i>default</i> .
Q <sub>a</sub> = Q(a,x)	Cuantil asociado con el a% de la distribución F(x).
RDS	Conjunto de datos de referencia. <i>Reference data set</i> .
RDS(f)	Conjunto de datos de referencia asociado a f.
RD	Variables explicativas. <i>Risk drivers</i> .
S(tr)	Estatus de una operación en la fecha de referencia tr.
t	Fecha actual.
td	Fecha de <i>default</i> .
tr	Fecha de referencia.
td – tr	Horizonte.