



Crisis bancarias como eventos infrecuentes

MARÍA DE FÁTIMA LEÓN, ASTRID VARGAS Y DIEGO WINKELRIED*

En este documento se argumenta que las estimaciones provenientes de la extensa literatura sobre los determinantes de la probabilidad de ocurrencia de crisis bancarias podrían estar significativamente sesgadas. El sesgo se asocia con el simple hecho de que las crisis bancarias son eventos raros o infrecuentes. De hecho, una vez que se corrigen los sesgos mencionados, los modelos probabilísticos comúnmente utilizados para anticipar la ocurrencia de un episodio de crisis bancaria ven incrementado su poder predictivo sustancialmente. Ello es relevante en el diseño de sistemas de alerta temprana, ya que los enormes costos asociados con la ocurrencia de una crisis bancaria podrían ser reducidos si ésta es oportunamente identificada.

Palabras Clave : Crisis bancaria, evento raro, sesgo, *relogit*.

Clasificación JEL : C25, C53, G21, H12.

Las crisis bancarias son, sin dudas, eventos de gran importancia en la actividad financiera y macroeconómica de un país. De acuerdo con [Reinhart y Rogoff \(2014\)](#), históricamente estas crisis anteceden recesiones caracterizadas por reducciones significativas del PBI per cápita (1.5 por ciento, en promedio), incluso varios años después de ocurrido el episodio (3 años, en promedio). Asimismo, lo usual es que estas crisis se traduzcan en cuantiosos costos de rescate de las instituciones financieras afectadas. Por ejemplo, y también tomando cifras de [Reinhart y Rogoff \(2014\)](#), en el mejor de los casos como en la crisis de 1984 en los Estados Unidos, estos costos ascenderían a 3 por ciento del PBI. No es raro, sin embargo, encontrar episodios mucho más traumáticos, como la crisis experimentada en Chile en 1981 (entre 29 y 41 por ciento del PBI) o la registrada en Japón durante 1992 (entre 8 y 24 por ciento del PBI). Más aún, las crisis bancarias usualmente predicen o acompañan crisis cambiarias ([Kaminsky y otros, 1998](#)) o crisis de deuda soberana ([Reinhart y Rogoff, 2011](#)), lo que podría incrementar significativamente (más que duplicar) sus efectos adversos sobre la economía.

No es de sorprender, luego, que exista un gran interés entre académicos y hacedores de política en explorar y entender cuáles son las precondiciones asociadas con la ocurrencia de una crisis bancaria. Si bien la literatura empírica al respecto es amplia (véase, entre otros, [Frydl, 1999](#); [Von Hagen y Ho, 2007](#); [Domac y Martínez, 2003](#); [Laeven, 2011](#)), en este estudio prestaremos especial atención a [Demirgüç-Kunt](#)

* León: Universidad del Pacífico (Lima, Perú), Av. Salaverry 2020, Lima 11, Perú. Teléfono: +511 219-0100 (e-mail: mariadefatima.leon@gmail.com). Vargas: Universidad del Pacífico (Lima, Perú) (e-mail: astrid.vargase@gmail.com). Winkelried: Universidad del Pacífico (Lima, Perú) (e-mail: winkelried_dm@up.edu.pe).

Agradecemos los comentarios de Daniel Barco, Renzo Castellares, Javier Torres, Pablo Lavado, participantes del XXXIII Encuentro de Economistas del BCRP y de un revisor anónimo. Las opiniones aquí vertidas, así como cualquier error u omisión, son de entera responsabilidad de los autores.

y Detragiache (1998, 2002, 2005a, 2010). Estas autoras, en una secuencia de influyentes estudios, estiman modelos probabilísticos logit, y encuentran que la probabilidad de ocurrencia de una crisis bancaria es explicada por los factores macroeconómicos (crecimiento económico, fluctuaciones en los términos de intercambio, inflación, depreciación, tasa de interés real), las variables financieras (crédito privado como porcentaje del PBI, crecimiento del crédito, ratio de reservas líquidas) y los factores institucionales (la existencia de un seguro de depósitos, la efectividad del sistema judicial) que son tomados, hoy por hoy, como los principales determinantes de crisis bancarias.

Las propiedades estadísticas de los estimadores de modelos probabilísticos, tipo logit, son bastante conocidas (Cordeiro y McCullagh, 1991). En particular, estos estimadores son consistentes por lo que el uso de muestras de gran tamaño lleva a desestimar, para fines de inferencia y predictivos, los sesgos típicos de muestras finitas. No obstante, King y Zeng (2001a,b, 2002) muestran que estos sesgos se ven exacerbados cuando el modelo probabilístico intenta explicar un *evento raro o infrecuente*. La magnitud del sesgo puede ser lo suficientemente significativa como para distorsionar los resultados derivados del modelo logit estándar; por ejemplo, la probabilidad de ocurrencia del evento infrecuente usualmente se encuentra subestimada. King y Zeng (2001a,b) proponen los ajustes necesarios para corregir estos sesgos, dando origen al estimador conocido como *relogit (Rare Events logit)*. Véase, además, Tomz y otros (2003).

Esta observación metodológica es relevante para el caso de crisis bancarias ya que, además de ser costosas, se tratan también de eventos infrecuentes. Por ejemplo, en los datos de Demirgüç-Kunt y Detragiache (1998), cuya muestra incluye 65 países en el periodo de 1980 a 1994, se observa que, en promedio, solo entre 3 y 5 por ciento de las observaciones representa el inicio de una crisis bancaria. Asimismo, para los datos de Reinhart y Rogoff (2014) de un total de 70 países durante un periodo de 211 años (de 1800 a 2010), en promedio, cada país ha experimentado cuatro crisis bancarias; es decir, una crisis cada cincuenta años. De este modo, existen indicios que sugieren que las estimaciones logit en la literatura de crisis bancarias (Demirgüç-Kunt y Detragiache, 1998, 2010) podrían presentar sesgos considerables, típicamente asociados con una *subestimación* de la probabilidad de ocurrencia de una crisis. Específicamente, King y Zeng (2001a) muestran que las correcciones a las estimaciones de los modelos probabilísticos que proponen son mayores cuando el número de observaciones es reducido (digamos, alrededor de unos cuantos miles) y los eventos bajo análisis presentan una frecuencia muestral menor al 5 por ciento, como es el caso de las crisis bancarias.

Es bueno mencionar que el estimador relogit ha sido aplicado con éxito en otras disciplinas, sobre todo en las ciencias políticas,¹ y en mucho menor medida en microeconomía.² No obstante, así como en el caso de las crisis bancarias, muchos eventos macroeconómicos de importancia son infrecuentes (recesiones, *booms* de precios, entre otros) y el uso del estimador relogit podría constituir una práctica más general en la profesión.³

De esta manera, el objetivo de este trabajo es estimar la probabilidad de una crisis bancaria y los efectos de sus determinantes, considerando explícitamente que las crisis bancarias son eventos infrecuentes. En

¹ Hurwitz y Stefko (2004) evalúan el número de veces en la que la Corte Suprema en EEUU vota a favor de un caso favorecido por la existencia de un precedente. Martin (2005) estudia el número de veces que, en los EEUU, un presidente opta por seguir un procedimiento legislativo para llevar a cabo acuerdos internacionales, los cuales requieren de la aprobación de 2/3 del senado. Buhag (2005) analiza la probabilidad de que dos países democráticos entren en un conflicto armado. Finalmente, Fuhrmann y Kreps (2010) estudian el número de veces en que un estado considera o decide atacar con tecnología nuclear a un país que no cuenta con tal tecnología.

² Hennig-Thurau y otros (2007) estudian el número de veces en el que un individuo ve una película en un canal comercial (cine, renta o compra de DVD). Röller y otros (2007) analizan el número de veces en que una compañía forma un *Research Joint Venture*. Finalmente, Cheng y otros (2013) estudian el atraso de salarios de trabajadores inmigrantes chinos.

³ Como aplicación en macroeconomía, solo conocemos el estudio de Hausmann y otros (2005), quienes analizan episodios en donde los países inician una etapa de crecimiento rápido y sostenido.

particular, indagamos sobre la importancia de los sesgos de muestras finitas en estimaciones logit de ocurrencia de crisis y, sobre todo, sobre el desempeño predictivo del estimador relogit. Utilizando datos similares a los de Demirgüç-Kunt y Detragiache (1998), observamos que la probabilidad no condicional de una crisis bancaria predicha con relogit se incrementa, en promedio, 4 puntos porcentuales con respecto de la probabilidad obtenida con estimaciones logit. Además, la capacidad de predicción del método relogit es significativamente mayor: relogit predice crisis bancarias más frecuentemente que logit. Asimismo, el modelo relogit es capaz de predecir la gran mayoría las crisis bancarias en la muestra utilizada, ya sea el mismo año de la crisis, uno o dos años antes. De hecho, las estimaciones hechas con relogit presentan un gran potencial para el diseño de sistemas de alerta temprana.

El resto del documento se organiza de la siguiente manera. En la sección 1 se describe la metodología relogit de King y Zeng (2001a,b). En la sección 2 se presenta la muestra de crisis bancarias, las variables explicativas consideradas en el análisis empírico y las fuentes de información utilizadas. En la sección 3 se presentan los resultados de las estimaciones logit y relogit, así como las predicciones obtenidas con cada método. La sección 4 presenta comentarios finales.

1 ASPECTOS METODOLÓGICOS

Sea y una variable aleatoria binaria, $y = \{0, 1\}$, que denota la ocurrencia de una crisis bancaria, y sea \mathbf{x} un vector de variables explicativas. El objeto de interés en modelos probabilísticos es la probabilidad del evento $y = 1$,

$$\Pr(y = 1 | \mathbf{x}) = G(\mathbf{x}'\boldsymbol{\beta}), \quad (1)$$

donde $\boldsymbol{\beta}$ es un vector de coeficientes desconocidos y $G(\cdot)$ es una *función de enlace* creciente y que toma valores entre 0 y 1. Para el caso del modelo logit, se tiene que

$$G(z) = \frac{1}{1 + \exp(-z)}. \quad (2)$$

El interés es predecir $\Pr(y = 1 | \mathbf{x})$ para lo que debe estimarse $\boldsymbol{\beta}$. La estimación de este vector se realiza por máxima verosimilitud, tomando como punto de partida $y | \mathbf{x} \sim \text{Bernoulli}(G(\mathbf{x}'\boldsymbol{\beta}))$. Así, para una muestra de n observaciones, la función de log-verosimilitud es

$$\mathcal{L}(\boldsymbol{\beta}) = \sum_{i=1}^n y_i \ln(G(\mathbf{x}_i'\boldsymbol{\beta})) + \sum_{i=1}^n (1 - y_i) \ln(1 - G(\mathbf{x}_i'\boldsymbol{\beta})). \quad (3)$$

El estimador de máxima verosimilitud de $\boldsymbol{\beta}$, el vector $\hat{\boldsymbol{\beta}}$ que maximiza (3), es en general sesgado, dada la no linealidad de la función de enlace. La forma del sesgo es, para casos generales, algo engorrosa (ver Cordeiro y McCullagh, 1991, para mayor detalle); sin embargo, King y Zeng (2001a) proveen una expresión bastante simple para un modelo de locación (únicamente se incluye un intercepto en el modelo), que ilustra con claridad cuáles son las fuentes de sesgo. A saber,

$$\mathbb{E}(\hat{\boldsymbol{\beta}}) \simeq \boldsymbol{\beta} + \frac{\boldsymbol{\pi} - 0.5}{n\pi(1 - \pi)}, \quad (4)$$

donde $\boldsymbol{\pi} = \Pr(y = 1)$. El segundo término de (4) es aproximadamente igual al sesgo. Esta expresión indica que, dado que los eventos infrecuentes satisfacen $\boldsymbol{\pi} < 0.5$, el sesgo de $\hat{\boldsymbol{\beta}}$ sería negativo, lo que lleva a concluir que la proporción muestral $\bar{y} = G(\hat{\boldsymbol{\beta}})$ sería también un estimador sesgado hacia abajo de la

verdadera probabilidad π . Por su parte, se aprecia que la magnitud del sesgo se relaciona con el conteo absoluto de éxitos del evento de estudio, cuya esperanza está dada por $n\pi$. Así, si el evento es infrecuente, en el sentido que el número de éxitos es muy pequeño, se exagera el sesgo en el estimador logit. Si, como es usual, se asume que $n \rightarrow \infty$ manteniendo π fijo, entonces el sesgo se desvanece, reflejando así la consistencia del estimador. Ocurre lo mismo si $n\pi \rightarrow \infty$ conforme $n \rightarrow \infty$. No obstante, una aproximación asintótica más apropiada para un evento infrecuente sería que $n\pi \rightarrow \alpha < \infty$ conforme $n \rightarrow \infty$, en cuyo caso el estimador logit sería incluso inconsistente.

Como se discute en King y Zeng (2002), podría pensarse que el sesgo resulta importante porque, en una muestra típica, las observaciones de éxito ($y = 1$) se encuentran subrepresentadas, dada la infrecuencia del evento. Y es precisamente el desbalance entre observaciones con $y = 1$ y observaciones con $y = 0$ lo que magnifica posibles sesgos en muestras finitas. Así, una posible solución es balancear artificialmente los datos, descartando al azar observaciones con $y = 0$. Ello es, en muchas dimensiones, equivalente a reponderar las observaciones y darle más importancia a las que cuenten con $y = 1$. En (3) cada observación contribuye a la verosimilitud de la muestra de la misma manera, es decir cada observación es tan importante como cualquier otra. La propuesta es replantear el problema de modo que la función objetivo del problema de estimación pase a ser

$$\tilde{\mathcal{L}}(\boldsymbol{\beta}) = \sum_{i=1}^n w_1 y_i \ln(G(\mathbf{x}_i' \boldsymbol{\beta})) + \sum_{i=1}^n w_0 (1 - y_i) \ln(1 - G(\mathbf{x}_i' \boldsymbol{\beta})), \quad (5)$$

donde se le asigna un peso w_1 a las observaciones con $y = 1$ y un peso w_0 a las observaciones con $y = 0$. Toda vez que $w_1 > w_0$, se estaría “rebalanceando” la estimación, y en (5) las observaciones con $y = 1$ se tratarían como más informativas que las con $y = 0$. La elección

$$w_1 = \frac{\pi}{\bar{y}} \quad \text{y} \quad w_0 = \frac{1 - \pi}{1 - \bar{y}}, \quad (6)$$

donde \bar{y} es la proporción muestral de unos, mientras que π es la probabilidad de éxito poblacional, elimina el sesgo en modelos de locación, y lo reduce a órdenes de magnitud insignificantes en modelos más generales (King y Zeng, 2001a). Note que si $\bar{y} < \pi$ (la condición de infrecuencia del evento), entonces $w_1 > w_0$. Luego, el estimador relogit es el vector $\tilde{\boldsymbol{\beta}}$ que maximiza (5).

Es bueno notar que π es información *a priori*, por lo que no puede ser inferida con los datos de la muestra, lo que da un sabor Bayesiano al procedimiento de corrección de sesgos. De hecho, lo usual es que \bar{y} se aproxime a π por lo que las diferencias entre (3) y (5) serían reducidas. Note que $\pi = \bar{y}$ produce exactamente el modelo logit. No obstante, si en efecto el evento está subrepresentado en la muestra, $\bar{y} < \pi$ con alta probabilidad, entonces conviene implementar la corrección. Es recomendable, en la práctica, estudiar la sensibilidad de los resultados a la elección de π .

El interés, como se mencionó, es predecir la probabilidad de ocurrencia de una crisis, que denotamos como $G(\boldsymbol{\beta}) \equiv G(\mathbf{x}' \boldsymbol{\beta})$ por brevedad. El predictor $G(\tilde{\boldsymbol{\beta}})$ es sesgado, dada la no linealidad de $G(\cdot)$, aunque es posible realizar una corrección por sesgo dado que se dispone del estimador relogit de $\boldsymbol{\beta}$, que es insesgado. Sea $g(\boldsymbol{\beta}) = G(\boldsymbol{\beta})[1 - G(\boldsymbol{\beta})]$ la primera derivada de $G(\cdot)$. Luego, por el teorema del valor medio,

$$G(\tilde{\boldsymbol{\beta}}) = G(\boldsymbol{\beta}) + g(\boldsymbol{\beta}) \mathbf{x}'(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) + \frac{1}{2} g(\bar{\boldsymbol{\beta}}) \left(2G(\bar{\boldsymbol{\beta}}) - 1 \right) \mathbf{x}'(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})' \mathbf{x},$$

donde cada elemento de $\bar{\boldsymbol{\beta}}$ se encuentra contenido entre los elementos correspondientes de $\tilde{\boldsymbol{\beta}}$ y $\boldsymbol{\beta}$. Al

tomar esperanzas y notar que $\tilde{\beta}$ es insesgado,

$$\mathbb{E} [G(\tilde{\beta})] = G(\beta) + g(\tilde{\beta}) \left(G(\tilde{\beta}) - \frac{1}{2} \right) \mathbf{x}' \mathbb{V}(\tilde{\beta}) \mathbf{x},$$

donde $\mathbb{V}(\tilde{\beta})$ es la matriz de covarianzas de $\tilde{\beta}$. Con ello, se deduce que

$$\tilde{\pi} = G(\tilde{\beta}) + g(\tilde{\beta}) \left(\frac{1}{2} - G(\tilde{\beta}) \right) \mathbf{x}' \mathbb{V}(\tilde{\beta}) \mathbf{x}, \quad (7)$$

es un predictor aproximadamente insesgado de la probabilidad deseada, $\mathbb{E} [\tilde{\pi}] \simeq G(\beta)$. Se aprecia que, en efecto, $G(\tilde{\beta})$ es sesgado. El segundo término de (7) es una corrección por sesgo. La dirección del sesgo en la probabilidad está determinada por el término $0.5 - G(\tilde{\beta})$. Típicamente, ante eventos infrecuentes ocurrirá que $G(\tilde{\beta})$ es menor que 0.5, por lo que la corrección incrementa la probabilidad predicha directamente, $G(\beta)$. Nos referiremos a $\tilde{\pi}$ como el predictor relógit de $\Pr(y = 1 | \mathbf{x})$.

2 DATOS Y VARIABLES

En nuestra exploración empírica se contrastan los resultados del método relógit aplicado a crisis bancarias contra los resultados obtenidos con regresiones logit similares a las de Demirgüç-Kunt y Detragiache (1998). La muestra corresponde inicialmente a los 65 países utilizados en Demirgüç-Kunt y Detragiache (1998), durante el periodo de 1980 a 1994. Sin embargo, 3 países fueron eliminados por contar con registros muy incompletos (El Salvador, Jamaica y Zaire), dando una muestra de 62 países.

Se trató de replicar la muestra de Demirgüç-Kunt y Detragiache (1998), utilizando las fuentes de información y definiciones sugeridas por estas autoras. Así, mayoritariamente se utiliza información de *International Financial Statistics* (IFS) del Fondo Monetario Internacional. No obstante, se recurrió a bases de datos adicionales como *World Bank DataBank* (DataBank), *IMF World Economic Outlook* (WEO), *World Development Indicators* (WDI) para completar valores omitidos en fuentes anteriores. Si bien no fuimos capaces de replicar exactamente los resultados de Demirgüç-Kunt y Detragiache (1998), claramente por no utilizar exactamente los mismos datos, las diferencias respecto al trabajo original son, para todo fin práctico, marginales.⁴

Variable dependiente

La variable dependiente es una variable binaria (*dummy*) que determina el *inicio* de una crisis bancaria. Demirgüç-Kunt y Detragiache (1998) declaran la existencia de una crisis bancaria a partir de cuatro componentes: (1) el ratio de activos improductivos respecto del total de activos del sector bancario debe ser mayor a 10 por ciento; (2) el costo de rescate de instituciones financieras debe ser al menos equivalente al 2 por ciento del PBI; (3) los problemas en el sector financiero resultan en nacionalización de bancos a gran escala; (4) ocurrencia de corridas bancarias y toma de medidas de emergencia como congelamiento de depósitos, feriados bancarios y elevación de las garantías de depósitos. La muestra contiene un total de 31 crisis, cuyos lugares, fechas de inicio y duración se muestran en el Cuadro 1 (p. 78).

Variables explicativas

Los determinantes de una crisis bancaria utilizados en esta investigación se agrupan en tres categorías: variables macroeconómicas, variables financieras y otras (variables institucionales o de largo plazo).

⁴ Dos de las variables consideradas por Demirgüç-Kunt y Detragiache no fueron incorporadas en nuestro análisis. Primero, el ratio del superávit del gobierno central entre PBI, ya que los registros en las diversas bases de datos resultaron ser bastante incompletos; segundo, un índice de aplicación de la ley, que se encuentra en una base de datos que no es de uso público.

CUADRO 1. Crisis bancarias por país, 1980 - 1994

| País | Año de inicio | Duración en años | País | Año de inicio | Duración en años |
|---|---------------|------------------|--------------------|---------------|------------------|
| Chile | 1981 | 4 | México | 1982 | 1 |
| Colombia | 1982 | 4 | | 1994 | 4 |
| Estados Unidos | 1981 | 12 | Nepal | 1988 | 7 |
| Filipinas | 1981 | 7 | Nigeria | 1991 | 4 |
| Finlandia | 1991 | 4 | Noruega | 1987 | 7 |
| Guyana | 1993 | 3 | Papúa Nueva Guinea | 1989 | 6 |
| India | 1991 | 4 | Senegal | 1983 | 6 |
| Indonesia | 1992 | 3 | Sri Lanka | 1989 | 5 |
| Israel | 1983 | 2 | Sudáfrica | 1985 | 1 |
| Italia | 1990 | 5 | Suecia | 1990 | 4 |
| Japón | 1992 | 3 | Tailandia | 1983 | 5 |
| Jordania | 1989 | 2 | Tanzania | 1988 | 7 |
| Kenia | 1993 | 1 | Turquía | 1994 | 1 |
| Malasia | 1985 | 4 | Uganda | 1990 | 5 |
| Mali | 1987 | 3 | Uruguay | 1981 | 5 |
| | | | Venezuela | 1993 | 2 |
| Número total de crisis | | | | 31 | |
| Proporción muestral \bar{y} en porcentaje | | | | 3.6 | |
| Duración promedio en años | | | | 4 | |

NOTAS: Respecto a las crisis estudiadas en Demirgüç-Kunt y Detragiache (1998), por falta de información financiera confiable para inicios de los 80s, se excluyen las crisis bancarias de Perú (1983 y 1986) y Turquía (1983).

En cuanto a las variables macroeconómicas, se plantea que bajas tasas de crecimiento del PBI, tasas de interés reales excesivamente altas, niveles altos de inflación y choques negativos en los términos de intercambio tienden a incrementar significativamente la probabilidad de crisis bancarias. Esta visión apunta a que las crisis no son consecuencias de profecías autocumplidas, sino de la evolución de variables fundamentales y del estado de la economía. Adicionalmente, la tasa de depreciación se utiliza para controlar por la posibilidad de que la crisis bancaria pueda ser impulsada por la exposición a excesivo riesgo cambiario, ya sea en el sistema bancario en sí o entre los prestatarios de los bancos.⁵ La definición de variables y fuentes de información de este grupo son:

- *Crecimiento*. Tasa de crecimiento del PBI real. Fuente: IFS línea 99, WEO código NGDP_RPCH.
- *Términos de intercambio*. Variación porcentual del índice precios de exportaciones, menos la variación porcentual del índice de precios de importaciones. Fuente: WEO, IFS línea 76.
- *Depreciación*. Variación porcentual del tipo de cambio nominal frente al US\$. Para EEUU, variación porcentual del tipo de cambio nominal efectivo. Fuente: IFS, línea AE.
- *Inflación*. Variación porcentual del deflactor del PBI. Fuente: IFS línea 99BI, Data Bank.
- *Tasa de interés*. Tasa de interés nominal menos la tasa de inflación contemporánea. Fuente: IFS: tasa nominal de títulos de corto plazo del gobierno (línea 61a); de no estar disponible, tasa que cobra el banco central a los bancos domésticos como tasa de descuento (línea 60a); de no estar disponible, tasa

⁵ Dada la incidencia de episodios hiperinflacionarios o de fuertes devaluaciones en la muestra, las variables de inflación y depreciación fueron transformadas de la siguiente manera. Si X denota el cambio porcentual en el índice de precios o en el tipo de cambio nominal, entonces $x = X/(1 + X)$, de modo que x se encuentra acotada entre 0 y 1.

de interés de depósitos comerciales (línea 60l).

En cuanto a las variables financieras, la proporción de M2/RIN se utiliza para controlar por los efectos de una salida repentina de capitales externos. Así, conforme esta proporción aumenta, el país sería más vulnerable a sufrir una crisis bancaria. Por otro lado, la liberalización financiera podría aumentar la fragilidad del sector bancario debido al aumento de las oportunidades para la toma de riesgos excesivos y fraude. Las variables *proxy* que capturan el progreso de la liberalización financiera son la relación entre el crédito al sector privado y el PBI y el crecimiento del crédito real. Finalmente, las circunstancias macroeconómicas adversas deben ser menos propensas a conducir a crisis bancarias en los países en que el sistema bancario es más líquido. Para captar el nivel de liquidez, se utiliza la relación de efectivo más reservas y activos bancarios. La definición de variables y fuentes de información de este grupo son:

- *M2/RIN*. Ratio de M2 (dinero más cuasidinero) a las reservas internacionales del banco central. Fuente: IFS líneas 34 (dinero) y 35 (cuasidinero) convertidas a US\$ (con el tipo de cambio de la línea AE), entre la línea 1dd (Reservas internacionales). De no estar disponible M2, DataBank.
- *Crédito/PBI*. Ratio del doméstico al sector privado entre el PBI nominal. Fuente: IFS línea 32d (crédito) y DataBank código NY.GDP.MKTP.CN (PBI nominal).
- *Liquidez*. Ratio de reservas líquidas bancarias al total de activos bancarios. Fuente: IFS línea 20 (reservas bancarias) entre la suma de las líneas 21 hasta 22f (activos bancarios).
- *Crecimiento del crédito*. Tasa de crecimiento del crédito doméstico real al sector privado. Esta variable entra a la regresión rezagada en dos años. Fuente: IFS línea 32d (crédito) y DataBank código NY.GDP.DEFL.ZS (índice de precios).

Respecto a las variables institucionales, se incorpora una variable *dummy* que indica la existencia de un seguro de depósitos explícito. Si bien Demirgüç-Kunt y Detragiache (2005b) mencionan que el signo esperado es ambiguo, en sus resultados el coeficiente presenta un signo positivo, lo que indica que un seguro de depósito explícito incrementa la probabilidad de que el sector bancario sea menos sólido por la intensificación del riesgo moral. Asimismo, se incluye el PBI per cápita como *proxy* del nivel de desarrollo del país, que presumiblemente correlaciona con la capacidad administrativa del gobierno y, en última instancia, con la efectividad de la supervisión prudencial del sistema bancario. La definición de variables y fuentes de información de este grupo son:

- *Seguro de depósitos*. Variable *dummy* construida por Demirgüç-Kunt y Detragiache (2005b).
- *PBI per cápita*. PBI real per cápita en US\$. Fuente: WDI código NY.GDP.PCAP.KD.

Finalmente, se incorpora la variable *Duración*, que está definida como el tiempo (en años) transcurrido desde la última crisis bancaria. Para ello, utilizamos las fechas de crisis bancarias desde 1900 a 1979 provistas por Reinhart y Rogoff (2014) y, para las crisis de 1980 a 1994, las encontradas en Demirgüç-Kunt y Detragiache (1998).

3 RESULTADOS

Dado que la evolución de las variables explicativas podría verse significativamente afectada luego del inicio una crisis bancaria, es necesario considerar que podría darse un efecto de doble causalidad (Kaminsky y otros, 1998). Por ejemplo, como consecuencia de la crisis bancaria, el crédito doméstico al sector privado podría reducirse, arrastrando luego al crecimiento del PBI. Asimismo, la tasa de interés real podría caer debido a una política monetaria expansiva que usualmente acompaña operaciones de rescate a instituciones financieras.

Para aliviar los efectos de la doble causalidad, Demirgüç-Kunt y Detragiache (1998) proponen dos

alternativas. La primera consiste en eliminar, para cada país, todas las observaciones después del inicio de la primera crisis bancaria registrada entre 1980 a 1994. La segunda consiste en identificar la duración de la crisis (después del primer año, ver Cuadro 1) y eliminar las observaciones correspondientes únicamente a dicho periodo. Además, dado que la probabilidad de crisis para un país que ha experimentado una crisis anteriormente puede ser diferente a la de un país que nunca ha experimentado problemas en el sector bancario, se incorpora la variable *Duración*, que es el tiempo (en años) transcurrido desde que terminó la última crisis bancaria. Para efectos comparativos, se aplica el método relogit considerando ambas alternativas.⁶

Asimismo, cada grupo de regresiones consiste en 3 especificaciones. La primera incluye únicamente a las variables macroeconómicas y al PBI per cápita; la segunda agrega las variables financieras; y la tercera incluye todas las variables explicativas. Finalmente, para efectos de las correcciones por sesgo se utiliza como probabilidad no condicional de una crisis $\pi = 6.3$ por ciento, lo cual corresponde al porcentaje de crisis bancarias respecto a las observaciones totales entre 1980 y 2000 en la muestra de países, mucho más amplia, de Reinhart y Rogoff (2014). Este número es significativamente mayor a la frecuencia muestral de $\bar{y} \approx 3$ por ciento. Luego se realiza un análisis de sensibilidad sobre la elección de π .

Los resultados de las estimaciones se reportan en el Cuadro 2 (p. 81, donde se eliminan las observaciones después de la primera crisis) y en el Cuadro 3 (p. 82, donde se eliminan las observaciones durante las crisis), usando logit y relogit. Observamos que, en conjunto, los determinantes considerados son altamente significativos, de acuerdo con el estadístico χ^2 de la regresión. Asimismo, cada grupo por separado también resulta ser significativo, lo que finalmente favorece a la especificación (3), que incluye todos los posibles determinantes. Así, en adelante, aunque nos referiremos a todos los resultados, prestaremos especial atención al modelo de regresión completo.

Significación de las variables explicativas

En todas las especificaciones, se observa que la variable *Crecimiento* está asociada significativamente con la probabilidad de crisis bancarias, lo cual confirma que los problemas en el sector real son causa de los problemas sistémicos en el sector bancario entre los años 1980s y 1990s. Además, la tasa de interés real es también significativa y presenta el signo positivo esperado en todas las especificaciones. Ello revela una potencial vulnerabilidad del sistema bancario ante fuertes desbalances macroeconómicos. El coeficiente de la inflación es siempre positivo, como se esperaba, aunque resulta ser significativo únicamente en el primer grupo de regresiones (Cuadro 2).

Por otro lado, en el Cuadro 2 se observa que el seguro de depósito explícito es una variable altamente significativa para explicar una crisis bancaria. Se reafirmaría que este mecanismo crea incentivos para que los banqueros y demás agentes tomen riesgos excesivos. No obstante, los resultados al respecto no son tan contundentes en el Cuadro 3.

Por último, en el Cuadro 3 se observa que la variable *Duración* es significativa y presenta signo

⁶ Sobre la posibilidad de utilizar técnicas de datos de panel, es decir controlar por efectos fijos, Demirgüç-Kunt y Detragiache (1998) señalan que la estimación logit con datos de panel (el denominado “logit condicional”) implica excluir del panel a todo país que no haya experimentado una crisis bancaria. Si bien ello permitiría controlar por efectos fijos, también conllevaría a descartar una gran cantidad de datos disponibles, puesto que los países que no han experimentado crisis bancarias representan más de la mitad de la muestra. Más aún, limitar el panel solo a países con crisis podría contaminar sustancialmente las estimaciones con sesgos de selección muestral. Por su parte, el ignorar efectos fijos importantes podrían ciertamente afectar la estimación puntual de β , pero sus efectos sobre la probabilidad predicha serían considerablemente menores. Es por estas razones que, al igual que Demirgüç-Kunt y Detragiache, optamos por estimar un modelo logit simple utilizando la muestra completa y prescindiendo de los efectos fijos. Dejamos la exploración con métodos de datos de panel para futuras investigaciones.

CUADRO 2. Estimaciones excluyendo observaciones después de la primera crisis

| | (1) | | (2) | | (3) | |
|----------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | logit | relogit | logit | relogit | logit | relogit |
| Constante | -3.528*** (0.423) | -3.154*** (0.443) | -3.166*** (0.537) | -2.916*** (0.573) | -3.311*** (0.565) | -3.070*** (0.611) |
| Variables Macroeconómicas | | | | | | |
| Crecimiento | -0.111*** (0.050) | -0.110*** (0.051) | -0.137*** (0.052) | -0.136*** (0.049) | -0.160*** (0.055) | -0.155*** (0.053) |
| Términos de intercambio | -0.025** (0.013) | -0.025** (0.013) | -0.016* (0.014) | -0.015* (0.013) | -0.019** (0.014) | -0.018** (0.013) |
| Depreciación | 0.343 (1.296) | 0.569 (1.147) | 0.000 (1.429) | 0.250 (1.387) | -0.436 (1.490) | -0.151 (1.422) |
| Inflación | 7.685*** (2.356) | 7.214*** (2.084) | 7.262*** (2.535) | 6.612*** (1.985) | 7.411*** (2.694) | 6.846*** (2.029) |
| Tasa de interés | 0.060*** (0.016) | 0.056*** (0.014) | 0.062*** (0.018) | 0.064*** (0.015) | 0.064*** (0.019) | 0.065*** (0.015) |
| Variables Financieras | | | | | | |
| M2/RIN | | | 0.000 (0.000) | 0.003*** (0.000) | 0.000 (0.000) | 0.003*** (0.000) |
| Crédito/PBI | | | 0.001 (0.001) | 0.001* (0.001) | 0.001* (0.001) | 0.001** (0.001) |
| Liquidez | | | -2.062 (2.504) | -1.602* (1.856) | -2.138* (2.317) | -1.500 (2.084) |
| Crecimiento del crédito | | | 0.782 (1.156) | 1.124 (1.360) | 1.040* (1.207) | 1.368** (1.058) |
| Otras variables | | | | | | |
| PBI per cápita | -0.012 (0.019) | -0.013 (0.022) | -0.022* (0.022) | -0.019 (0.026) | -0.050*** (0.024) | -0.045*** (0.019) |
| Seguro de depósitos | | | | | 1.387*** (0.508) | 1.307*** (0.516) |
| Número de crisis | 29 | | 28 | | 28 | |
| Número de países | 60 | | 59 | | 59 | |
| Número de observaciones | 587 | | 570 | | 570 | |
| Estadístico χ^2 | 33.87*** | | 31.42*** | | 38.65*** | |
| Crisis predichas | 18 | 21 | 17 | 24 | 19 | 23 |
| Crisis predichas (%) | 62 | 72 | 61 | 86 | 68 | 82 |

NOTAS: Al incluir variables financieras (es decir, al pasar de la especificación 1 a la especificación 2) se pierde el caso de Tanzania, que reporta una crisis bancaria en 1988, por falta de información. Se excluyen, además, las crisis de México y Turquía en 1994 por tratarse de eventos posteriores a la ocurrencia de una primera crisis bancaria en los 80s. De hecho, este criterio hace que Turquía sea completamente eliminada de esta muestra de estimación. Un caso similar ocurre para Perú, quien reporta una crisis bancaria en 1983 que no es incluida por no contar con información completa de los determinantes.

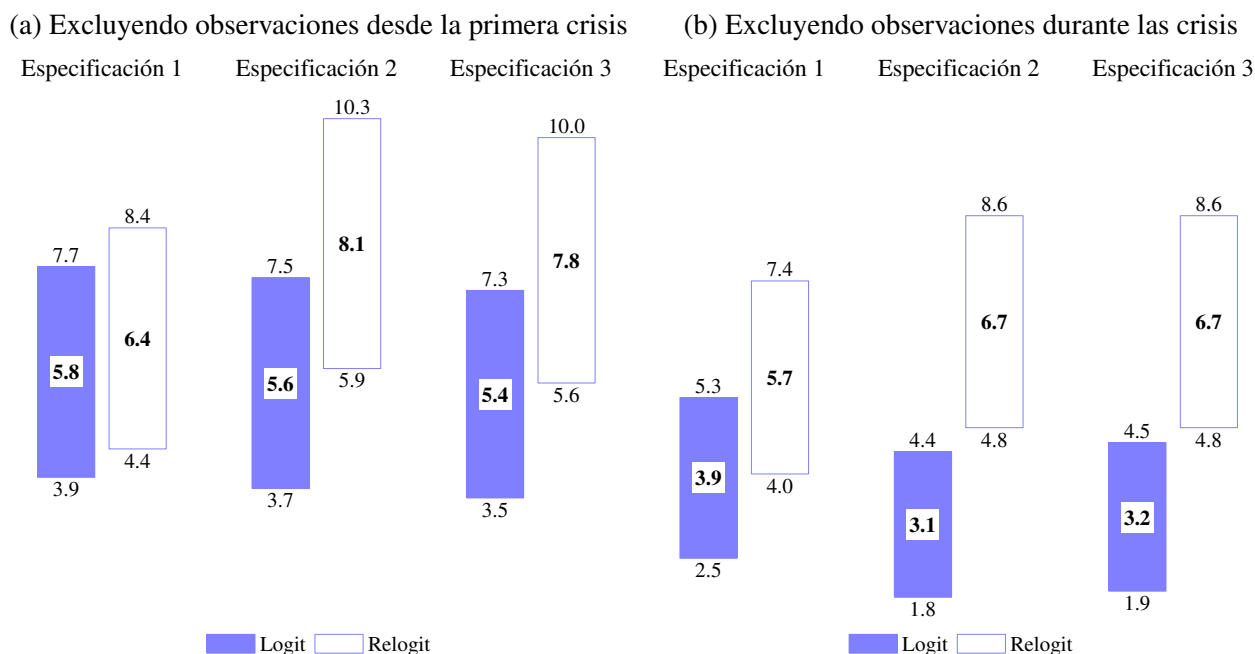
Errores estándares en paréntesis. * [**] {***} denota un nivel de 10% [5%] {1%} de significación estadística.

CUADRO 3. Estimaciones excluyendo observaciones durante las crisis

| | (1) | | (2) | | (3) | |
|----------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | logit | relogit | logit | relogit | logit | relogit |
| Constante | -3.009*** (0.323) | 2.579*** (0.312) | -2.553*** (0.487) | -2.306*** (0.460) | -2.666*** (0.518) | -2.322*** (0.510) |
| Variables Macroeconómicas | | | | | | |
| Crecimiento | -0.122*** (0.046) | -0.119*** (0.043) | -0.166*** (0.050) | -0.155*** (0.046) | -0.171*** (0.051) | -0.162*** (0.047) |
| Términos de intercambio | -0.015* (0.012) | -0.015* (0.013) | -0.007 (0.013) | -0.006 (0.013) | -0.007 (0.013) | -0.007 (0.014) |
| Depreciación | 1.292* (1.170) | 1.396* (1.101) | 1.098* (1.179) | 1.059* (1.159) | 1.076* (1.178) | 1.038* (1.138) |
| Inflación | 0.841 (1.543) | 0.774 (1.340) | 1.584* (1.866) | 1.830* (1.475) | 1.467 (1.871) | 1.578* (1.504) |
| Tasa de interés | 0.019*** (0.007) | 0.017*** (0.008) | 0.023*** (0.008) | 0.020** (0.011) | 0.023*** (0.008) | 0.021** (0.011) |
| Variables Financieras | | | | | | |
| M2/RIN | | | -0.001 (0.003) | 0.004*** (0.001) | 0.001 (0.003) | 0.002*** (0.001) |
| Crédito/PBI | | | 0.000 (0.001) | 0.000 (0.001) | 0.000 (0.001) | 0.000 (0.001) |
| Liquidez | | | -4.281** (2.512) | -3.318** (2.201) | -3.988** (2.552) | -3.180** (2.307) |
| Crecimiento del crédito | | | 2.025** (1.144) | 2.241** (1.190) | 2.042** (1.155) | 2.221** (1.186) |
| Otras variables | | | | | | |
| PBI per cápita | -0.067*** (0.024) | -0.065*** (0.021) | -0.091*** (0.028) | -0.082*** (0.024) | -0.091*** (0.028) | -0.083*** (0.023) |
| Seguro de depósitos | | | | | 0.419* (0.478) | 0.394 (0.521) |
| Duración | 0.053*** (0.012) | 0.052*** (0.009) | 0.062*** (0.013) | 0.055*** (0.010) | 0.057*** (0.014) | 0.052*** (0.012) |
| Número de crisis | 31 | | 30 | | 30 | |
| Número de países | 62 | | 61 | | 61 | |
| Número de observaciones | 690 | | 668 | | 668 | |
| Estadístico χ^2 | 39.75 | | 46.70 | | 47.46 | |
| Crisis predichas | 13 | 20 | 16 | 24 | 19 | 23 |
| Crisis predichas (%) | 42 | 65 | 53 | 80 | 63 | 76 |

NOTAS: Al incluir variables financieras (es decir, al pasar de la especificación 1 a la especificación 2) se pierde el caso de Tanzania, que reporta una crisis bancaria en 1988, por falta de información.

Errores estándares en paréntesis. * [**] [***] denota un nivel de 10% [5%] [1%] de significación estadística.

GRÁFICO 1. Probabilidad no condicional de crisis bancaria según método de estimación

NOTA: Se presenta la estimación de $\Pr(y = 1) = \mathbb{E}_x(\Pr(y = 1 | x))$, en porcentajes, junto con su intervalo de 95% de confianza.

positivo, lo cual indica que un país que ha experimentado una crisis bancaria en el pasado se encontraría más propenso a tener problemas financieros en el futuro.

En general, los resultados de las estimaciones relogit son similares a las estimaciones logit en cuanto a magnitud y signos de coeficientes de las variables explicativas. En el caso de M2/RIN se observa que la significación es mayor en las estimaciones relogit. Esto indicaría que un ambiente macroeconómico débil no es el único factor detrás de problemas en el sector financiero, sino que es necesario prestar atención también a la vulnerabilidad del sector financiero ante factores externos, como la salida repentina de capitales.

Probabilidades de crisis

Una de las diferencias más saltantes entre las estimaciones logit y relogit es el porcentaje de crisis correctamente predichas. Mientras que entre el 40 y 70 por ciento de las crisis son predichas correctamente con el método logit, este número asciende a entre 65 y 85 por ciento para relogit.

Es interesante notar que el desempeño predictivo de ambos métodos mejora conforme se introducen regresores adicionales al modelo probabilístico, y que las mayores mejoras se registran en el modelo logit. Ello sugiere que las especificaciones iniciales, que excluyen variables explicativas relevantes, estarían fuertemente sesgadas. El comportamiento más estable de relogit a lo largo de especificaciones es una simple manifestación de que éste es un método que inherentemente corrige por sesgos. Así, el relogit no solo puede ser útil en el contexto de eventos infrecuentes, sino en contextos más generales en donde se sospecha de la omisión, deliberada o no, de algún determinante importante.

Por otro lado, a fin de continuar con la comparación entre logit y relogit, se determinó la probabilidad no condicional de crisis bancaria en cada especificación. Para ello, se obtuvo primero la probabilidad de crisis bancaria de cada observación y luego el promedio de estos valores junto con sus intervalos de confianza. El Gráfico 1 muestra los resultados obtenidos para un nivel de significación del 5 por ciento.

Se aprecia claramente cómo *relogit* ajusta las predicciones del modelo probabilístico hacia arriba. Este hecho es más elocuente en las especificaciones mejor controladas, que incluyen un mayor número de determinantes. Se aprecia, de hecho, que *relogit* no rechaza regiones de probabilidad altas como sí lo hace *logit*.

Alertas tempranas

Las medidas de capacidad predictiva de los Cuadros 2 y 3 indican si el modelo probabilístico predijo la ocurrencia de una crisis bancaria *en el año en que la crisis ocurrió*. Sin embargo, dado que en general los determinantes considerados son persistentes, es posible que estos modelos provean de señales anticipadas sobre la ocurrencia de la crisis en años previos al inicio del episodio. En otras palabras, los elementos que contribuyen a aumentar la fragilidad del sistema bancario pueden manifestarse uno o más años antes de que los problemas se pongan en manifiesto. En este sentido, el conteo de crisis correctamente predichas anterior podría estar subestimando la capacidad predictiva de cada método.

En los Cuadros 4 (p. 85) y 5 (p. 86) se muestra el conteo de crisis correctamente predichas tanto en el año de la crisis como en los dos años previos. Se muestra, asimismo, si es que alguna crisis nunca fue predicha. Para ahorrar espacio, únicamente se presentan las predicciones del modelo con mayor número de regresores, es decir la especificación 3. En ambos Cuadros, los resultados de la columna $h = 0$ son los correspondientes a los reportados en los Cuadros 2 y 3.

En el Cuadro 4 se aprecia que, del total de 28 crisis utilizadas en la estimación, el modelo *logit* predice 22 por lo menos una vez desde dos años antes de iniciada la crisis hasta el momento en que ésta ocurre, mientras que los 6 episodios restantes no son predichos en absoluto. En el caso del modelo *relogit*, únicamente una crisis no es predicha, mientras que las 27 restantes sí son anticipadas. Más aún, todo evento anticipado con *logit* es también anticipado con *relogit*, y el método corregido provee señales más categóricas ya que los eventos son anticipados por lo menos en dos años consecutivos, mientras que el caso del *logit* se registran algunas señales aisladas (por ejemplo, en las crisis de Portugal en 1986 o de Nepal en 1988). El caso no predicho por *relogit*, Suecia en 1990, tampoco es predicho por *logit*.

El Cuadro 5 presenta resultados similares y confirma la mayor capacidad predictiva de *relogit*. Del total de 30 crisis utilizadas en la estimación, *logit* predice 24 por lo menos una vez desde dos años antes de iniciada la crisis, y *relogit* 28. Las dos crisis no anticipadas por *relogit*, Nigeria en 1991 y Guyana en 1993, tampoco son anticipadas por *logit*. Nuevamente, se tiene que *relogit* no registra señales aisladas. Todo ello apunta a que *relogit* podría ser una herramienta útil en el diseño de sistemas de alertas tempranas.

Efectos marginales

En todo análisis de regresión, el interés se centra en determinar cómo se ve afectada la variable dependiente cuando se altera el valor de las variables explicativas. Dada la naturaleza no lineal de los modelos probabilísticos, los efectos marginales de interés (esto es, los cambios en las probabilidades predichas ante cambios en las variables explicativas) no son iguales a los coeficientes estimados en los Cuadros 2 y 3, como sí ocurriría con modelos lineales, sino que involucran además estimaciones de las probabilidades de crisis condicionales a los valores de las variables explicativas. Así, aunque las estimaciones puntuales de los coeficientes *logit* y *relogit* puedan parecer similares, es posible que los efectos marginales derivados de cada método sean distintos, toda vez que el sesgo corregido por *relogit* sea empíricamente relevante.

El Cuadro 6 (p. 87) presenta algunos cálculos ilustrativos donde, utilizando la especificación completa del segundo grupo de regresiones, se presentan los efectos marginales en torno a episodios de crisis bancarias de tres variables que resultaron ser estadísticamente significativas (crecimiento, tasa de interés

CUADRO 4. Predicciones del modelo que excluye observaciones desde la primera crisis

| | | relogit | | | logit | | | | |
|-------|-------------------|------------|----------|----------|----------|------------|---------|---------|---------|
| | | No predice | $h = 0$ | $h = 1$ | $h = 2$ | No predice | $h = 0$ | $h = 1$ | $h = 2$ |
| 1981 | Chile | | X | X | | | X | X | |
| 1981 | Estados Unidos | | X | X | | | X | X | |
| 1981 | Filipinas | | X | <u>X</u> | | | X | | |
| 1981 | Uruguay | | X | X | | | X | X | |
| 1982 | Colombia | | X | X | X | | X | X | X |
| 1982 | México | | X | <u>X</u> | | | X | | |
| 1983 | Israel | | X | X | | | X | X | |
| 1983 | Senegal | | X | X | X | | X | X | X |
| 1983 | Tailandia | | | <u>X</u> | <u>X</u> | X | | | |
| 1985 | Malasia | | X | <u>X</u> | | | X | | |
| 1985 | Sudáfrica | | X | X | X | | X | X | X |
| 1986 | Portugal | | | <u>X</u> | X | | | | X |
| 1987 | Mali | | <u>X</u> | <u>X</u> | <u>X</u> | X | | | |
| 1987 | Noruega | | X | <u>X</u> | | X | | | |
| 1988 | Nepal | | | X | <u>X</u> | | | X | |
| 1989 | Jordania | | X | <u>X</u> | <u>X</u> | | X | | |
| 1989 | Papúa Nueva Guina | | X | <u>X</u> | | | X | | |
| 1989 | Sri Lanka | | <u>X</u> | <u>X</u> | <u>X</u> | X | | | |
| 1990 | Italia | | X | X | X | | X | X | X |
| 1990 | Suecia | X | | | | X | | | |
| 1991 | Finlandia | | X | X | | | X | X | |
| 1991 | India | | X | X | X | | X | X | X |
| 1991 | Nigeria | | X | X | X | | X | X | X |
| 1992 | Indonesia | | X | X | | | X | X | |
| 1992 | Japón | | X | <u>X</u> | | X | | | |
| 1993 | Guyana | | | X | <u>X</u> | | | X | |
| 1993 | Kenia | | X | X | X | | X | X | X |
| 1993 | Venezuela | | X | X | X | | X | X | X |
| Total | | 1 | 23 | 27 | 15 | 6 | 19 | 16 | 9 |

NOTAS: “X” denota que la crisis fue predicha h años antes del momento de ocurrencia. “X” indica los casos en donde un método señala la ocurrencia de una crisis, y el otro no. Se utiliza la especificación 3 del Cuadro 2.

y crecimiento del crédito). El Cuadro muestra estimaciones de $\Pr(y = 1 | z_1, \mathbf{x}_{-z}) - \Pr(y = 1 | z_0, \mathbf{x}_{-z})$ donde un regresor z cambia de valor mientras que el resto de variables, recolectadas en el vector \mathbf{x}_{-z} , se mantienen constantes en sus promedios muestrales. La variable z cambia de un valor dado inicialmente por el promedio en los años 3, 4 y 5 antes de la ocurrencia de la crisis bancaria (z_0) a un valor final dado por el promedio 0, 1 y 2 años antes de este episodio (z_1).

Es interesante notar que los efectos marginales de relogit tienden a ser mayores a los de logit, sobre todo cuando éstos son positivos. El hecho de que las estimaciones detrás de relogit sean más sensibles a los cambios en z en periodos cercanos a la ocurrencia de una crisis bancaria sugiere que es precisamente esta mayor sensibilidad, que es a su vez un resultado de la corrección del sesgo inherente en el método, la fuente de la mayor capacidad predictiva discutida en los Cuadros 4 y 5. Toda vez que z cambie de manera relativamente predecible en anticipación a una crisis bancaria, relogit responderá más energicamente a esta

CUADRO 5. Predicciones del modelo que excluye observaciones durante las crisis

| | | relogit | | | logit | | | | |
|-------|--------------------|------------|----------|----------|----------|------------|---------|---------|---------|
| | | No predice | $h = 0$ | $h = 1$ | $h = 2$ | No predice | $h = 0$ | $h = 1$ | $h = 2$ |
| 1981 | Chile | | X | X | | X | X | | |
| 1981 | Estados Unidos | | X | X | | X | X | | |
| 1981 | Filipinas | | X | <u>X</u> | | X | | | |
| 1981 | Uruguay | | X | X | | X | X | | |
| 1982 | Colombia | | <u>X</u> | <u>X</u> | X | | | | |
| 1982 | México | | X | X | | X | X | | |
| 1983 | Israel | | X | <u>X</u> | | X | | | |
| 1983 | Senegal | | X | X | X | X | X | X | |
| 1983 | Tailandia | | | | <u>X</u> | X | | | |
| 1985 | Malaysia | | X | <u>X</u> | | X | | | |
| 1985 | Sudáfrica | | X | X | X | X | X | X | |
| 1986 | Portugal | | X | X | X | X | X | X | |
| 1987 | Mali | | <u>X</u> | <u>X</u> | | X | | | |
| 1987 | Noruega | | <u>X</u> | <u>X</u> | | X | | | |
| 1988 | Nepal | | | X | <u>X</u> | | X | | |
| 1989 | Jordania | | X | <u>X</u> | <u>X</u> | X | | | |
| 1989 | Papúa Nueva Guinea | | X | <u>X</u> | <u>X</u> | X | | | |
| 1989 | Sri Lanka | | <u>X</u> | <u>X</u> | X | | | X | |
| 1990 | Italia | | X | X | X | X | X | X | |
| 1990 | Suecia | | X | X | X | X | X | X | |
| 1991 | Finlandia | | X | X | X | X | X | X | |
| 1991 | India | | X | X | X | X | X | X | |
| 1991 | Nigeria | X | | | | X | | | |
| 1992 | Indonesia | | X | <u>X</u> | | X | | | |
| 1992 | Japón | | X | X | X | X | X | X | |
| 1993 | Guyana | X | | | | X | | | |
| 1993 | Kenia | | X | X | X | X | X | X | |
| 1993 | Venezuela | | X | <u>X</u> | | X | | | |
| 1994 | México | | X | X | X | X | X | X | |
| 1994 | Turquía | | X | <u>X</u> | | X | | | |
| Total | | 2 | 26 | 27 | 15 | 6 | 22 | 15 | 11 |

NOTAS: "X" denota que la crisis fue predicha h años antes del momento de ocurrencia. "X" indica los casos en donde un método señala la ocurrencia de una crisis, y el otro no. Se utiliza la especificación 3 del Cuadro 3.

variación, enviando así una señal más fuerte sobre el posible desarrollo de una crisis. Es bueno recordar que, de los eventos reportados en el Cuadro 6, logit no fue capaz de predecir las crisis de Noruega en 1987 ni Japón en 1992.

Calibración de π

Como se ha discutido, los sesgos por la infrecuencia del evento se originan cuando existe una fuerte discrepancia entre la frecuencia muestral \bar{y} y la verdadera probabilidad no condicional de ocurrencia del evento π . Al reponderar las observaciones involucradas en la estimación, relogit trata de cerrar esta brecha y así corregir el sesgo. La elección de π es una elección subjetiva del investigador, basada en su

CUADRO 6. Efectos marginales en torno a crisis selectas

| Variable | Valor inicial | Valor final | Cambio | Efecto marginal | |
|-------------------------|---------------|-------------|--------|-----------------|-------|
| | (z_0) | (z_1) | | Relogit | Logit |
| Noruega 1987 | | | | | |
| Crecimiento | 3.30 | 3.72 | 0.43 | -0.11 | -0.10 |
| Tasa de interés | -2.64 | 7.13 | 9.77 | 0.70 | 0.38 |
| Crecimiento del crédito | 0.02 | 0.17 | 0.15 | 1.02 | 0.37 |
| Japón 1992 | | | | | |
| Crecimiento | 5.53 | 3.28 | -2.25 | 1.64 | 0.65 |
| Tasa de interés | 2.26 | 2.47 | 0.21 | 0.05 | 0.01 |
| Crecimiento del crédito | 0.10 | 0.09 | -0.01 | -0.06 | -0.05 |
| México 1994 | | | | | |
| Crecimiento | 4.50 | 3.33 | -1.16 | 2.79 | 2.61 |
| Tasa de interés | 1.28 | 5.14 | 3.87 | 1.13 | 1.10 |
| Crecimiento del crédito | 0.11 | 0.27 | 0.16 | 4.76 | 3.17 |

NOTAS: Los efectos marginales son $\Pr(y = 1 | z_1, \mathbf{x}_{-z}) - \Pr(y = 1 | z_0, \mathbf{x}_{-z})$ donde el regresor z cambia de valor, de z_0 a z_1 , mientras que el resto (\mathbf{x}_{-z}) se mantienen constantes en sus promedios muestrales. El valor de z_1 es el promedio en la ventana de 3 años que incluye el año de la crisis y los dos años anteriores (Noruega: 1985-1987, Japón: 1990-1992, México: 1992-1994), mientras que el valor de z_0 es el promedio en la ventana de 3 años que abarca desde 5 hasta 3 años antes de la crisis (Noruega: 1982-1984, Japón: 1987-1989, México: 1989-1991). Se utiliza la especificación 3 del Cuadro 3.

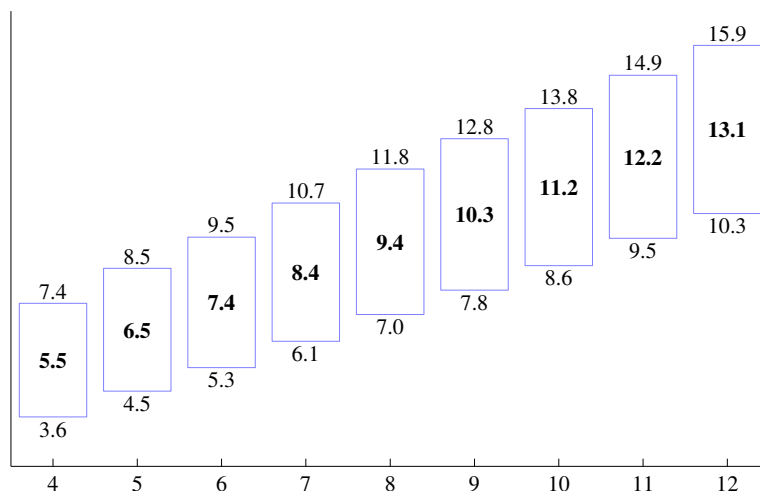
conocimiento previo u otra fuente de información ajena a la muestra utilizada en la estimación, cuya mejor inferencia sobre π es el estimador sesgado \bar{y} .

En nuestro análisis hemos utilizado una muestra similar a la utilizada en [Demirgüç-Kunt y Detragiache \(1998\)](#), tanto en el número de países incluidos (62 países) como en el periodo de análisis (15 años, de 1980 a 1994). No obstante, se entiende que esta muestra subrepresenta la ocurrencia de crisis bancarias. Por ello, se recurrió a la base de datos de [Reinhart y Rogoff \(2011\)](#), que incluye un periodo más largo (21 años, entre 1980 y 2000) y un número mayor de países para la calibración de π . Ello dio como resultado $\pi = 6.3$ por ciento, lo cual corresponde al porcentaje de crisis bancarias en esta base de datos, que es significativamente mayor al $\bar{y} \approx 3$ en la muestra de estimación. En otras palabras, implícitamente asumimos que nuestra muestra captura una fracción del universo de crisis bancarias correctamente documentado por [Reinhart y Rogoff](#), por ejemplo por no contar con suficiente información histórica sobre los determinantes \mathbf{x} .

Si bien consideramos que esta manera de calibrar π es poco controversial, podría argumentarse – por lo menos conceptualmente – que la base de datos de [Reinhart y Rogoff \(2011\)](#) es, a su vez, la muestra de un universo aún mayor, y así sucesivamente. La pregunta de fondo, no obstante, es cómo la inferencia relógit es afectada por la elección subjetiva de π . El Gráfico 2 (p. 88) provee una primera respuesta. Este Gráfico muestra la estimación de la probabilidad no condicional de crisis, similar a las presentadas en el Gráfico 1, obtenida con la tercera especificación del Cuadro 2, para distintos valores de π . Es interesante notar que los intervalos de confianza resultantes, como era de esperarse, contienen a π pero no contienen necesariamente a \bar{y} , especialmente cuando π es mucho mayor que \bar{y} .

4 COMENTARIOS FINALES

Así como muchos eventos de primera importancia en el desempeño de las economías, las crisis bancarias son eventos de estudio inusuales: diversas bases de datos de trabajos empíricos relacionados apuntan a

GRÁFICO 2. Sensibilidad ante calibración de π 

NOTA: Se presenta la estimación de $\Pr(y = 1) = \mathbb{E}_x(\Pr(y = 1 | x))$, en porcentajes, junto con su intervalo de 95% de confianza, para distintos valores de π (eje horizontal, también en porcentajes), utilizando la especificación 3 del Cuadro 2.

que éstas ocurren entre 3 y 5 por ciento de los periodos considerados. Esta baja frecuencia de ocurrencia sugiere que las estimaciones logit consideradas en la literatura podrían estar significativamente sesgadas, en particular subestimando la probabilidad predicha de crisis.

Nuestros resultados indican que las estimaciones con relogit son similares a las estimaciones logit en cuanto a magnitud y signos de coeficientes de las variables explicativas. No obstante, se verifica que la probabilidad no condicional de crisis bancaria se encuentra subestimada al emplear logit. La probabilidad no condicional de crisis bancaria se incrementa cerca de 4 por ciento (en promedio) al utilizar relogit. En todas las especificaciones utilizadas (empleando variables macroeconómicas, financieras e institucionales) relogit no rechaza regiones altas de probabilidad, a diferencia del logit convencional.

Además, el modelo estimado con relogit presenta una mejor capacidad de predicción de crisis bancarias que el estimado con logit. En un primer grupo de estimaciones, de un total de 28 crisis bancarias, relogit no predice únicamente una crisis mientras que logit no predice seis; en un segundo grupo de resultados, ahora de un total de 30 crisis bancarias, relogit no predice dos crisis mientras que logit no predice, nuevamente, seis. Más aún, relogit tiende a predecir las crisis uno o dos años *antes* de que éstas sean registradas, lo que ilustra el potencial del método en el diseño de sistemas de alerta tempranas.

Relogit es un método que ha sido desarrollado pensando en modelos de regresión con características de cortes transversales. Consideramos que la adecuación de este método a contextos relevantes para el estudio de fenómenos macroeconómicos, especialmente series temporales o paneles de datos, abre una interesante agenda de investigación futura.

REFERENCIAS

- Buhag, H. (2005), "Dangerous dyads revisited: Democracies may not be that peaceful at all", *Conflict Management and Peace Science*, 22(2), 95-111
- Cheng, Z., I. Nielsen y R. Smyth (2013), "Determinants of wage arrears and their implications for the socioeconomic wellbeing of China's migrant workers: Evidence from Guangdong province", Monash University, Documento de Trabajo 50/12.

- Cordeiro, G. y P. McCullagh (1991), "Bias correction in generalized linear models", *Journal of the Royal Statistical Society*, 53(3), 629-643.
- Demirgüç-Kunt, A. y E. Detragiache (1998), "The determinants of banking crises", *IMF Staff Papers*, 45(1), 81-109.
- Demirgüç-Kunt, A. y E. Detragiache (2002), "Does deposit insurance increase banking system stability?", *Journal of Monetary Economics*, 49(7), 1373-1406.
- Demirgüç-Kunt, A. y E. Detragiache (2005a), "Cross-country empirical studies of systemic bank distress: A survey", *National Institute Economic Review*, 192(1), 68-83.
- Demirgüç-Kunt, A. y E. Detragiache (2005b), "Deposit insurance around the world: A comprehensive database", World Bank Policy Research Working Paper 3628.
- Demirgüç-Kunt, A. y E. Detragiache (2010), "Bank capital: Lessons from financial crisis", *Journal of Money, Banking and Credit*, 45(6), 1147-1164.
- Domac, I. y M. Martínez (2000), "Banking crisis and exchange rate regimes: Is there a link?", *Journal of International Economics*, 61(1), 41-72
- Fondo Monetario Internacional (2014), *World Economic Outlook Database*.
- Frydl, E. J. (1999), "The length and cost of banking crises", IMF Working Paper 99/30.
- Fuhrmann, M. y S. Kreps (2010), "Targeting nuclear programs in war and peace: A quantitative empirical analysis, 1941-2000", *Journal of Conflict Resolution*, 54(6), 831-859.
- Hausmann, R., L. Pritchett y D. Rodrik (2005), "Growth accelerations", *Journal of Economic Growth*, 10(4), 303-329.
- Hennig-Thurau, T., V. Henning y H. Sattlersource (2007), "Consumer file sharing of motion pictures", *Journal of Marketing*, 71(10), 1-18.
- Hurwitz, M. y J. Stefko (2004), "Acclimation and attitudes: 'Newcomer' justices and precedent conformance on the Supreme Court", *Political Research Quarterly*, 57(1), 121-129.
- Kaminsky, G., S. Lizondo y C. M. Reinhart (1998), "Leading indicators of currency crises", *IMF Staff Papers*, 45(1), 1-48.
- King, G. y L. Zeng (2001a), "Logistic Regression in rare events data", *Political Analysis*, 9(2), 137-163.
- King, G. y L. Zeng (2001b), "Explaining rare events in international relations", *International Organization*, 55 (3), 693-715.
- King, G. y L. Zeng (2002) "Estimating risk and rate levels, ratios and differences in case-control studies", *Statistics in Medicine*, 21, 1409-1427.
- Laeven, L. (2011), "Banking crises: A review", *Annual Review of Financial Economics*, 3(1), 17-40.
- Martin, L. (2005), "The president and international commitments: Treaties as signaling devices", *Presidential Studies Quarterly*, 35(3), 440-465.
- Reinhart, C. M. y K. Rogoff (2011), "From financial crash to debt crisis", *American Economic Review*, 101(5), 1676-1706.
- Reinhart, C. M. y K. Rogoff (2014), *This Time Is Different: Eight Centuries of Financial Folly*, Princeton University Press.

Röller, L., R. Siebert, y M. Tombak (2007), “Why firms form (or do not form) RJVS?”, *Economic Journal*, 117(522), 1122-1144.

Tomz, M., G. King y L. Zeng (2003), “ReLogit: Rare events logistic regression”, *Journal of Statistical Software*, 8(2), 1-27.

Von Hagen, J. y T. K. Ho (2007), “Money market pressure and the determinants of banking crises”, *Journal of Money, Credit and Banking*, 39(5), 1037-1066.