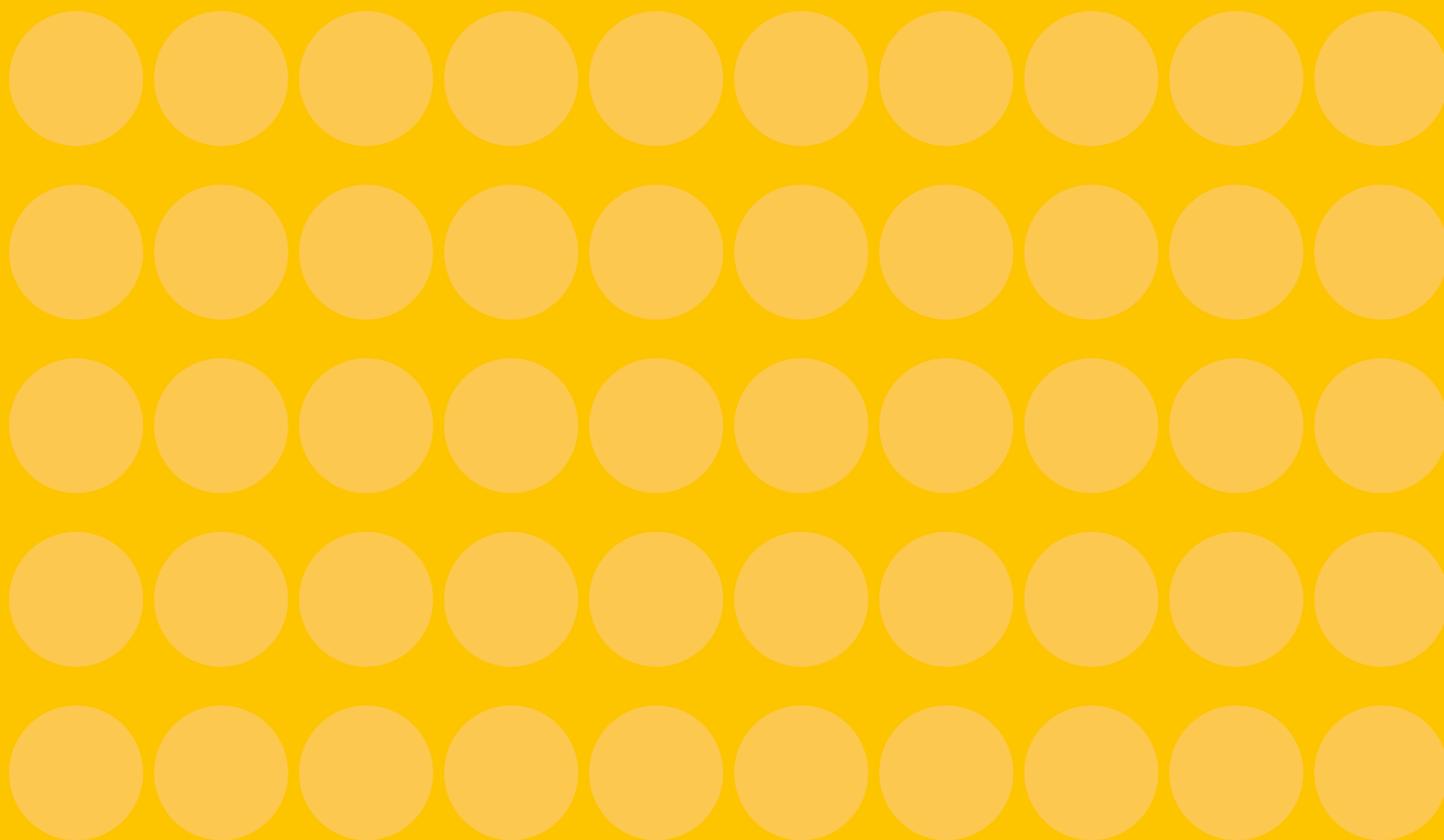


BOLETÍN

Volumen LXII

Número 3, julio-septiembre de 2016



Centro de Estudios Monetarios Latinoamericanos

Boletín

Volumen LXII, número 3,
julio-septiembre de 2016

ASAMBLEA

Bancos centrales asociados (*vox et votum*)
Miembros colaboradores (*vox*)

JUNTA DE GOBIERNO 2015-2017

Presidente

Banco Central de Chile

Miembros

Central Bank of The Bahamas
Banco Central de Bolivia
Banco Central de Costa Rica
Banco de México
Banco Central del Paraguay
Banco Central de Venezuela
Banco de España
Board of Governors
of the Federal Reserve System

COMITÉ EDITORIAL

Fernando Tenjo Galarza
Director general
Dalmir Sergio Louzada
Subdirector general
Fernando Sánchez Cuadros
Reuniones Técnicas de Banca Central
Ana Laura Sibaja Jiménez
Servicios de Información
María José Roa
Investigación Económica

ÍNDICE

215 Minería de textos para bancos centrales

David Bholat
Stephen Hansen
Pedro Santos
Cheryl Schonhardt-Bailey

245 Iniciativa del G20 sobre deficiencias de los datos II: frente al desafío de política

Robert Heath
Evrin Bese Goksu

273 Modelos del riesgo de crédito

Somnath Chatterjee

Boletín es una publicación del Centro de Estudios Monetarios Latinoamericanos, Durango núm. 54, colonia Roma, Ciudad de México, 06700. ISSN: 0186-7229. *Los trabajos firmados son responsabilidad exclusiva de los autores y no coinciden necesariamente con el criterio del CEMLA.*

MINERÍA DE TEXTOS PARA BANCOS CENTRALES

David Bholat
Stephen Hansen
Pedro Santos
Cheryl Schonhardt-Bailey

INTRODUCCIÓN

La minería de textos (también llamada *procesamiento de lenguaje natural*¹ o *lingüística computacional*) es un término amplio para una gama de herramientas de cómputo y técnicas estadísticas que cuantifican texto.² La minería de textos se parece a la lectura porque ambas actividades implican extraer significado de series de letras. Sin embargo, el análisis computacional y estadístico del texto difiere de la lectura en dos aspectos importantes. Primero, los enfoques permitidos por la computadora pueden procesar y resumir mucho más texto en comparación con el tiempo que una persona tiene para leer. Y segundo, tales técnicas pueden ser capaces de extraer significado de texto que escapa a los

¹ El procesamiento de lenguaje natural es el procesamiento y análisis por computadora de los lenguajes naturales en el ser humano, en contraste con los lenguajes de programación, como Java.

² También hay métodos asistidos por computadora para el análisis cualitativo de textos. Sin embargo, estos se salen del ámbito de aplicación de este manual. Sin embargo, en este vínculo se encuentra una introducción y contraste de algunas de las herramientas cualitativas de la minería de textos: <<http://www.surrey.ac.uk/sociology/research/researchcentres/caqdas/support/choo>>. Ver también Upshall (2014).

Traduce y publica el CEMLA con la debida autorización el Handbook, núm. 33, del Centre for Central Banking Studies, Banco de Inglaterra, 2015. Agradecemos a Ayeh Bandeh-Ahmadi, Aude Bicquelet, David Bradnum, Peter Eckley, Jo Gill, David Gregory, Sujit Kapadia, Tom Khabaza, Christopher Lovell, Rickard Nyman, Paul Ormerod, Paul Robinson, Robert Smith, David Tuckett, Iulian Udrea y Derek Vallès por sus comentarios. Las opiniones expresadas en este manual son las del autor y no necesariamente las de nuestros empleadores. <david.bholat@bankofengland.co.uk>, <stephen.hansen@upf.edu>, <pedro.santos@bankofengland.co.uk> y <c.m.schonhardt-bailey@lse.ac.uk>.

lectores humanos, quienes pasan por alto ciertos patrones porque no se apegan a sus creencias y expectativas previas.

Aunque son muchas sus aplicaciones en campos como las ciencias políticas y la mercadotecnia, la minería de textos históricamente ha sido menos utilizada como técnica en la economía. En particular, este es el caso de la investigación que realizan los bancos centrales internamente. Esto ha sido así por un par de razones al menos. La primera es que tal vez no sea obvio que el texto puede describirse y analizarse como datos cuantitativos.³ Por tal motivo, los bancos centrales no están familiarizados con las herramientas y técnicas estadísticas que hacen posible lo anterior. La segunda es que, incluso si los banqueros centrales han oído hablar de la minería de textos, ya tienen acceso a otros datos cuantitativos a su disposición. Tal vez piensan que el costo de oportunidad y otros tipos de costos asociados con la transformación de textos en datos cuantitativos, así como tener que aprender nuevas herramientas y técnicas para analizarlos, superan los beneficios esperados.

Sin embargo, la minería de textos bien pudiera valer la inversión de los bancos centrales porque vuelve manejable una variedad de fuentes de datos importantes para evaluar la estabilidad monetaria y financiera, y que no pueden analizarse cuantitativamente de otras maneras. Los textos fundamentales para los bancos centrales son, entre otros, artículos de prensa, contratos financieros, redes sociales, inteligencia de supervisión y de mercado, así como informes escritos de distinta índole. Con las técnicas de minería de textos, podemos analizar un documento o una serie de documentos (un corpus). Un documento puede ser cierto discurso de un miembro del Comité de Política Monetaria

³ Al texto a veces se le denomina *datos no estructurados*, en contraste con los datos estructurados (los números). Sin embargo, calificar a un texto como no estructurado es un poco engañoso. Un texto sí tiene estructura, siendo la gramática la más obvia, pero también patrones estructurales de distinto tipo que se extraen mediante las técnicas de minería de textos.

(CPM) del Banco de Inglaterra (en lo sucesivo “el Banco”), una nota del personal o un informe de campo presentado por un agente.⁴ El corpus correspondiente sería la totalidad de discursos de miembros del CPM, de notas del personal y de informes de campo.

Aunque los bancos centrales no han hecho mucho uso intencional de las técnicas de minería de textos, los banqueros centrales sí cosechan todos los días los beneficios de las aplicaciones de la minería de textos. Considérese, por ejemplo, la frecuencia con que los banqueros centrales (o cualquiera) usan Google para buscar información o utilizan el corrector de ortografía antes de publicar un documento. Considérense también los cortafuegos para aislar a los bancos centrales de los ataques cibernéticos, o la funcionalidad de búsqueda en las bases de datos de citas para recuperar las publicaciones académicas sobre cualquier tema. En este y otros casos, las técnicas de minería de textos funcionan en un segundo plano para ayudar a los bancos centrales a realizar su trabajo más eficientemente.

Otra finalidad de este documento es, entonces, demostrar el valor que los bancos centrales podrían obtener de una aplicación más concienzuda de las técnicas de minería de textos y explicar algunas de ellas utilizando ejemplos relevantes para esas instituciones. Este documento consta de dos partes principales. En la primera se explica cómo puede aplicarse la minería de textos a la investigación y la formulación de políticas que realiza el banco central, con base en ejemplos tomados de publicaciones. En la segunda parte del documento se proporcionan los primeros pasos para la minería de textos. Empezaremos con una explicación de cómo preparar los textos para su análisis. Luego comentamos varias técnicas de minería de textos, empezando por algunos métodos intuitivos, como las técnicas booleanas y las basadas en diccionario, antes de proceder a explicar los más

⁴ Los agentes son empleados del Banco en todo el Reino Unido que proporcionan inteligencia sobre las condiciones de la economía local.

complejos, como el análisis semántico latente, la asignación de Dirichlet latente y la clasificación jerárquica descendente.

La minería de textos booleana y la basada en diccionarios, por un lado, y el análisis semántico latente, la asignación de Dirichlet latente y la clasificación jerárquica descendente, por el otro, son técnicas que se sirven de epistemologías diferentes, es decir, aproximaciones diferentes de la generación de conocimiento: la deducción y la inducción.⁵ La deducción parte de una teoría general y luego comprueba si es válida utilizando ciertos conjuntos de datos. En contraste, la abducción se basa en algunos datos para intentar inferir la mejor explicación de cierto suceso, pero sin el propósito de obtener una explicación generalizable a otros casos.⁶ La minería de textos booleana y la basada en diccionarios son técnicas deductivas porque parten de una lista predefinida de palabras, motivada por una teoría general respecto al porqué de su importancia. Las fortalezas de esta técnica son su sencillez y su escalabilidad. El código para su instrumentación por lo general tiene pocas líneas de longitud y puede aplicarse fácilmente a archivos de texto enormes. La debilidad de esta técnica es que se centra sólo en las palabras que el investigador consideró previamente como informativas; todas las demás son ignoradas. Por otro lado, el análisis semántico latente, la asignación de Dirichlet latente y la clasificación jerárquica descendente infieren patrones temáticos en cierto corpus sin afirmar que tales patrones se presentan en otros documentos. La principal fortaleza de estas técnicas es que analizan todas las palabras dentro de la muestra y arrojan resultados estadísticos más

complejos. Su principal desventaja es la complejidad de la programación.

La minería de textos es un tema vasto. Por necesidad, hemos tenido que ser selectivos con las técnicas que tratamos en este documento. Principalmente nos concentramos en las técnicas de aprendizaje automático sin supervisión. *El aprendizaje automático sin supervisión implica* tomar observaciones *no clasificadas* y sacar a la luz los patrones ocultos que las estructuran en cierta manera significativa.⁷ Estas técnicas pueden contrastarse con el *aprendizaje automático supervisado*. El aprendizaje automático supervisado inicia con las observaciones que *clasifica* el investigador para *entrenar* un algoritmo con *supervisión* humana, con el fin de que *aprenda* la correlación entre las clasificaciones asignadas por el investigador y las palabras características de los documentos en esas clasificaciones (Grimmer y Stewart, 2013). Aunque en la conclusión tratamos brevemente el aprendizaje automático supervisado, el punto focal de este documento son las técnicas de aprendizaje automático sin supervisión porque guardan relación con las prácticas cambiantes del Banco respecto a los *grandes volúmenes de datos* (Bholat, 2015; Haldane, 2015). En el documento, pondremos en cursivas los términos donde aparezcan definidos, tal como lo hemos hecho en esta introducción.

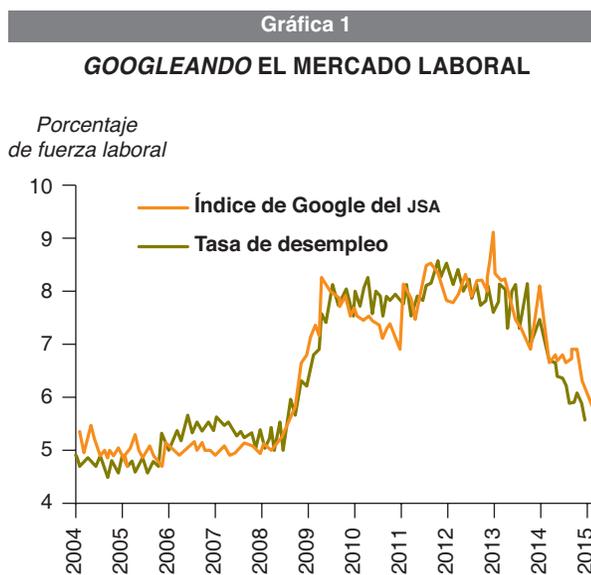
⁵ Por supuesto, la deducción y la abducción son tipos ideales. En realidad, todos los métodos explicativos están mezclados. No obstante, creemos que esta clasificación ayuda a situar las distintas técnicas de minería de textos en términos de sus similitudes y sus diferencias.

⁶ La *inducción* es una tercera epistemología que, como la abducción, parte de datos sin creencia previa; sin embargo, al igual que la deducción, busca producir afirmaciones teóricas generales.

⁷ Los resultados de algoritmos para el aprendizaje automático no supervisado pueden ingresarse en modelos econométricos para predecir alguna variable de interés, pero este es un método distinto a elegir intencionalmente las dimensiones del contenido con base en su capacidad predictiva.

1. EL TEXTO COMO DATOS PARA LA INVESTIGACIÓN QUE REALIZA EL BANCO CENTRAL

Con el fin de estimular un análisis de la minería de textos, empezaremos considerando los temas de investigación medulares que son temas de interés para los bancos centrales, utilizando como variable sustituta el estudio analítico One Bank Research Agenda (OBRA, Banco de Inglaterra, 2015). De hecho, el documento para discusión de OBRA identifica texto como fuente de datos cuyo potencial analítico no ha sido aprovechado del todo. Esto se debe al raudal de nuevos textos con datos disponibles mediante los registros de búsqueda en internet y las redes sociales. McLaren y Shanbhogue (2011) proporcionan un buen ejemplo de lo que puede hacerse. Utilizando datos de Google sobre volúmenes de búsquedas, descubrieron que tales datos proporcionan un seguimiento más oportuno de las variables económicas clave que las estadísticas oficiales. Por ejemplo, la gráfica 1 muestra que las búsquedas en Google del Jobseeker's Allowance (JSA)⁸ prácticamente replican la cifra oficial de desempleo.



Fuente: McLaren and Shanbhogue (2011).

Un tema que interesa a los bancos centrales es la medición del riesgo y de la incertidumbre en la economía y el sistema financiero. La investigación

⁸ Subsidio por desempleo que paga el gobierno del Reino Unido.

de Nyman *et al.* (2015) fue una contribución reciente en este sentido. Nyman y sus coautores parten de una teoría general: la hipótesis de las finanzas emocionales. Conforme a esta, las personas se convencen de tomar posiciones en los mercados financieros creando narraciones respecto a los posibles resultados de sus acciones. Estas narraciones de convencimiento incorporan emociones como excitación por las ganancias esperadas y ansiedad por las posibles pérdidas. De acuerdo con Nyman y sus coautores, estas narraciones no las fabrican las personas aisladamente. Más bien, se construyen socialmente mediante interacciones como son las conversaciones entre personas. Es por medio de estas interacciones sociales que las narraciones se crean y se difunden, pudiendo así repercutir en el precio de los activos.

Para probar su hipótesis, analizaron tres fuentes de datos de texto: el comentario del mercado que proporciona diariamente el Banco (2000-2010), los informes de investigación de inversionistas (2010-2013) y el archivo de noticias de Reuters (1996-2014). El sentimiento se mide mediante el coeficiente de sentimiento construido en la ecuación 1 (ver Nyman *et al.*, 2015).

$$1 \quad SI[T] = \frac{(|Excitement| - |Anxiety|)}{|T|},$$

donde $SI[T]$ es el coeficiente de sentimiento del documento T ; $|Excitement|$ es el número de palabras que expresan “excitación”; $|Anxiety|$ es el número de palabras que expresan “ansiedad”; y $|T|$ es el número total de palabras en el documento T .

El signo del coeficiente indica el sentimiento del mercado: alcista, si es positivo; bajista, si es negativo. El coeficiente luego se compara con los acontecimientos históricos y otros indicadores financieros.

Los autores también miden el consenso de la narración. En particular, su método consiste en agrupar artículos en grupos temáticos.⁹ La

⁹ En particular, los autores utilizan un algoritmo de agrupamiento de X medias, que emplea criterios de información bayesianos para determinar el número óptimo de

incertidumbre en la distribución de los temas entonces actúa como variable sustituta de la incertidumbre. En otras palabras, la entropía reducida en la distribución de temas se utiliza como indicación de la concentración de los temas o consenso. En la gráfica 2 se muestra la serie de tiempo del índice de sentimiento y la medida del consenso. Los autores encontraron evidencia de comportamiento de manada (entropía disminuida) y mayor excitación antes de la crisis financiera reciente.

Una vez que se mide la incertidumbre en la economía, el objetivo de los bancos centrales es manejarla. Esta es una de las motivaciones principales de la reciente política del Banco de guía prospectiva (Carney, 2013), por la cual el Banco dirige las expectativas respecto a la orientación futura de la política mediante la comunicación de sus intenciones futuras y sus pronósticos oficiales. La minería de textos puede ayudar en este caso y otros parecidos al medir en qué grado los funcionarios del Banco están comunicando un mensaje congruente al resto del mundo.¹⁰ Y la evaluación de la eficacia de las comunicaciones del Banco es un área de investigación identificada mediante el documento de discusión OBRA.

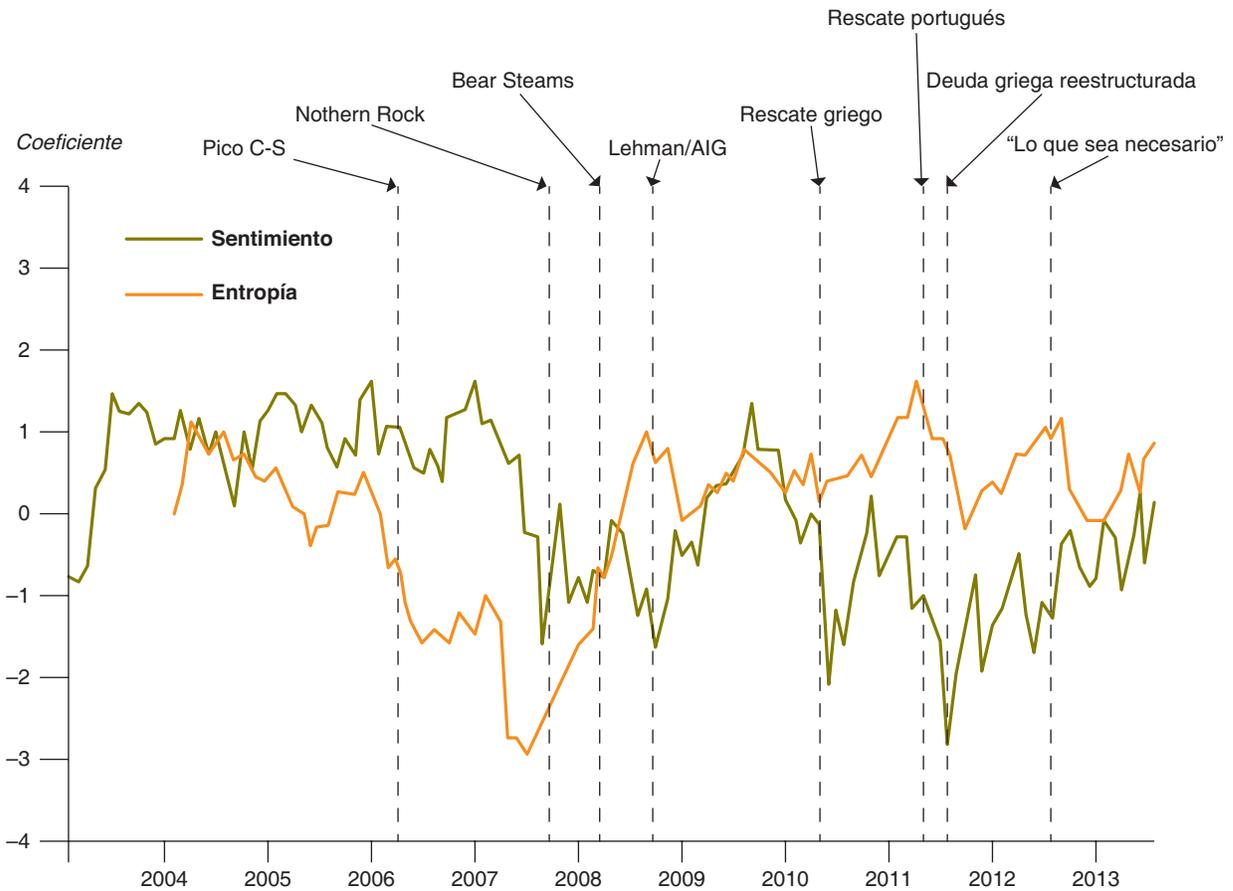
La gráfica 3 de Vallès y Schonhardt-Bailey (2015) ejemplifica el tipo de investigación que puede efectuarse. En la gráfica se muestra el contenido temático de los discursos del CPM y las minutas durante

agrupamientos. Luego, usan la entropía de Shannon como medida de la distribución de temas. El consenso incrementado se mide mediante 1) una reducción en el número de agrupamientos de temas, cuando el tamaño real de cada agrupamiento no cambia y 2) el crecimiento relativo de un tema determinado, para un número fijo de agrupamientos de temas.

¹⁰ Ver Rosa y Verga (2006), Blinder *et al.* (2008), Jansen y Haan (2010), Bennani y Farvaque (2014) para investigaciones similares sobre la congruencia de la comunicación de los bancos centrales. Sin embargo, no siempre es bueno que la comunicación sea congruente. Por ejemplo, Humpherys *et al.* (2011) crearon modelos para detectar declaraciones financieras fraudulentas en las comunicaciones de la dirección y encontraron evidencia de que las declaraciones fraudulentas tienen más probabilidad de contener menos diversidad léxica.

Gráfica 2

SENTIMIENTO Y ENTROPÍA EN LOS ARCHIVOS DE NOTICIAS REUTERS



Fuente: Nyman *et al.* (2015).

el último año de la gestión de Mervin King y el primero de la de Mark Carney.

Cada gráfica representa espacialmente co-ocurrencias, es decir, la convergencia y la divergencia de personas al hablar sobre ciertos temas. La proximidad espacial sugiere un mayor grado de coocurrencia. Por ejemplo, en ambas gráficas, la categoría temática *economía real* está muy relacionada con el CPM cuando habla como comité en sus minutas.

Puede identificarse una diferenciación evidente en los temas tratados por los miembros del CPM en sus discursos externos durante la era de Carney.

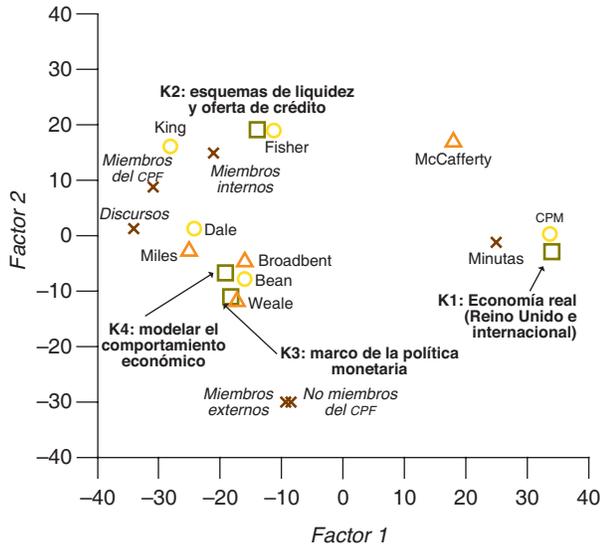
Algunos miembros utilizaron sus discursos para comentar la guía prospectiva, pero otros no. Así, Vallès y Schonhardt-Bailey arrojaron luz respecto a dónde transmite un mensaje el comité como un todo, mientras que los miembros individuales están mandando mensajes más variados.¹¹

¹¹ Otros estudios recientes que emplean la minería de textos para entender las comunicaciones de los bancos centrales son el análisis de Bulir *et al.* (2014) respecto a los informes de inflación de los bancos centrales y el que hace Siklos (2013) de las minutas de cinco bancos centrales, para demostrar la manera como cambió su tono después de la crisis financiera. También Nergues *et al.* (2014) obtienen mediciones de red para investigar cambios en el discurso del Banco

Gráfica 3

CONTENIDO TEMÁTICO DE LAS MINUTAS DEL CPM^a

ANÁLISIS CORRESPONDIENTE DE VARIABLES DE CLASES Y PASIVOS.
MINUTOS Y DISCURSOS DEL CPM, GOBERNACIÓN DE KING
(JULIO 2012-JUNIO 2013)

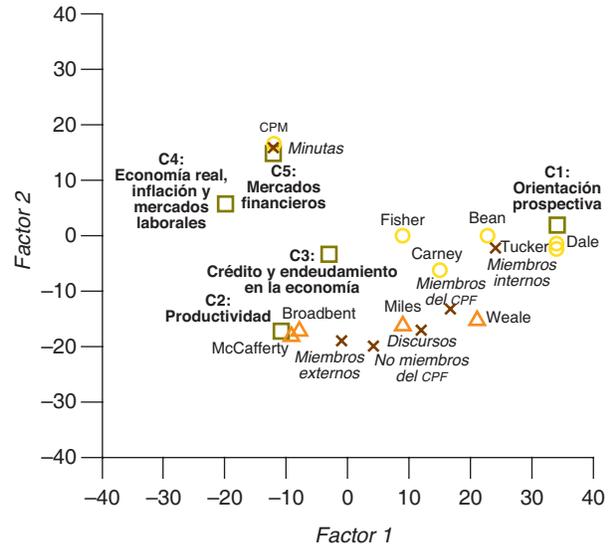


- Clase
- × Atributo
- Miembro interno
- △ Miembro externo

	Asociación	Acumulativo
Factor 1	44.3%	44.3%
Factor 2	31.2%	75.5%

- Clase 1** (40%) — Economía real (RU e internacional)
- Clase 2** (26%) — Esquemas de liquidez y oferta de crédito
- Clase 3** (18%) — Marco de la política monetaria
- Clase 4** (16%) — Modelar el comportamiento económico

ANÁLISIS CORRESPONDIENTE DE VARIABLES DE TEMAS Y PASIVOS.
MINUTOS Y DISCURSOS DEL CPM, GOBERNACIÓN DE CARNEY
(JULIO 2013-JUNIO 2014)



- Clase
- × Atributo
- Miembro interno
- △ Miembro externo

	Asociación	Acumulativo
Factor 1	33.5%	33.5%
Factor 2	26.3%	59.8%

- Clase 1** (27%) — Orientación prospectiva
- Clase 2** (27%) — Productividad
- Clase 3** (10%) — Crédito y endeudamiento en la economía
- Clase 4** (24%) — Economía real, inflación y mercados laborales
- Clase 5** (12%) — Mercados financieros

^a Estas gráficas representan las correlaciones entre temas y oradores en la gobernación de King y Carney. Las posiciones de los puntos y la distancia entre puntos reflejan el grado de co-ocurrencias. Los ejes identifican la cantidad máxima de asociación por factores, como se explica con más detalle en la sección 2.

Fuente: Vallès y Schonhardt-Bailey (2015).

Los bancos centrales desean saber si están comunicando un mensaje congruente, pero también si las distintas políticas que proclaman se complementan o están en conflicto. De hecho, el documento de discusión OBRA señala la comprensión de las interacciones de la política monetaria, la macroprudencial y la microprudencial como un

Central Europeo antes de la crisis financiera y después de ella.

importante tema de investigación para el Banco. La minería de textos podría ser útil para entender estas interacciones. Aquí me inspiro en el ensayo reciente de William Li y coautores titulado “Law is Code” (Li *et al.*, 2015), quienes rastrean la complejidad creciente de la jurisprudencia estadounidense con el paso del tiempo mediante un análisis de todo el código legal de Estados Unidos (el US

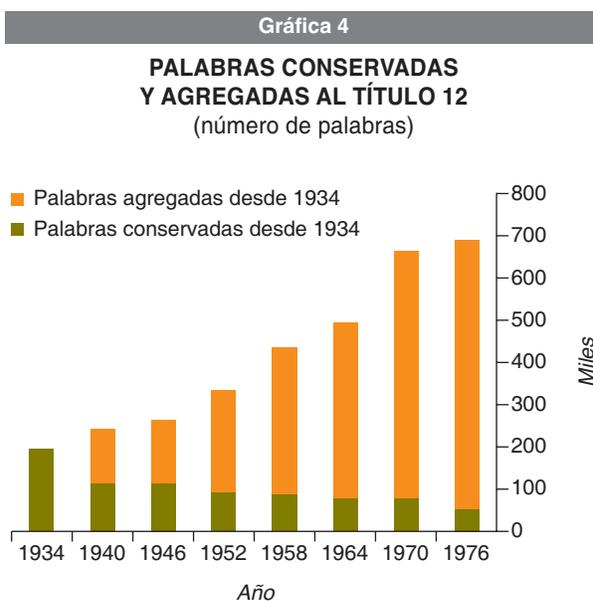
Code) desde 1926 hasta la fecha. Los autores, que evocan el tono de Haldane (2012) en su crítica a la regulación financiera compleja, argumentan que la complejidad creciente de esa compilación de leyes dificulta su comprensión, tiene consecuencias negativas involuntarias y es un posible lastre para la productividad. Con el fin de consignar la complejidad creciente del Código de Estados Unidos, los autores produjeron varias medidas basadas en textos, entre otras:

1) Las medidas de tamaño y contenido del Código con el tiempo. Los autores interpretan que el aumento del Código, conforme a su número de palabras, significa que cada vez es más gravoso. Señalan que el tamaño bruto del Código y su tasa de aumento han crecido en los últimos decenios. Asimismo, dan seguimiento a los cambios en el contenido de secciones específicas del Código mediante la comparación de las palabras agregadas con las eliminadas con el paso del tiempo. Por ejemplo, la gráfica 4 muestra los cambios en el Título 12 (Bancos y Banca) del Código entre 1934 y 1976.

2) Medidas de complejidad ciclomática. Para este estudio consideraremos que la complejidad ciclomática se refiere al conteo de los enunciados condicionales en el Código, es decir, las cláusulas que empiezan con la palabra “si”, “salvo”, “a menos que” y palabras elusivas similares. Los autores argumentan que los enunciados condicionales causan incertidumbre respecto a cuándo se aplican las leyes y, por lo tanto, su presencia indica una mayor complejidad.

3) Medidas de interrelación. Como en el sistema financiero, la interrelación en el derecho puede generar diseminaciones y consecuencias involuntarias (Gai *et al.*, 2011). Li y sus coautores miden la interrelación legal buscando las citas entre distintas partes del Código. Las secciones que tienen muchas referencias y se citan en otras partes son nodos *sistémicamente importantes* en la red jurídica. En el cuadro 1 se enuncian las secciones más importantes del Título 12. Este análisis identifica qué secciones del Código, si se reformaran, repercutirían mucho en otros aspectos de la legislación.

Son varias las aplicaciones obvias del método de Li y sus coautores para los asuntos que competen a los bancos centrales. Por ejemplo, podrían realizarse mediciones de la interrelación textual antes de cualquier cambio propuesto a la regulación o a las formas de informe regulatorio. Esto podría ayudar a cuantificar de antemano las posibles interacciones adversas de los cambios monetarios, macroprudenciales y microprudenciales. En términos más generales, las mediciones que dan seguimiento a los cambios en el tamaño, el contenido, la complejidad ciclomática y la interrelación de la regulación podrían calcularse en el Reino Unido y en otros entornos para someter a prueba la hipótesis de que la normatividad financiera se ha vuelto más compleja con el tiempo.



Fuente: Li *et al.* (2015).

SECCIONES DEL TÍTULO 12 DEL CÓDIGO DE ESTADOS UNIDOS CON MAYOR INTERRELACIÓN

Número de sección	Nombre
1841	Definiciones de la Ley de Controladoras Bancarias
101	Abrogada (entrega de billetes en circulación)
1818	Terminación del Estatus como Institución de Depósito Asegurada
1709	Seguro de Hipotecas
1813	Definiciones de la Ley Federal de Seguros de Depósitos

Fuente: W. P. Li, P. Azar, D. Larochele, P. Hill y A. W. Lo (2015), "Law is Code: A Software Engineering Approach to Analyzing the United States Code", *Journal of Business & Technology Law*, vol. 10, núm. 2, pp. 297-374.

2. INTRODUCCIÓN A LAS TÉCNICAS DE MINERÍA DE TEXTOS

En la primera sección de este texto se planteó que la minería de textos es prometedora para los bancos centrales. En esta sección enumeramos los pasos principales de cualquier proyecto de minería de textos y proporcionamos una visión general de algunas técnicas específicas de la minería de textos. Todas las técnicas descritas en esta sección son una especie de *análisis del contenido* que resume de qué tratan los textos. Lo hacen contando el número de palabras de un corpus. La lógica subyacente es que la frecuencia de las palabras y su co-ocurrencia son buenos indicadores del tema o del sentimiento expresado en los textos.

Para hacer más concreta esta intuición, considérense las *nubes de palabras* en la gráfica 5. La nube de palabras en la parte superior proviene del discurso inaugural del gobernador Carney en la conferencia sobre el OBRA organizada por el Banco (Carney, 2015); la de la parte inferior proviene del artículo sobre tal evento publicado en el *Financial Times* (Giles, 2015).

Como puede verse en la gráfica 5, los problemas ocurren con las palabras muy frecuentes y las poco frecuentes. Por ejemplo, las palabras *the* y *and* aparecen muy a menudo y, por ello, no nos ayudan a distinguir un documento de otro. Y al contrario, hay muchas palabras que sólo aparecen una vez.

La gráfica 6 simplifica las nubes de palabras de la gráfica 5 al mostrar únicamente aquellas que aparecen por lo menos dos veces en los textos y eliminar las palabras *the* y *and*. El contenido de ambas nubes es similar salvo en la mención que hizo Mark Carney de sí mismo en tercera persona, por lo que *Mr. Mark Carney* no aparece en la nube de la parte superior.

columnas que corresponden a cada documento. El elemento $X_{i,j}$ representa la cuenta del i -ésimo término en el documento j .

Cuadro 2

**REPRESENTACIÓN MATRICIAL
TÉRMINO-DOCUMENTO DE LAS NUBES
DE PALABRAS**

Término	Comentarios de Mark Carney sobre la agenda de investigación del Banco	Artículo en el <i>Financial Times</i> sobre la agenda de investigación del Banco de Inglaterra
policy	37	8
research	35	11
bank	28	10
that	28	3
our	25	2
financial	18	3
central	16	1
thinks	0	1
uses	0	1
wanted	0	1
won't	0	1

La tendencia a ser altamente dimensional y dispersa suele caracterizar a las matrices de término-documento como la del cuadro 2. Cualquier documento dado contendrá sólo un subconjunto de todos los términos únicos y las filas correspondientes a los términos no utilizados serán todas cero. Por ejemplo, la matriz completa de término-documento para los textos relacionados con el OBRA tiene 906 dimensiones (o términos únicos) y 42% del recuento de palabras es cero. Por lo anterior, la tarea fundamental consiste en extraer la información escasamente dimensional de los documentos que, por naturaleza, son altamente dimensionales. Esto equivale a una situación en la

que un investigador tiene una base de datos con miles de *covariantes* y está intentando elegir cuál subconjunto de estas, o cuál resumen estadístico, debería incluirse en el análisis de regresión.

Por lo tanto, aunque la lógica de utilizar el recuento de palabras como indicador del contenido de los textos es simple, su puesta en práctica es más complicada. En las siguientes subsecciones nos centraremos en las técnicas para reducir la *dimensionalidad* (el número de palabras o variables) y la dispersión.

2.1 PREPROCESAMIENTO ANALÍTICO

El primer paso en la minería de textos es definir el corpus en cuestión. Si el objetivo es generalizar, a partir del corpus, una población más grande de documentos, entonces se aplican las reglas estándar de muestreo. Los documentos deben ser representativos y seleccionarse aleatoriamente o mediante algún otro método de muestreo probabilístico. Un problema que a veces se presenta en esta etapa es la *duplicación*, que un corpus incluya el mismo documento varias veces. Por ejemplo, en las grandes bases de datos de periódicos, el mismo artículo aparece más de una vez, tal vez porque existen ediciones del mismo periódico en países diferentes. Tal duplicación puede distorsionar la inferencia al dotar de mayor representatividad a ciertos documentos. La eliminación de duplicaciones puede hacerse manualmente mediante la revisión del corpus para asegurarse de que cada documento sea una observación única, o mediante un algoritmo, como en Eckley (2015).

Una vez que el corpus se ha especificado, el siguiente paso es pasarlo a un formato analizable. Esto suele ser un proceso engorroso y tardado. Por ejemplo, los documentos en formato PDF pudieran convertirse y guardarse como archivos de texto (txt). Por otro lado, si los textos están sólo impresos, como sucede en muchos archivos, pudiera ser necesario digitalizarlos y convertirlos a archivos de texto mediante *programas para reconocimiento óptico de caracteres*.

Una vez que los textos se han convertido al formato apropiado, el siguiente paso consiste en segmentar el documento en *tokens*. Este paso implica representar el texto como una lista de palabras, números, signos de puntuación y posiblemente otros símbolos, como los de monedas o el de marca registrada.¹² Reiteramos que, en teoría, parece sencillo, pero no lo es tanto en la práctica.

Considérese este extracto, de una sola oración, del enunciado de la misión del Banco: *Promoting the good of the people of the United Kingdom* [Promover el bienestar del pueblo del Reino Unido]. El número de palabras en esta oración podría contarse de dos maneras por lo menos. Una es contar cada palabra discreta o token. Haciéndolo así, el extracto tiene diez tokens. Pero otra manera de contar el número de palabras en la oración es contando las palabras distintas. Así que, si bien el extracto incluye diez tokens, hay sólo siete palabras distintas, porque el token *the* aparece tres veces y *of* aparece dos veces. Podemos representar el número de tipos de palabras como un vector [1,3,1,2,1,1,1], como en el cuadro 3. En la minería de textos, las representaciones vectoriales de texto se denominan *representaciones de bolsa de palabras*.

Obsérvese que la representación vectorial en el cuadro 3 trata a *United* y *Kingdom* como palabras independientes. Sin embargo, otra representación, que supuestamente refleja mejor el significado o *semántica* de la oración, pudiera tratar ambos tokens como un único concepto, es decir, *United Kingdom*.

Escribir un algoritmo para dividir en tokens el texto de manera que siempre transmita el significado correcto es difícil. Por ejemplo, si un algoritmo representara cada caso en el que *United* está seguido de la palabra *Kingdom* como una aparición

¹² Una manera de segmentar en *tokens* es mediante una función de expresión regular que permita a los investigadores buscar patrones en el texto y dividir el texto con base en tales patrones. Por ejemplo, una búsqueda de las palabras *Bank of England* pudiera utilizar una expresión estándar que busque una palabra que comience con la letra *B* seguida de tres caracteres alfabéticos, un espacio, la palabra *of*, un espacio y una palabra que empiece con la letra *E*.

Cuadro 3

UNA REPRESENTACIÓN VECTORIAL DE PALABRAS

<i>Término</i>	<i>Documento: Enunciado de la misión del Banco (extracto)</i>
promoting	1
the	3
good	1
of	2
people	1
United	1
Kingdom	1

Cuadro 4

UNA REPRESENTACIÓN VECTORIAL DE TOKENS

<i>Término</i>	<i>Documento: Enunciado de la misión del Banco (extracto)</i>
Promoting	1
the	3
good	1
of	2
people	1
United Kingdom	1

del término *United Kingdom*, entonces el algoritmo trataría incorrectamente “united” y “kingdom” en la siguiente oración como un token en vez de dos: *The marriage of Isabella of Castile to Ferdinand of Aragon created a united kingdom in Spain* [El matrimonio de Isabel de Castilla y Fernando de Aragón creó un reino unido en España].

Además de ejemplificar lo difícil que es segmentar en tokens el texto para que el recuento de palabras transmita significado con precisión, la oración anterior también muestra que un recuento

de palabras desajustado pudiera ser un mal indicador del contenido distintivo de un documento.

En el ejemplo, la palabra *the* es la moda. Así que, como dijimos al principio de esta sección, un aspecto fundamental de la minería consiste en reducir la dimensionalidad de las representaciones de bolsa de palabras para eliminar el *ruido* y centrarse en el contenido distintivo de los documentos. Hay varias técnicas para lidiar con las palabras que son irrelevantes para el contenido del corpus. Mencionamos algunas a continuación.

1) *Eliminar los signos de puntuación y las palabras raras.* Palabras como *el, de, que* y similares son muy comunes, pero ayudan poco a distinguir el contenido de un documento del contenido de otro. Así que tales palabras vacías, como se les conoce, con frecuencia se descartan de la muestra.¹³ Por ejemplo, si descartáramos las palabras *of* y *the*, entonces el extracto del enunciado de la misión del Banco de Inglaterra quedaría representado de la siguiente manera:

Tabla 5

REPRESENTACIÓN VECTORIAL UTILIZANDO PALABRAS VACÍAS

<i>Término</i>	<i>Documento: Enunciado de la misión del Banco (extracto)</i>
Promoting	1
good	1
people	1
United Kingdom	1

2) *Representar las palabras con su raíz lingüística común.* Otro procedimiento para reducir la dimensionalidad es la *lematización*. La

¹³ Una lista de palabras vacías se puede conseguir en: <<http://snowball.tartarus.org/algorithms/english/stop.txt>>.

lematización utiliza el etiquetado de partes del discurso para determinar la categoría gramatical (parte del discurso) a la que pertenece cada palabra (sustantivo, pronombre, verbo, etc.) y convertirla a su forma base. Por ejemplo, *pienso*, etiquetado como verbo, se convertiría en *pensar*, pero no si fuera etiquetado como sustantivo. El etiquetado de partes del discurso es difícil porque las palabras suelen pertenecer a más de una categoría. Por ejemplo, la palabra *libro* es tanto un sustantivo como un verbo conjugado. Cuando una palabra recibe un etiquetado múltiple, pueden utilizarse algoritmos sintácticos para determinar el etiquetado correcto con base en las palabras vecinas. Considérese la oración: *Fed increases interest rate* [La Fed sube la tasa de interés] (Jurafsky y Manning, 2012). En inglés cada una de estas palabras puede ser un sustantivo o un verbo. Por ejemplo, la palabra *Fed* puede referirse a la Reserva Federal o al pasado del verbo *feed*. Un algoritmo sintáctico podría desambiguar cada uno de estos tokens mediante referencia a otro, hasta que a cada token se le haya asignado una categoría gramatical.

3) *Extracción de raíces.* En la práctica, muchos mineros de textos simplemente extraen las raíces de las palabras porque este procedimiento tiende a ser más rápido y más sencillo que la lematización.¹⁴ La extracción de raíces implica cortar los afijos y contar sólo las raíces. Por ejemplo, la palabra “bancario” contiene la raíz “banca” y el afijo “rio”. Por lo tanto, “bancario” y “banca”, una vez extraída su raíz, serían tratadas como dos apariciones del mismo token.¹⁵

¹⁴ El algoritmo de Porter es popular para los textos en inglés.

¹⁵ El resultado después de la extracción de raíces tal vez no sea una palabra que ocurra naturalmente y pudiera carecer de significado interpretativo. Por ejemplo, la raíz de *inflation* es *inflat*. Como ya dijimos, la lematización proporciona un resultado más refinado, es decir, por lo general arroja la forma de diccionario de una palabra.

4) *Conversión a minúsculas*. La conversión a minúsculas implica pasar todos los tokens alfabéticos a minúscula. Aunque en algunos casos esto podría oscurecer el significado de algunos nombres propios (por ejemplo, un acrónimo como *US* para *United States* podría convertirse erróneamente en el pronombre *us*, nosotros en inglés), este procedimiento casi siempre considera irrelevante que la palabra se encuentre al principio de una oración. No obstante, hay casos en los que la conversión a minúsculas puede resultar engañosa. Considérese la oración: *The Bank of England is the main regulator of XYZ bank* [El Banco de Inglaterra es el principal regulador del banco XYZ]. La conversión a minúsculas trataría *Bank* y *bank* como apariciones del mismo token cuando, de hecho, se trata de entidades jurídicas diferentes.

2.2 TÉCNICAS BOOLEANAS

Hasta ahora hemos considerado los pasos de preprocesamiento comunes en cualquier investigación de minería de textos. Ahora toca el turno a los distintos tipos de técnicas de minería de textos.

Tal vez el método de minería de textos más sencillo es realizar una búsqueda booleana de uno o más términos. *Las técnicas de búsqueda booleana* combinan términos individuales o frases con operadores lógicos como AND (y), OR (o) y NOT (no) para formar expresiones de búsqueda. Después, a cada documento se le asigna un 1 o un 0 dependiendo de si la expresión es verdadera (= 1) o falsa (= 0). Dado que la búsqueda booleana ya forma parte de los principales buscadores de internet, tiene la gran ventaja de que los investigadores no tienen que ingresar en los textos sin procesar si estos ya han sido indizados. Por ejemplo, Google y Yahoo arrojan un número aproximado de sitios web que cumplen con los criterios de la búsqueda booleana.

Un ejemplo de investigación relevante para los bancos centrales que utilizan técnicas booleanas es la realizada por Baker, Bloom y Davis (2013), quienes calcularon la incertidumbre asociada a la política económica analizando los principales

periódicos de Estados Unidos y Europa. Contaron diariamente cuántos de los artículos publicados a partir de 1985 contenían palabras relacionadas con la incertidumbre y la economía. Específicamente, los autores computaron el número de artículos diarios que cumplían con los siguientes criterios de búsqueda:

- 1) el artículo contiene *incierto* OR *incertidumbre*, AND
- 2) el artículo contiene *económica* OR *economía*, AND
- 3) el artículo contiene *congreso* OR *déficit* OR *reserva federal* OR *legislación* OR *regulación* OR *casa blanca*.

Las series de tiempo de este *recuento normalizado*—conforme al número total de artículos de periódico cada mes— actúa entonces como variable sustituta de la incertidumbre respecto a la política.¹⁶ Como se observa en la gráfica 7, su índice pudo consignar importantes sucesos políticos y de los mercados financieros.

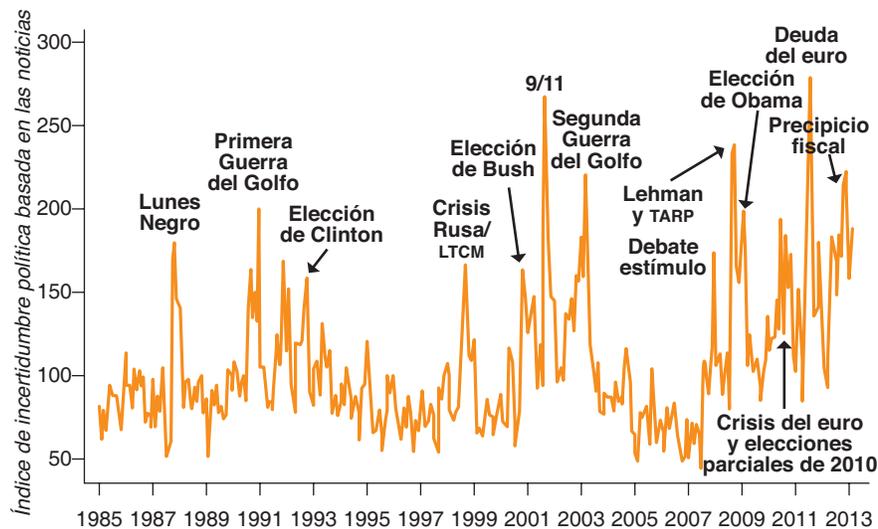
Otro ejemplo de minería booleana de textos es el trabajo reciente del Banco en la antesala del referendo por la independencia de Escocia. En ese entonces existía preocupación por las secuelas de una votación a favor, como por ejemplo, corridas contra los bancos escoceses (Banco de Inglaterra, 2014). Dadas las fluctuaciones en los resultados de las encuestas de opinión y el sentimiento público, el Banco monitoreó estos riesgos mediante un análisis del tráfico en Twitter. Se dio seguimiento a los tuits que contenían términos y combinaciones de términos relativos a corridas bancarias durante una semana completa, pero particularmente en la noche de la votación. Al final, el tráfico fue poco y resultó evidente el voto opuesto a la independencia.¹⁷

¹⁶ Ver <<http://www.policyuncertainty.com>> para más detalles.

¹⁷ En general, las redes sociales tienen propiedades que las hacen atractivas como fuente de datos. (O'Connor *et al.*, 2010; Bollen *et al.*, 2011). En particular, los datos de las redes sociales se generan durante la noche, gracias a lo cual las autoridades pueden vigilar los riesgos después del cierre

Gráfica 7

ÍNDICE DE INCERTIDUMBRE POLÍTICA BASADA EN LAS NOTICIAS



Fuente: Baker, Bloom, and Davis (2013).

En conclusión, podría decirse que las mediciones booleanas del riesgo y la incertidumbre basadas en textos tienen cuatro puntos fuertes en comparación con las mediciones más usuales, como el índice VIX. Primero, al modificar los términos de búsqueda utilizados para los artículos, se puede consignar la incertidumbre más allá de los mercados financieros. Segundo, la fuente de incertidumbre suele enunciarse explícitamente en texto: “la incertidumbre causada por el referendo”, por ejemplo. En contraste, con las medidas de incertidumbre basadas en el precio de los activos financieros, la fuente de la incertidumbre no siempre está clara. Tercero, si buscamos texto en los medios con fechas más antiguas, podemos obtener series de tiempo mucho más amplias que las basadas en el precio de las opciones. Por ejemplo, el VIX sólo está disponible a partir de finales de los años ochenta, mientras que Baker, Bloom y Davis remontaron su índice hasta 1900. Cuarto y último, podemos obtener mediciones multinacionales de

de los mercados y antes de su apertura.

la incertidumbre centrándonos en el texto publicado en los medios de distintos países, mientras que el VIX simplemente es una variable sustituta de la incertidumbre en los precios de las acciones en Estados Unidos, que pueden ser un mal indicador de la incertidumbre en otros países.

2.3 TÉCNICAS DE DICCIONARIO

Otra serie de estrategias para la minería de textos son las técnicas de diccionario, como las utilizadas por Nyman y coautores, tal como señalamos en la sección 1. Las técnicas de diccionario por lo general consisten de dos pasos:

- 1) Primero, definir una lista de palabras clave para consignar el contenido de interés.
- 2) Luego, representar cada documento en términos de la frecuencia (normalizada) de las palabras en el diccionario.

A diferencia de las técnicas booleanas, las técnicas de diccionario miden la intensidad del uso

de las palabras, que pudiera ser una medida más adecuada del contenido del corpus. Sin embargo, la aplicación de las técnicas de diccionario implica que el investigador tenga acceso a los textos sin procesar.

Por ejemplo, permitamos que el diccionario $D = \{\text{labour, wage, employ}\}$ [trabajo, salario, empleo] sea las raíces de diversas palabras relacionadas con los mercados laborales. Podemos entonces representar cada documento d como la participación s_d de palabras que están en el diccionario (ecuación 2).

Frecuencia normalizada de palabras en un documento

$$2 \quad s_d = \frac{\left(\begin{array}{l} \text{Número de ocurrencias de } \textit{trabajo} + \\ \text{número de ocurrencias de } \textit{salario} + \\ \text{número de ocurrencias de } \textit{empleo} \end{array} \right)}{\text{Total de palabras en el documento } d}$$

Los estudios publicados proporcionan varios ejemplos en los que se han empleado técnicas de diccionario para analizar textos sobre finanzas y economía. Por ejemplo, Tetlock (2007) escribió un artículo muy citado que utiliza diccionarios (es decir, léxicos) para medir el tono de la columna “Abreast of the Market” que se publica en el *Wall Street Journal*.¹⁸ En particular, emplea los diccionarios Harvard IV – 4, cuyas listas de palabras reflejan muchas categorías, incluido el sentimiento positivo y negativo, dolor y placer, y rituales y procesos naturales, entre otros.¹⁹ Tetlock luego cuenta el número de palabras en la columna de cada día entre 1984 y 1999. Su primer hallazgo es que la mayor parte de la variación en el recuento de palabras entre columnas refleja virajes entre el optimismo y el pesimismo. Su segundo hallazgo es que un aumento en las noticias negativas influye de manera estadísticamente significativa en los resultados del día siguiente. Después de Tetlock (2007), otros estudios han empleado la misma estrategia básica de contar palabras de

interés en el contexto de los mercados financieros y correlacionarlas con el precio de los activos (por ejemplo, Aase, 2011).

Sin embargo, reiteramos que hay muchos casos en los que un mero recuento de palabras puede ser engañoso. Considérese la clase de minería de textos conocida como *análisis del sentimiento*, llamada también *minería de opinión* o *análisis de subjetividad*. La finalidad del análisis del sentimiento consiste en detectar el sentir expresado respecto a cierto objeto, o como se dice en el medio, un objetivo. El objetivo del sentimiento pudiera ser una persona, un suceso, una institución, un objeto, por nombrar unos cuantos. La manera más sencilla de analizar el sentimiento es la detección de polaridad, es decir, una clasificación binaria de sentimiento positivo y negativo.

Sin embargo, la medición precisa del sentimiento es complicada por varias sutilezas lingüísticas comunes como la negación, la ironía, la ambigüedad, las expresiones idiomáticas y los neologismos (Jurafsky y Manning, 2012). Considérese, por ejemplo, el siguiente párrafo:

One would have expected building society X to have been an excellent institution. It had a top-notch CEO and world-renowned analysts. Its services were highly rated by retail clients and its operations were efficient. Yet it was a total failure. [Uno pensaría que la sociedad de constructores X habría sido una institución excelente. Tenía un director general de primera y analistas de renombre mundial. Sus servicios recibían buena calificación de los clientes minoristas y sus operaciones eran eficientes. Aun así, fue un fracaso total.]

Un mero recuento del total de las palabras positivas y negativas no transmitiría el sentimiento del párrafo anterior.

2.4 PONDERACIÓN DE PALABRAS

Un simple recuento pudiera no ser apropiado porque puede exagerar la importancia de un número

¹⁸ Ver <<http://www.wsj.com/news/types/abreast-of-the-market>>.

¹⁹ Ver <<http://www.wjh.harvard.edu/~inquirer>>.

pequeño de palabras muy frecuentes. Esto puede ser causa de problemas por dos razones:

- 1) La ley de Zipf, una observación científica de que la frecuencia de una palabra es inversamente proporcional a su jerarquía relativa en un corpus. Esto significa que una diferencia mínima en la jerarquía relativa de una palabra puede significar una gran diferencia en términos del recuento real de palabras. Así que, depender del recuento simple de palabras pudiera exagerar su importancia comparativa.
- 2) Si una palabra se utiliza en muchos documentos de un corpus, entonces su poder para discriminar entre dos documentos es menor que si sólo aparece en unos pocos documentos. No obstante, pudiera ser recomendable dar mayor peso a las palabras que aparecen en pocos documentos porque pudieran indicar diferencias reales en el contenido.

Para abordar estos problemas, una manera común de ponderación que se utiliza en la minería de textos es la *frecuencia de término-frecuencia inversa de documento* (*tf.idf*), conforme a la ecuación 3.

$$3 \quad tf.idf_{t,d} = (1 + \log f_{t,d}) \cdot \log \left(\frac{D}{d_t} \right),$$

donde D es número total de documentos en el corpus, d_t es el número de documentos en los que aparece el término t y $f_{t,d}$ es la frecuencia del término t en el documento d .

El primer factor en la ecuación 3 es la frecuencia del término t en el documento d , que otorga menor peso a las palabras que aparecen con más frecuencia. El segundo factor es la frecuencia inversa en el documento del término t , que asigna mayor peso a las palabras que aparecen con menos frecuencia.

Un ejemplo de la ponderación *tf.idf* aparece en Loughran y McDonald (2011). Su punto de partida es una crítica a los diccionarios Harvard IV-4 utilizados por Tetlock (2007). Los diccionarios de Tetlock contienen palabras como *impuesto*, *costo* y

pasivo que, si bien transmiten un sentimiento negativo en un contexto general, tienen un tono más neutral en el contexto de los mercados financieros, pues describen prácticas contables cotidianas. Por lo anterior, Loughran y McDonald proponen un diccionario específico de las finanzas y demuestran que puede predecir mejor el rendimiento de los activos que los diccionarios genéricos.²⁰ Sin embargo, tras una ponderación *tf.idf*, el desempeño de los diccionarios genéricos mejora de manera notable.²¹ Esto se debe a que palabras como *impuesto* aparecen en muchos documentos y, por lo tanto, tienen menos peso que otras palabras *realmente* negativas.

2.5 MODELOS DE ESPACIO VECTORIAL

Hasta ahora hemos considerado técnicas que identifican los principales temas *dentro de* los textos mediante una serie predefinida de palabras clave. Ahora consideraremos las técnicas para medir la similitud de los temas *entre* textos.

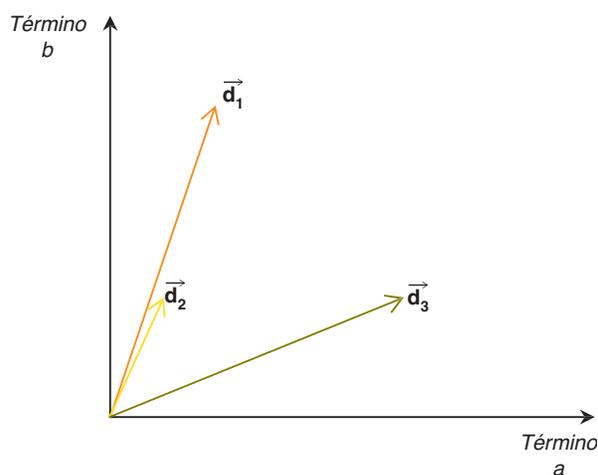
Una manera de medir la similitud de los documentos es utilizando la distancia euclidiana simple. Por ejemplo, Kloptchenko *et al.* (2004) utilizan la distancia euclidiana para encontrar agrupaciones de informes financieros. Sin embargo, como muestra la gráfica 8, esta medición de la distancia tiene limitaciones. La gráfica representa tres documentos hipotéticos, cada uno de los cuales contiene dos términos: a y b . Supóngase que los documentos 1 y 2 utilizan los términos a y b casi en la misma proporción. Sin embargo, debido a que el documento 1 pudiera ser mucho más largo que el documento 2, su distancia es bastante significativa. De hecho, el documento 3, que utiliza el término a en relación con el término b mucho más que el documento 2, se mediría como más parecido al documento 1, simplemente por su longitud similar.

²⁰ Disponible en <http://www3.nd.edu/~mcdonald/Word_Lists.html>.

²¹ Aunque todavía encuentran que su diccionario específico de finanzas tiene un mayor poder explicativo.

Gráfica 8

ESTA GRÁFICA REPRESENTA TRES DOCUMENTOS, CADA UNO CONTIENE DOS TÉRMINOS A Y B.



Nota: documentos 1 y 2 tienen contenido muy similar aunque todavía se encuentran muy separados por la diferencia de longitud.

Este ejemplo muestra las distorsiones que pueden darse cuando se utiliza la distancia euclidiana para medir la distancia entre documentos. Una medida que evita estos problemas es *la similitud coseno*, que refleja el ángulo formado por dos vectores. Volviendo a la gráfica 8, podemos ver que el ángulo formado entre los documentos 1 y 2 es muy pequeño; apuntan a la misma dirección dado que utilizan los dos términos en proporciones casi idénticas. Sin embargo, debido a que las frecuencias de los términos difieren en los documentos 1 y 3, el ángulo es mayor. Si un vector del documento contuviera sólo el término *a* y el otro sólo el término *b*, los vectores serían *ortogonales*. Así que, midiendo el coseno del ángulo θ formado por dos documentos en el espacio vectorial se obtiene una medida de similitud independiente de la longitud de los documentos. La fórmula para computar la similitud coseno se proporciona en la ecuación 4.

4

$$\cos \theta = \frac{\vec{d}_1 \cdot \vec{d}_2}{\|\vec{d}_1\| \|\vec{d}_2\|}$$

donde \cdot es el operador producto punto y $\|\vec{d}_i\|$ es la longitud del vector que representa al documento *i*.

Hoberg y Phillips (2010) proporcionan un ejemplo de modelación con espacio vectorial en economía. Tomaron declaraciones de productos de empresas de los formularios 10K presentados ante la Comisión de Valores y Bolsas de Estados Unidos y computaron su similitud de coseno. Luego utilizaron esta puntuación de similitud como variable sustituta del sector al que pertenecen las empresas. Los autores argumentaron que su análisis proporciona una medida mucho más rica y continua de la intercambiabilidad de productos que los códigos tradicionales de clasificación sectorial.²²

Sin embargo, la similitud de coseno no es una panacea. El cuadro 6 muestra dos documentos hipotéticos. Ambos tratan sobre la educación, pero emplean palabras diferentes para abordar el tema. Esta característica del lenguaje se conoce como *sinonimia*, es decir, el mismo tema subyacente puede describirse haciendo uso de muchas palabras diferentes. Aunque el tema es el mismo en ambos documentos, por emplear un vocabulario diferente su similitud de coseno será baja.

Cuadro 6

SINONIMIA EN UNA MUESTRA DE DOCUMENTOS

Documento 1				
escuela	universidad	colegio	maestro	profesor
0	5	5	1	2
Documento 2				
escuela	universidad	colegio	maestro	profesor
10	0	0	4	0

Nota: este cuadro muestra dos documentos con contenido diferente pero una similitud coseno baja debido a polisemia.

²² Ver, por ejemplo, el sistema británico de clasificación estándar de actividades económicas industriales (Office for National Statistics, 2007).

La polisemia también está relacionada con el problema de la sinonimia: una misma palabra puede tener significados diferentes en contextos diferentes. Considérense los dos documentos en el cuadro 7.

Cuadro 7

POLISEMIA EN UNA MUESTRA DE DOCUMENTOS

		Documento 1					
tanque	marino	rana	animal	naval	guerra		
5	5	3	2	0	0		
		Documento 2					
tanque	marino	rana	animal	naval	guerra		
5	5	0	0	4	3		

Nota: este cuadro muestra dos documentos con contenido diferente pero una similitud de coseno relativamente alta debido a polisemia.

Supóngase que el documento 1 es sobre animales y el documento 2, sobre cuestiones bélicas. El problema es que la misma palabra tiene diferentes significados dependiendo del contexto. Por ejemplo, en una conversación sobre animales, *tanque* se refiere al lugar donde viven peces o anfibios; en una conversación sobre guerra, se refiere a un arma mecanizada. Esto significa que dos documentos esencialmente sin relación alguna pueden mostrar una similitud de coseno elevada.²³

²³ Además de la sinonimia y la polisemia, también podemos considerar la relación jerárquica entre palabras utilizando algoritmos basados en tesauros (Jurafsky y Manning, 2012). Decimos que una palabra es el *hipónimo* de otra si la primera palabra es una subcategoría de la otra. Por ejemplo, en la gráfica que sigue, obsérvese que *receptoras de depósitos* es un hipónimo de *institución financiera*. Y, a su vez, *institución financiera* es un hiperónimo de *receptora de depósitos*. Una de las maneras más sencillas de saber si dos palabras son similares es viendo qué tan próximas están en la jerarquía. Definimos la longitud de ruta, $pathlen(w_1, w_2)$, como uno más el número de orillas en la ruta más corta en el gráfico de hiperónimos entre w_1 y w_2 . Luego, la *similitud de*

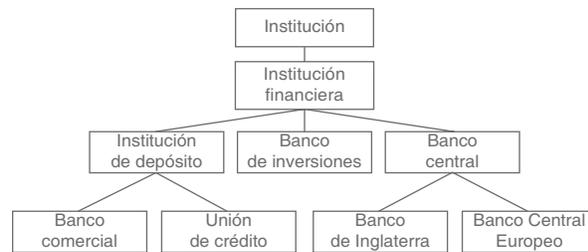
2.6 ANÁLISIS SEMÁNTICO LATENTE

Un supuesto de los modelos de espacio vectorial es que las palabras en un texto son conceptualmente independientes de las demás. Pero esto no siempre es así. Por ejemplo, *naval* y *marino* pudieran considerarse como expresiones superficiales de un tema latente más fundamental, como es *guerra*.

Los *modelos de variable latente* que ahora vamos a comentar adoptan este enfoque. Suponen que las palabras no son independientes y, más bien, están interrelacionadas por temas subyacentes no observados. Estos modelos tienen cuatro virtudes. Abordan la sinonimia al asociar cada palabra en el vocabulario con un tema latente dado. Consignan la polisemia al permitir que cada palabra tenga asociaciones con múltiples temas. También asocian cada documento con temas más que con

longitud de rutas puede definirse mediante

$$pathsim(w_1, w_2) = \frac{1}{pathlen(w_1, w_2)}$$



En el diagrama superior, la longitud de ruta entre *banco comercial* e *institución* es igual a la que existe entre *banco comercial* y *banco de inversión*. Sin embargo, semánticamente, *banco comercial* está relacionado más específicamente con *banco de inversión* que con la idea genérica de *institución*. Esto sugiere que necesitamos una medida más adecuada. Una posible mejora es aumentar palabras en el tesauro con valores de contenido de información (CI). Los valores de CI se computan con base en el conteo de frecuencia de las palabras que se encuentran en el texto. Para cada palabra w en el tesauro, el CI se define como el logaritmo negativo de la probabilidad de esa palabra. Otras técnicas más complejas, como la medida de Resnik (res), toman como dato de entrada dos palabras w_1 y w_2 y producen una medida de similitud. La técnica se basa en la idea del mínimo ancestro común (least common subsumer, LCS), es decir, la palabra más específica que es un ancestro común de las dos palabras. Por ejemplo, el LCS de *banco comercial* y *unión de crédito* es *receptora de depósitos*. El LCS se evalúa directamente desde el CI: $res(w_1, w_2) = CI(LCS(w_1, w_2))$.

palabras. Y permiten a los algoritmos encontrar la mejor asociación entre palabras y variables latentes, sin utilizar listas o categorías de palabras predefinidas, en contraste con la técnicas booleanas y basadas en diccionarios.

Análisis semántico latente (latent semantic analysis, LSA; Deerwester *et al.* 1990) constituye uno de los primeros ejemplos del método de variable latente. El LSA comienza por representar la matriz término-documento mediante la descomposición de valores singulares, que consiste en la búsqueda de los componentes principales de las filas y las columnas de una matriz término-documento. Es decir, el LSA calcula las combinaciones lineales de términos que explican gran parte de la variación de términos entre documentos, así como las combinaciones lineales de documentos que explican gran parte de la variación de documentos entre términos. Entonces, la idea consiste en aproximar la matriz de término-documento utilizando únicamente los componentes principales y medir la similitud de los documentos con la aproximación, más que con la matriz de término-documento verdadera.²⁴ La hipótesis es que los componentes principales representan los temas en común y que los componentes descartados representan las elecciones de palabras idiosincrásicas.

Una aplicación que hizo un banco central recientemente del análisis semántico latente se muestra en un artículo de Acosta (2014), quien estudia el efecto de una mayor transparencia en las reuniones del Comité Federal de Mercado Abierto (Federal Open Market Committee, FOMC) de la Reserva Federal de Estados Unidos.²⁵ La Reserva ha publicado varias versiones estenográficas de las reuniones del FOMC desde octubre de 1993. Antes

de esa fecha, los miembros del FOMC no estaban conscientes de que sus deliberaciones estaban siendo grabadas. Pero después de que el Congreso ejerciera presión para que la Reserva se volviera más transparente, el expresidente de esta, Alan Greenspan, descubrió que el personal había estado transcribiendo palabra por palabra las reuniones desde mediados de los años setenta. Así, convino en publicar las transcripciones previas y darlas a conocer en lo sucesivo con un rezago de cinco años.

Acosta analizó si esa mayor conciencia de los miembros del FOMC respecto a que los comentarios en las reuniones estaban siendo grabados y serían divulgados cambió su comportamiento. El autor aplica la descomposición de valores singulares y utiliza los 200 componentes principales para medir la similitud de documentos. Lo que descubrió fue un mayor apego a las convenciones después de que empezaron a publicarse las transcripciones.

2.7 LA ASIGNACIÓN DE DIRICHLET LATENTE

Un problema de la descomposición de valores singulares es que los temas que produce no son probabilísticos. La *asignación de Dirichlet latente* (latent Dirichlet allocation, LDA) corrige esto.

La LDA es un *modelo de elementos mixtos* en el que palabras y documentos son probabilidades asignadas y se asocian a múltiples temas. Esto contrasta con los *modelos deterministas de membresía* única en el que palabras y documentos se asignan únicamente a un tema. Este aspecto probabilístico de la LDA es importante. Supóngase un tema sobre inflación y otro sobre desempleo. Ahora considérese la palabra *tasa*. De entrada, no se sabe bien a bien a qué tema asociar *tasa*, pues un tema sobre inflación o desempleo podría incluir la *tasa de inflación* o de la *tasa de participación* en la fuerza laboral, respectivamente. Permitir la asignación probabilística de palabras a temas permite esta flexibilidad semántica.

En términos formales cada documento tiene su propia distribución de probabilidad en los temas.

²⁴ Sin embargo, hay otras maneras de realizar la descomposición de la matriz. Por ejemplo, Hendry (2012) aplica la *factorización de matrices no negativas* para estudiar si la comunicación de los bancos centrales afecta a los mercados y, de ser así, cómo lo hace.

²⁵ Hendry y Madeley (2010) y Masawi *et al.* (2014) aplican el análisis semántico latente a documentos de bancos centrales.

Luego, para cada palabra en cada documento se realiza una asignación de tema y, dependiendo de la asignación, una palabra del tema correspondiente.²⁶ Un ejemplo ilustra lo anterior. Supóngase que el hablante A y el hablante B hablan acerca de los temas 1 y 2. El hablante A dedica dos terceras partes de su tiempo al tema 1 y el hablante B dedica dos terceras partes de su tiempo al tema 2. En este caso, las palabras observadas por A pueden pensarse que se eligieron de la siguiente manera. Para cada palabra, A elabora un tema y, siendo la probabilidad del 0.67, este es el tema 1; y con una probabilidad del 0.33, este es el tema 2. Una vez que se elabora un tema para una palabra, la propia palabra se elabora a partir de la distribución de probabilidad asociada a cada tema. Para el hablante B, la probabilidad de que una palabra trate sobre el tema 1 es 0.33 y sobre el tema 2 es 0.67. Pero una vez que una palabra se asigna a un tema dado, se obtiene de la distribución de ese tema que es común a ambos hablantes.

La LDA ha encontrado amplia aplicación en las ciencias computacionales y la estadística, y está empezando a aparecer en la economía. Por ejemplo, Hansen y sus coautores (2014) estudiaron los efectos de una mayor transparencia del banco central. Al igual que Acosta, encontraron evidencia de un mayor apego a las convenciones después de difundirse las transcripciones del FOMC.

Los dos temas que calcularon aparecen en la gráfica 9, que representa temas como nubes de palabras en las que el tamaño de la palabra es

proporcional a su probabilidad en el tema. No se imponen categorías durante el cálculo. Estas son sencillamente las agrupaciones 25^a y 40^a de palabras estimadas. El tema 25 (arriba) claramente trata acerca de la inflación, mientras que el tema 40 (abajo) claramente trata acerca del riesgo. Así que podemos utilizar el resultado de la LDA para medir el contenido subyacente de una manera directamente significativa.

2.8 CLASIFICACIÓN JERÁRQUICA DESCENDENTE

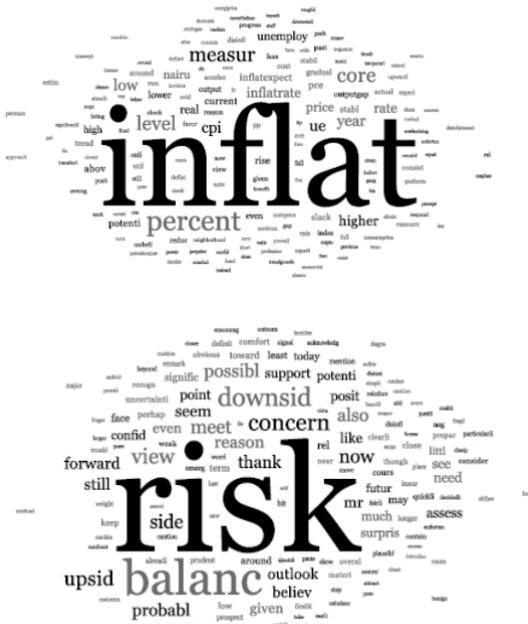
Además de la LSA y la LDA, hay otras técnicas de minería de textos que parten de la base de que las palabras no son independientes unas de otras, sino reflejan temas subyacentes. Por ejemplo, Schonhardt-Bailey (2013) y Vallès y Schonhardt-Bailey (2015) estudiaron los bancos centrales que utilizan el software Alceste.²⁷ En pocas palabras, Alceste intenta descubrir categorías estables de términos que estén asociadas máximamente en el interior, mientras que están mínimamente asociadas a otras

²⁶ Una descripción más formal del modelo es la siguiente. Supóngase un corpus conformado por K temas, donde un tema individual β_k es un vector de probabilidad sobre los elementos únicos V en el vocabulario. Por lo tanto, todas las palabras pueden aparecer en todos los temas, aunque con un peso diferente. Al mismo tiempo, cada documento individual en el corpus es una distribución de probabilidad θ_d sobre los temas K . Cada palabra individual $w_{d,n}$ en el documento d se genera mediante un proceso de dos pasos. Primero, una asignación temática $z_{d,n}$ se obtiene de θ_d . Segundo, una palabra se obtiene de $\beta_{z_{d,n}}$. Esto contrasta con el modelo de membresía única en el que a un documento se le asignaría un único tema $z_{d,n}$, y cada palabra en el documento d se obtendría de β_{z_d} .

²⁷ Alceste es el acrónimo de *Analyse des Lexèmes Co-occurrents dans les Énoncés Simples d'un Texte* (Análisis de los Lexemas Coocurrentes dentro de los Enunciados Simples de Texto). Image-Zafar distribuye el programa. Ver <<http://www.image-zafar.com/>>. El programa fue desarrollado originalmente por Max Reinert. Sus diversas publicaciones en más de 20 años, principalmente en francés, documentan los primeros pasos de Alceste (Reinert, 1983; Reinert, 1987; Reinert, 1990; Reinert, 1993; Reinert, 1998; y Reinert, 2003). Una reproducción de código abierto, basada en R, de Alceste se encuentra en el programa informático de Iramuteq <<http://www.iramuteq.org/>>. Desde 1983, una comunidad de investigadores y analistas de textos cada vez más interdisciplinaria ha empleado Alceste como técnica de minería de textos (Noel-Jorand *et al.*, 1995; Lahlou, 1996; Jenny, 1997; Noel-Jorand *et al.*, 1997; Brugidou, 1998; Guerin-Pace, 1998; Bauer, 2000; Brugidou, 2000; Brugidou, 2003; Noel-Jorand *et al.*, 2004; Schonhardt-Bailey, 2005; Schonhardt-Bailey, 2006; Bara *et al.*, 2007; Schonhardt-Bailey, 2008; Schonhardt-Bailey *et al.*, 2012; Weale *et al.*, 2012; Schonhardt-Bailey, 2013; Vallès y Schonhardt-Bailey, 2015). Peart (2013) utiliza Alceste para estudiar si las preferencias de los miembros del GPM permanecen estables con el tiempo.

Gráfica 9

ESTIMACIÓN DE DOS TEMAS DE LAS TRASCRIPTIONES DEL CORPUS DEL FOMC DURANTE EL PERIODO DE ALAN GREENSPAN



Fuente: Hansen, McMahon y Prat (2014).

categorías.²⁸ Para hacer esto, Alceste construye una matriz que cruza todos los enunciados –denominados unidades elementales de contexto

²⁸ Una manera de proceder más sencilla sería simplemente contar el número de ocurrencias y coocurrencias. Por ejemplo, Ronnqvist y Sarlin (2012) investigan la co-ocurrencia de los nombres de bancos finlandeses en un importante foro financiero en línea. Meten nombres de bancos y sus co-ocurrencias en el mismo registro en una red, donde el tamaño del nodo y el peso de la arista lo proporcionan el número de ocurrencias y de coocurrencias, respectivamente. Aunque es posible percibir algunos cambios en la concentración y fuerza de las conexiones, un proceso más objetivo consiste en consignar esos cambios temporales utilizando las medidas de la centralidad de la red. Los autores detectan un incremento en el número de ocasiones en las que se mencionan a los bancos finlandeses juntos, durante y después de la crisis financiera.

(elementary context units, ECU)²⁹– y todas las palabras, y cuyas celdas indican la presencia o ausencia de esa palabra en el enunciado, de manera parecida a una matriz de término-documento, pero con una unidad más pequeña de análisis textual. Las celdas indican la ausencia o presencia de esa palabra en el enunciado, representada por un 0 o un 1, respectivamente. El cuadro 8 muestra un ejemplo de este tipo de matriz.

Alceste luego divide el contenido de esta tabla en dos categorías, con el fin de maximizar la similitud de las ECU en la misma categoría y, al mismo tiempo, maximizar la diferencia entre categorías. El conjunto total de ECU en la matriz inicial constituye la primera categoría. El algoritmo luego busca una división que minimice el número de palabras traslapadas. El traslape se mide mediante el *valor ji cuadrado* (χ^2) de un cuadro con dos filas, comparando las distribuciones observadas con las esperadas. El algoritmo luego intenta maximizar los valores χ^2 repitiendo el proceso de división (*clasificación jerárquica descendente*), es decir, al probar si mediante dividir las clases en subclases más pequeñas mejora los valores χ^2 . El proceso iterativo de clasificación jerárquica descendente

²⁹ Las unidades elementales de contexto o enunciados son *oraciones medidas* que el programa construye automáticamente con base en la puntuación del texto. En el análisis de textos, un problema persistente y difícil consiste en la longitud óptima de un *enunciado*. Y es que el lenguaje podría analizarse en términos de oraciones, párrafos, frases y demás. Alceste resuelve este problema al no tratar de identificar directamente la longitud del enunciado. Más bien, produce clasificaciones que son independientes de la longitud de los enunciados. Se crean dos clasificaciones, cada una de las cuales utiliza unidades de contexto de distinta longitud; sólo las categorías que aparecen en ambas clasificaciones se conservan para análisis y tales categorías son independientes de la longitud de los enunciados. Esto deja sin clasificar cierto número de ECU, con lo cual se aproxima una medición de la bondad del ajuste. La calidad de la división se mide construyendo un cuadro que cruza todas las categorías obtenidas en la primera clasificación y todas las categorías obtenidas en la segunda clasificación. El resultado es un *cuadro de χ^2 señaladas*, es decir, un cuadro de datos con vínculos positivos y negativos entre la categorías. Este cuadro ayuda a elegir las categorías que comparten el mayor número de ECU.

Cuadro 8

PALABRA MEDIANTE UNA MATRIZ DE ECU

Palabra	ECU						<i>j</i>
	1	2	3	4	5	...	
1	0	1	1	1	1	...	0
2	1	1	0	0	0	...	1
...
<i>i</i>	0	1	1	0	1	...	1
Totales	2	31	5	10	67	...	2

finaliza cuando un número predeterminado de iteraciones ya no produce divisiones estadísticamente significativas.

Por otro lado, para cada categoría se produce una lista de palabras y la fuerza de asociación entre cada palabra y la categoría se expresa mediante un valor χ^2 y un coeficiente fi (ϕ), donde la distribución observada de palabras se compara con una esperada. Por ejemplo, si el vocabulario es diferente en ambas categorías, la distribución observada se desviará sistemáticamente de una distribución esperada consistente de independencia de palabras. Las relaciones entre categorías también pueden descomponerse y examinarse de forma espacial utilizando el factor de *análisis de correspondencia*. Un ejemplo de producción de análisis de correspondencia ya se mostró en la gráfica 3. Las posiciones de los puntos dependen de las correlaciones, donde la distancia refleja el grado de coocurrencia.³⁰ Con respecto a los ejes,

³⁰ Para hacer esto, el análisis de correspondencia utiliza la *distancia ji cuadrado*, que se asemeja a la distancia euclidiana entre puntos en el espacio físico. Cada diferencia al cuadrado entre coordenadas se divide entre el elemento correspondiente del perfil promedio (donde el perfil es una serie de frecuencias divididas entre su total). La razón para utilizar el concepto de ji cuadrado es que permite transformar las frecuencias dividiendo las raíces cuadradas de las frecuencias esperadas, con lo cual se normalizan las varianzas. Esto puede compararse con el análisis factorial, donde los datos en diferentes escalas se estandarizan.

el análisis de correspondencia pretende identificar una cantidad máxima de asociación en el primer eje (horizontal). El segundo eje (vertical) busca dar cuenta de un máximo de la asociación restante.³¹

2.9 APRENDIZAJE AUTOMÁTICO SUPERVISADO

El análisis semántico latente, la asignación Dirichlet latente y la clasificación jerárquica descendente son ejemplos de algoritmos de aprendizaje automático sin supervisión. En contraste, los algoritmos de aprendizaje automático supervisados comienzan cuando el investigador clasifica manualmente los datos de entrenamiento en categorías predefinidas, como en los métodos basados en diccionarios. Con el fin de evitar el problema de sobreajuste, el algoritmo es posteriormente validado en otra serie de documentos, denominados datos de comprobación, antes de que sea aplicado al resto del corpus (Grimmer y Stewart, 2013).

Tal vez la aplicación más fructífera de las técnicas de aprendizaje supervisado en economía es cuando el investigador tiene categorías de textos bien justificadas.³² Una posible aplicación para los bancos centrales consiste en asociar texto con una orientación rigurosa o condescendiente, partiendo de Apel y Grimaldi (2012).³³ Una aplicación bien conocida en economía utilizando datos de textos se encuentra en Gentzkow y Shapiro (2010). Sus datos de entrenamiento consisten en una muestra grande de discursos ante el Congreso de Estados Unidos. Cada discurso lleva una categoría que corresponde al partido del orador e identifica

³¹ Sin embargo, muchos casos requieren más de dos dimensiones para consignar la dimensionalidad de los datos. Así, Alceste proporciona el porcentaje que se acumula a cada dimensión, pero limita la representación gráfica a dos y tres dimensiones.

³² Alternativamente, pudieran haber sido ya asignadas categorías como parte de los metadatos.

³³ Sin embargo, pudiera ser difícil clasificar los documentos de antemano (Grimmer y Stewart, 2013). Por ejemplo, la naturaleza multifacética de las reuniones, los discursos y las conversaciones pudiera dificultar reducir un documento a un único tema.

las frases partidistas. Luego, asignan una calificación a otro corpus conformado por artículos de prensa, como izquierdista o derechista, con base en la presencia o ausencia de frases partidistas.

Un algoritmo popular de aprendizaje automático supervisado es Naïve Bayes.³⁴ Naïve Bayes aplica la regla de Bayes de que la categoría más probable de un documento es aquella que maximiza el producto de dos factores, $P(c)$ y $P(d|c)$, donde d es el documento y c es la categoría. El factor $P(c)$ se conoce como la probabilidad previa de la categoría. Consigna con qué frecuencia ocurre una categoría en los datos de entrenamiento. El factor $P(d|c)$ se conoce como la probabilidad. Consigna la probabilidad de que un documento d dada la categoría, cuando d puede representarse como un vector de palabras $d = x_1 + x_2 + x_3 + \dots + x_n$, donde n es el número total de palabras. Para cada palabra, el factor de probabilidad puede estimarse viendo el número de veces que la palabra aparece en esa categoría particular, como coeficiente de todas las palabras asociadas con esa categoría en los datos de entrenamiento. En la práctica, estas probabilidades se calculan llevando todos los textos dentro de una categoría particular a un único documento combinado de entrenamiento para la categoría y, posteriormente, contando las frecuencias relativas de w_i como un coeficiente del número general de palabras w en los datos de entrenamiento.³⁵

³⁴ Naïve Bayes es *ingenuo* en dos sentidos. Primero, parte del supuesto de una bolsa de palabras sencilla de que el orden de las palabras no importa, por lo que sólo considera la frecuencia de las palabras en un documento. Segundo, supone que la probabilidad de cada palabra que aparece en una categoría dada es independiente de la presencia de otras palabras, a pesar de que ya hemos señalado que es probable que se trate de un supuesto erróneo (Jurafsky y Manning, 2012).

³⁵ Por supuesto, es posible que una palabra asociada a una categoría no aparezca en los datos de entrenamiento. Supóngase que el banco X tiene sus oficinas principales en Londres. Pero que la palabra *Londres* no aparece en un documento de entrenamiento de 2,000 palabras clasificado como *banco X*. Como resultado, la probabilidad estimada de Londres dado el tema del banco X sería cero. Con el fin de evitar este resultado falso, podemos utilizar un procedimiento denominado alisamiento de Laplace agregando 1, que

Un ejemplo de investigación utilizando Naïve Bayes es el estudio de Moniz y Jong (2011) sobre los efectos de las minutas del CPM del Banco respecto a las expectativas de la tasa de interés futura. Los autores emplean Naïve Bayes en combinación con otras técnicas de minería de textos comentadas en este documento. Primero, buscaron las palabras en las páginas de Wikipedia sobre *banca central e inflación* y consignaron las palabras que se asocian³⁶ a *crecimiento económico, precios, tasas de interés y crédito bancario*. Estas palabras luego se utilizan como categorías en un modelo de Naïve Bayes asignándolas a oraciones en las minutas del CPM. Las categorías asignadas se utilizan posteriormente para construir un índice de sentimiento utilizando un método sencillo basado en diccionario.³⁷ Por último, el algoritmo de LDA se utiliza para encontrar palabras que pudieran actuar como intensificadores y atenuadores del sentimiento, por ejemplo, “aumento” y “moderado”, respectivamente. La combinación de métodos de Moniz y Jong destaca un aspecto importante de la minería de textos en la práctica: las técnicas supervisadas y no supervisadas suelen complementarse y emplearse en distintas etapas del proceso de minería de textos.

simplemente implica sumar 1 a la siguiente ecuación:

$$\hat{P}(w_i|c_j) = \frac{\text{cuenta}(w_i, c_j) + 1}{\sum_{w \in V} \text{cuenta}(w, c_j) + V}$$

donde w_i es una palabra determinada, c_j es una categoría determinada y V representa todas las palabras en el corpus.

³⁶ Los autores utilizan TextRank, un algoritmo de clasificación basado en gráficas para determinar grupos de palabras asociadas (Mihalcea y Tarau, 2004).

³⁷ Los autores utilizan el diccionario General Inquirer <<http://www.wjh.harvard.edu/~inquirer/>>, como ya se señaló.

El propósito de este documento ha sido demostrar el valor adicional que pueden obtener los bancos centrales si aplican las distintas técnicas de minería de textos, así como ejemplificar el uso de estas por quienes formulan las políticas y abordar los principales temas de investigación que interesan a los bancos centrales. Para concluir, deseamos destacar que la promesa de la minería de textos para los bancos centrales no es sólo hipotética sino que se ha comprobado. Por ejemplo, Kevin Warsh (2014) citó publicaciones sobre minería de textos (Schonhardt-Bailey, 2013; Hansen, McMahon y Prat, 2014) como influencias importantes de las recomendaciones finales de política que formaron parte de su análisis, a saber, que el Banco divulgara más información sobre sus deliberaciones. Warsh destacó que si bien “los estudios que buscan dar sentido a millones de palabras habladas” son “intimidantes e imperfectos”, la minería de textos “ha hecho grandes avances en nuestra comprensión” de los bancos centrales (Warsh, 2014). En términos más generales, este texto ha mostrado de qué manera la minería de textos puede ser una contribución útil al arsenal analítico de los bancos centrales y ayudarles a alcanzar sus objetivos de política.

BIBLIOGRAFÍA

- Aase, K. G. (2011), *Text Mining of News Articles for Stock Price Predictions*, tesis de maestría, Norwegian University of Science and Technology.
- Acosta, J. M. (2014), *FOMC Responses to Calls for Transparency: Evidence from the Minutes and Transcripts Using Latent Semantic Analysis*, tesis distinguida, Department of Economics, Stanford University, disponible en <<http://economics.stanford.edu/content/honors-thesis-2014>>.
- Apel, M., y M. Grimaldi (2012), *The Information Content of Central Bank Minutes*, Sveriges Riksbank Working Paper Series, núm. 261.
- Baker, S. R., N. Bloom y S. J. Davis (2013), *Measuring Economic Policy Uncertainty*, Chicago Booth Research Paper, núm. 13-02.
- Banco de Inglaterra (2014), “Inflation Report Q&A, 13th August 2014”, disponible en <<http://www.bankofengland.co.uk/publications/Documents/inflationreport/2014/conf130814.pdf>>.

- Banco de Inglaterra (2015), "One Bank Research Agenda Discussion Paper", disponible en <<http://www.bankofengland.co.uk/research/Documents/onebank/discussion.pdf>>
- Bara, J., A. Weale y A. Biquelet (2007), "Analysing Parliamentary Debate with Computer Assistance", *Swiss Political Science Review*, vol. 13, núm. 4, pp. 577-605.
- Bauer, M. (2000), "Qualitative Researching with Text, Image and Sound: A Practical Handbook", en M. W. Bauery G. Gaskell, *Classical Content Analysis: A Review*, Sage Publications, Londres, pp. 131-151.
- Bennani, H., y E. Farvaque (2014), "Speaking in Tongues? Diagnosing the Consistency of Central Banks' Official Communication", disponible en <http://www.econ.cam.ac.uk/eps2014/openconf/modules/request.php?module=oc_program&action=view.php&id=198>.
- Bholat, D. (2015), "Big Data and Central Banks", *Bank of England Quarterly Bulletin*, vol. 55, núm. 1, pp. 86-93.
- Blinder, A. S., M. Ehrmann, M. Fratzscher, J. de Haan y D. J. Jansen (2008), "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence", ECB Working Paper Series, núm. 898.
- Bollen, J., H. Mao y X. Zeng (2011), "Twitter Mood Predicts the Stock Market", *Journal of Computational Science*, vol. 2, núm. 1, pp. 1-8.
- Brugidou, M. (1998), "Epitaphes, l'image de Francois Mitterrand à travers l'analyse d'une question ouverte posée à sa mort", *Revue Française de Science Politique*, vol. 48, núm. 1, pp. 97-120.
- Brugidou, M. (2000), "Les discours de la revendication et de l'action dans les éditoriaux de la presse syndicale (1996-1998)", *Revue Française de Science Politique*, vol. 50, núm. 6, pp. 962-992.
- Brugidou, M. (2003), "Argumentation and Values: An Analysis of Ordinary Political Competence Via an Open-ended Question", *International Journal of Public Opinion Research*, vol. 15, núm. 4, pp. 413-430.
- Bulir, A., M. Cihak y D. J. Jansen (2014), *Does the Clarity of Inflation Reports Affect Volatility in Financial Markets?*, IMF Working Paper, núm. 14/175.
- Carney, M. (2013), "Crossing the Threshold to Recovery", discurso en el Banco de Inglaterra, 28 de agosto.
- Carney, M. (2015), "One Bank Research Agenda: Launch Conference", discurso en el Banco de Inglaterra, 25 de febrero.
- Deerwester, S., S. T. Dumais, G. W. Furnas, T. K. Landauer y R. Harshman (1990), "Indexing by Latent Semantic Analysis", *Journal of the American Society for Information Science*, vol. 41, núm. 6, pp. 391-407.
- Eckley, P. (2015), *Measuring Economic Uncertainty Using News-media Textual Data*, MPRA Paper, núm. 64874, disponible en <<http://mpra.ub.uni-muenchen.de/64874/>>.
- Gai, P., A. Haldane y S. Kapadia (2011), "Complexity, Concentration and Contagion", *Journal of Monetary Economics*, vol. 58, núm. 5, pp. 453-470.
- Gentzkow, M., y J. M. Shapiro (2010), "What Drives Media Slant? Evidence from U.S. Daily Newspapers", *Econometrica*, vol. 78, núm. 1, pp. 35-71.
- Giles, C. (2015), "Bank of England Mark Carney Expands Research Agenda", *Financial Times*, 25 de febrero de 2015.
- Grimmer, J., y B. M. Stewart (2013), "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts", *Political Analysis*, vol. 21, pp. 267-297.
- Guerin-Pace, F. (1998), "Textual Analysis. An Exploratory Tool for the Social Sciences", *Population: An English Selection*, número especial de *New Methodological Approaches in the Social Sciences*, vol. 10, núm. 1, pp. 73-95.
- Haldane, A. (2012), "The Dog and the Frisbee", discurso en el Banco de Inglaterra, 31 de agosto.
- Haldane, A. (2015), "The Promise of New Data and Advanced Analytics", discurso en el Banco de Inglaterra, 25 de febrero.

- Hansen, S., M. McMahon y A. Prat (2014), "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach", CEP Discussion Paper, Centre for Economic Performance, LSE, núm. DP1276.
- Hendry, S. (2012), *Central Bank Communication or the Media's Interpretation: What Moves Markets?*, Bank of Canada Working Paper, núm. 2012-9.
- Hendry, S., y A. Madeley (2010), *Text Mining and the Information Content of Bank of Canada Communications*, Bank of Canada Working Paper, núm. 2010-31.
- Hoberg, G., y G. M. Phillips (2010), "Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis", *The Review of Financial Studies*, vol. 23, núm. 10, pp. 3773-3811.
- Humpherys, S., K. C. Moffitt, M. B. Burns, J. K. Burgoon y W. F. Felix (2011), "Identification of Fraudulent Financial Statements Using Linguistic Credibility Analysis", *Decision Support Systems*, vol. 50, pp. 585-594.
- Jansen, D. J., y J. de Haan (2010), *An Assessment of the Consistency of ECB Communication Using Wordscores*, Nederlandsche Bank Working Paper, núm. 259.
- Jenny, J. (1997), "Techniques and Formalized Practices for Content and Discourse Analysis in Contemporary French Sociological Research", *Bulletin de Méthodologie Sociologique*, vol. 54, pp. 64-112.
- Jurafsky, D., y C. Manning (2012), "Natural Language Processing", discurso en línea, disponible en <<https://www.coursera.org/course/nlp>>.
- Kloptchenko, A., C. Magnusson, B. Back, A. Visa y H. Vanharanta (2004), "Mining Textual Contents of Financial Reports", *The International Journal of Digital Accounting Research*, vol. 4, núm. 7, pp. 1-29.
- Lahlou, S. (1996), "A Method to Extract Social Representations from Linguistic Corpora", *Japanese Journal of Experimental Social Psychology*, vol. 35, núm. 3, pp. 278-291.
- Li, W. P., P. Azar, D. Larochelle, P. Hill y A. W. Lo (2015), "Law is Code: A Software Engineering Approach to Analyzing the United States Code", *Journal of Business & Technology Law*, vol. 10, núm. 2, pp. 297-374.
- Loughran, T., y B. McDonald (2011), "When Is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks", *Journal of Finance*, vol. 66, núm. 1, pp. 35-65.
- Masawi, B., S. Bhattacharya y T. Boulter (2014), "The Power of Words: A Content Analytical Approach Examining Whether Central Bank Speeches Become Financial News", *Journal of Information Science*, vol. 40, núm. 2, pp. 198-210.
- McLaren, N., y R. Shanbhogue (2011), "Using Internet Search Data as Economic Indicators", *Bank of England Quarterly Bulletin*, vol. 51, núm. 2.
- Mihalcea, R., y P. Tarau (2004), "TextRank: Bringing Order into Texts", en Association for Computational Linguistics, *Proceedings of EMNLP 2004*, pp. 404-411.
- Moniz, A., y F. de Jong (2011), "Predicting the Impact of Central Bank Communications on Financial Market Investors' Interest Rate Expectations", *Lecture Notes in Computer Science*, vol. 8798, pp. 144-155.
- Nergues, A., J. Lee, P. Groenewegen y I. Hellesten (2014), "The Shifting Discourse of the European Central Bank: Exploring Structural Space in Semantic Networks", en *Signal-image Technology and Internet-Based Systems (SITIS), 2014 Tenth International Conference*, pp. 447, 455, 23-27.
- Nivre, J., J. Hall y J. Nilsson (2006), "MaltParser: A Data-driven Parser-generator for Dependency Parsing", en *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC2006)*, pp. 2216-2219.
- Noel-Jorand, M. C., M. Reinert, M. Bonnon y P. Therme (1995), "Discourse Analysis and Psychological Adaptation to High Altitude Hypoxia", *Stress Medicine*, vol. 11, pp. 27-39.

- Noel-Jorand, M. C., M. Reinert, S. Giudicelli y D. Dassa (1997), "A New Approach to Discourse Analysis in Psychiatry, Applied to Schizophrenic Patient Speech", *Schizophrenia Research*, vol. 25, pp. 183-198.
- Noel-Jorand, M. C., M. Reinert, S. Giudicelli y D. Dassa (2004), "Schizophrenia: The Quest for a Minimum Sense of Identity to Ward Off Delusional Psychosis", *The Canadian Journal of Psychiatry*, vol. 49, núm. 6, pp. 394-398.
- Nyman, R., D. Gregory, S. Kapadia, P. Ormerod, D. Tuckett y R. Smith (2015), *News and Narratives in Financial Systems: Exploiting Big Data for Systemic Risk Assessment*, material mimeografiado.
- O'Connor, B., R. Balasubramanian, B. R. Routledge y N. A. Smith (2010), "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series", *Proceedings of the 4th International Conference on Weblogs and Social Media*, pp. 122-129.
- Office for National Statistics (2007), *UK Standard Industrial Classification of Economic Activities 2007*.
- Peart, J. (2013), "How Do Appointment Processes Affect the Policy Outputs of Monetary Policy Committees?", disponible en <<http://johnpeart.org/dissertation/>>.
- Reinert, M. (1983), "Une methode de classification descendante hierarchique: application a l'analyse lexicale par contexte", *Les Cahiers de l'Analyse des Donnees*, vol. 8, núm. 2, pp. 187-198.
- Reinert, M. (1987), "Classification descendante hiérarchique et analyse lexicale par contexte: application au corpus des poésies d'Arthur Rimbaud", *Bulletin de Méthodologie Sociologique*, vol. 13, pp. 53-90.
- Reinert, M. (1990), "Alceste. Une methodologie d'analyse des donnees textuelles et une application: Aurelia de Gerard de Nerval", *Bulletin de Methodologie Sociologique*, vol. 26, pp. 24-54.
- Reinert, M. (1993), "Les 'mondes lexicaux' et leur 'logique' à travers l'analyse statistique d'un corpus de récits de cauchemars", *Langage et Société*, vol. 66, pp. 5-39.
- Reinert, M. (1998a), "Quel objet pour une analyse statistique du discours? Quelques réflexions à propos de la réponse Alceste", *Proceedings of the 4th JADT (Journées d'Analyse des Données Textuelles)*, Universidad de Niza, JADT.
- Reinert, M. (1998b), *Alceste Users' Manual (English version)*, Image, Toulouse.
- Reinert, M. (2003), "Le rôle de la répétition dans la représentation du sens et son approche statistique dans la méthode Alceste", *Semiotica*, vol. 147, núm. 1-4, pp. 389-420.
- Ronnqvist, S., y P. Sarlin (2012), "From Text to Bank Interrelation Maps", *Computational Intelligence for Financial Engineering & Economics*, 2104 IEEE Conference, pp. 48-54.
- Rosa, C., y G. Verga (2006), "On the Consistency and Effectiveness of Central Bank Communication: Evidence from the ECB", *European Journal of Political Economy*, vol. 23, núm. 1, pp. 146-175.
- Schonhardt-Bailey, C (2005), "Measuring Ideas More Effectively: An Analysis of Bush and Kerry's National Security Speeches", *Political Science and Politics*, vol. 38, núm. 3, pp. 701-711.
- Schonhardt-Bailey, C. (2006), *From the Corn Laws to Free Trade: Interests, Ideas and Institutions in Historical Perspective*, MIT Press, Cambridge, MA.
- Schonhardt-Bailey, C. (2008), "The Congressional Debate on Partial-birth Abortion: Constitutional Gravitas and Moral Passion", *British Journal of Political Science*, vol. 38, pp. 383-410.
- Schonhardt-Bailey, C., E. Yager y S. Lahlou (2012), "Yes, Ronald Reagan's Rhetoric Was Unique – But Statistically, How Unique?", *Presidential Studies Quarterly*, vol. 42, núm. 3, pp. 482-513.
- Schonhardt-Bailey, C. (2013), *Deliberating American Policy: A Textual Analysis*, MIT Press, Cambridge, MA.
- Siklos, P. L. (2013), *The Global Financial Crisis and the Language of Central Banking: Central*

Bank Guidance in Good Times and in Bad,
CAMA Working Paper, núm. 58/2013.

Tetlock, P. C. (2007), "Giving Content to Investor Sentiment: The Role of Media in the Stock Market", *The Journal of Finance*, Vol. LXII, núm. 3.

Upshall, M. (2014), "Text Mining: Using Search to Provide Solutions", *Business Information Review*, vol. 31, pp. 91-99.

Vallès, D. W., y C. Schonhardt-Bailey (2015), "Forward Guidance as Central Bank Discourse: MPC Minutes and Speeches under King and Carney", presentado en el Political Leadership and Economic Crisis Symposium, Yale University (febrero).

Warsh, K. (2014), "Transparency and the Bank of England's Monetary Policy Committee", reseña de Kevin Warsh.

Weale, A., A. Bicquelet y B. Judith (2012), "Debating Abortion, Deliberative Reciprocity and Parliamentary Advocacy", *Political Studies*, vol. 60, pp. 643-667.

INICIATIVA DEL G20 SOBRE DEFICIENCIAS DE LOS DATOS II: FRENTE AL DESAFÍO DE POLÍTICA

Robert Heath
Evrin Bese Goksu

I. INTRODUCCIÓN

Una vieja lección ampliamente conocida es que “los buenos datos y el buen análisis son los elementos vitales para una eficaz supervisión y respuesta de políticas, tanto en escala nacional como internacional”.¹ Ciertamente, la información confiable, completa y oportuna es esencial para evaluar los riesgos y las vulnerabilidades que enfrentan las economías, ya que la formulación de la política económica depende de una evaluación correcta de estos riesgos y vulnerabilidades.

En 2007/2008 los problemas en los sistemas financieros en varias economías avanzadas, entre ellas Estados Unidos de América (EUA), se desbordaron a través de las fronteras y afectaron al resto del mundo. Como el sector financiero estuvo en el centro de la crisis, las economías del G20 respaldaron una cantidad de acciones para la reforma del marco regulatorio del sector financiero. Aunque la falta de datos no fue la razón central de la crisis, hubiese sido posible detectar la acumulación de riesgos si se hubiese contado con los datos correctos en el momento oportuno. A estos efectos, la identificación y la manera de abordar las deficiencias de información fueron algunas de las tareas de la reforma del sector financiero, las cuales dieron como resultado la Iniciativa sobre Deficiencias de los Datos del G20 (Data Gap Initiative, o DGI, por sus siglas en inglés). Ya que la Iniciativa está a punto de comenzar su segunda fase, el presente documento explica cómo la DGI está satisfaciendo las necesidades de política.

¹ Ver el informe del Consejo de Estabilidad Financiera (CEF) del FMI, *Financial Crisis and Information Gaps*, octubre de 2009.

Traduce y publica el CEMLA con la debida autorización, el artículo original *G20 Data Gaps Initiative II: Meeting the Policy Challenge*, IMF Working Paper, núm. WP/16/43. Las opiniones expresadas en este documento son exclusivas de los autores y no necesariamente representan la política del FMI o sus ideas. Cualquier error derivado de la traducción es responsabilidad del CEMLA.

2. RESPUESTA A LAS NECESIDADES DE POLÍTICA: LA INICIATIVA SOBRE DEFICIENCIAS DE LOS DATOS

2.1 CÓMO LA EVOLUCIÓN DEL PENSAMIENTO ECONÓMICO A LO LARGO DEL TIEMPO AFECTÓ LAS ESTADÍSTICAS

No es nuevo el reconocimiento de la necesidad de consolidar los datos económicos y financieros luego de la crisis. A medida que los sistemas económicos y financieros evolucionan como consecuencia de acontecimientos del mercado y de la innovación financiera, cambian las necesidades de información. Si nos remontamos a la historia, los eventos de crisis siempre han actuado como desencadenantes para cuestionar la naturaleza, la calidad y la disponibilidad de los datos necesarios para elaborar políticas económicas.

La Gran Depresión de los años treinta es un buen ejemplo de avances fundamentales de las estadísticas económicas. A medida que los responsables de la política comenzaron a gestionar de manera más activa la economía, y en especial la demanda agregada, el foco intelectual y de política pasó a estar concentrado en los factores de oferta y de demanda en la economía y en las transacciones más que en los acervos. El resultado fue la creación a fines de los años cuarenta del Sistema de Cuentas Nacionales, que hasta hoy es el marco de referencia predominante para la estadística macroeconómica; por la misma época se publicó el primer *Manual de balanza de pagos* del FMI.

La tendencia a la liberación del capital que se inició en los ochenta trajo nuevas oportunidades para la inversión, pero también trajo aparejado nuevos riesgos y vulnerabilidades, tanto en el ámbito nacional como en el internacional, lo cual llevó a una mayor atención de la política en la estabilidad financiera.² Estos acontecimientos han orillado a replantear las políticas macroprudenciales y monetarias,³ como así también de los marcos estadísticos.

Durante la crisis mexicana de 1994-1995, fueron elementos clave los flujos de capital internacional y la falta de información relevante. El FMI respondió con el establecimiento de dos estándares fundamentales para la divulgación de un conjunto fundamental de datos económicos y financieros centrales: las Normas Especiales para la Divulgación de Datos (NEDD) y el Sistema General de Divulgación de Datos (SGDD). Las NEDD tenían como meta a los países con acceso a los mercados internacionales de capital

² Ver *What Has Capital Flow Liberalization Meant for Economic and Financial Statistics?*, IMF Working Paper, núm. WP/15/88.

³ Por ejemplo, ver la serie del seminario *Rethinking Macro Policy III* disponible en <<http://www.imf.org/external/np/seminars/eng/2015/macro3/>>.

mientras que el SGDD se enfocó en los países que necesitaban desarrollar sus sistemas estadísticos.

En 1997-1998 la crisis asiática reveló la necesidad de mejorar la información sobre las reservas y sobre las actividades vinculadas a estas, ya que se consideró que los contratos de ventas a plazo de divisas por parte del Banco de Tailandia habían enmascarado la verdadera presión sobre las reservas internacionales. Como resultado de esta situación se elaboró una *plantilla de reservas* y se reforzaron las NEED al agregar datos sobre requerimientos de reservas y liquidez en divisas.

Debido a los desequilibrios mundiales y a las consiguientes discrepancias en los flujos de ingresos en escala mundial, a fines de 1997 se emprendió el primer Estudio Coordinado sobre Inversión de Cartera (ECIC) del FMI, con el fin de mejorar las estadísticas sobre los acervos de valores de cartera en forma de capital accionario y títulos de deuda a corto y largo plazos. El fortalecimiento de las estadísticas bancarias internacionales del Banco de Pagos Internacionales (BPI) y la adopción creciente del *Manual de la balanza de pagos* del FMI han continuado durante los últimos dos decenios.

2.2 LA CRISIS FINANCIERA MUNDIAL DE 2007-2008 Y LA DGI

La crisis financiera que se desató en 2007 por los problemas en el mercado de las hipotecas de alto riesgo o *subprime* en EUA se expandió al resto del mundo hasta convertirse en la crisis mundial más grave desde la Gran Depresión. Una diferencia entre la crisis financiera mundial y las anteriores crisis de posguerra fue que la crisis afectó la médula misma del sistema financiero mundial diseminándose por toda la economía mundial.

Fue necesario un esfuerzo mundial para la recuperación. Como parte de la respuesta mundial ante esta crisis, en octubre de 2009, los ministros de Finanzas y gobernadores de los bancos centrales del G20 firmaron la DGI liderada por la Secretaría del Consejo de Estabilidad Financiera (CEF) y el personal del FMI. La DGI se emitió como una iniciativa

predominante de 20 recomendaciones para zanjar las deficiencias de información reveladas por la crisis financiera mundial.

Desde que se emprendió, se ha logrado un progreso significativo para zanjar estas deficiencias.⁴ Ante este avance, en septiembre de 2015, en su sexto año de aplicación, los ministros de Finanzas y los gobernadores del G20 cerraron la primera acción (DGI-1) y abrieron una segunda acción de la DGI (DGI-2).

El éxito de la DGI se atribuye principalmente al sólido respaldo de política, un atributo común a las economías del G20, y a la cooperación estrecha entre las partes relevantes. El Inter-Agency Group on Economic and Financial Statistics (IAG)⁵ ha actuado como el facilitador mundial y coordinador del ejercicio, colaborando con otros grupos e iniciativas. El personal del FMI ha estado vigilando año con año la aplicación de las recomendaciones de la DGI realizada por las economías de G20 e informando, junto con la Secretaría del CEF, el progreso logrado por los ministerios y bancos del G20. Desde 2009, se han entregado seis informes de progreso a estos.⁶

En 2009 muchas de las recomendaciones se escribieron como aspiraciones, ya que eran poco claras las implicaciones de la crisis para las políticas regulatoria y financiera. Se esbozaron luego de consultas exhaustivas con usuarios y compiladores, entre ellas una conferencia de usuarios de julio de 2009,⁷ y se estructuraron en torno a cuatro temas: acumulación de riesgo en el sector financiero, vulnerabilidades financieras transfronterizas,

⁴ Ver *Sixth Progress Report on the Implementation of the G-20 Data Gaps Initiative*, septiembre de 2015 <<http://www.imf.org/external/np/g20/pdf/2015/6thprogressrep.pdf>>.

⁵ Los miembros de la IAG son el BPI, el BCE, el Eurostat, el FMI (director), la OCDE, las Naciones Unidas y el Banco Mundial. Se invita al CEF a participar en temas en los cuales está directamente involucrado.

⁶ Los informes de progreso de la Iniciativa sobre Deficiencias de los Datos del G20 están disponibles en <<http://www.imf.org/external/ns/cs.aspx?id=290>>.

⁷ Los artículos de la conferencia de usuarios de 2009 están disponibles en <<http://www.imf.org/external/np/seminars/eng/2009/usersconf/index.htm>>.

vulnerabilidad de las economías nacionales ante los choques y comunicación de estadísticas oficiales.

Con el paso del tiempo, se hicieron cada vez más claras las implicaciones de la crisis en las políticas regulatoria y macroprudencial, y por lo tanto las necesidades de datos. Como reflejo de esta situación, las recomendaciones de la DGI-2 se enfocan en bases de datos que respaldan la estabilidad del sistema financiero tanto en escala nacional como internacional. Sin embargo, como las 20 recomendaciones de la DGI-1 han superado la prueba del tiempo, la DGI-2 representa una evolución y no un replanteo del proyecto DGI. La DGI-2 tiene como meta reforzar y consolidar el progreso logrado en la DGI-1, alcanzar el potencial para la provisión de los datos plasmados en la Iniciativa y promover estadísticas de alta calidad para el uso de política (ver cuadro 1). Las recomendaciones de DGI-2 se precisan en el anexo 1.

La intención de este documento de trabajo es demostrar cómo la DGI-2 es integral para el cumplimiento de las necesidades emergentes de política, tanto a nivel regulatorio como microfinanciero. A estos efectos, las recomendaciones de la DGI-2 son más específicas que las de la DGI-1, y algunas de ellas se identifican como prioridades mundiales (ver cuadro 1) a partir de las consultas con usuarios y compiladores en 2015. Se les solicita a las economías del G20 que se comprometan en planes de acción acordados en escala nacional que toman en cuenta las circunstancias internas pero que se basan en las metas fijadas para cada recomendación. La meta es avanzar en la agenda estadística acordada con las economías del G20 y respaldadas por ministros y gobernadores del G20 en el entorno mundial. Esta agenda se elaboró para ayudar a que los sistemas financieros nacionales e internacionales sean más estables en un mundo con creciente interconectividad financiera. Sigue siendo relevante para la DGI-2 un documento de trabajo anterior que presenta la justificación analítica para las recomendaciones de la DGI-1.⁸

⁸ Ver *Why Are the G20 Data Gaps Initiative and the NEDD Plus*

2.3 EVOLUCIÓN DE LAS NECESIDADES DE POLÍTICA DESPUÉS DE LA CRISIS FINANCIERA MUNDIAL

2.3.1 Agenda de la reforma regulatoria

Después de la crisis financiera mundial, en 2008, los líderes del G20, en su reunión en Washington,⁹ se comprometió a aplicar una reforma fundamental del sistema financiero mundial para fortalecer los mercados financieros y los regímenes regulatorios para evitar crisis futuras.¹⁰ Como parte de la reforma de la agenda, se estableció el CEF en abril de 2009 como el organismo sucesor del Foro de Estabilidad Financiera y comenzó a trabajar como centro neurálgico para llevar adelante el programa de reforma financiera tal como se elaboró con los organismos relevantes. Se establecieron las obligaciones de los miembros del CEF que incluían revisiones periódicas por parte de los pares, usando, entre otros aportes, los informes del programa de evaluación del sector financiero del FMI/Banco Mundial. Los líderes del G20 señalaron la importancia de los esfuerzos mundiales para aplicar la reforma regulatoria mundial a fin de protegerse contra acontecimientos mundiales, regionales y transfronterizos que puedan afectar la estabilidad financiera internacional.

Los componentes de la agenda de reforma regulatoria del G20 se complementan entre sí con la finalidad de fortalecer el sistema financiero internacional. La DGI ha sido un elemento importante de esta agenda, ya que los temas de la agenda de reforma regulatoria requieren es su mayoría de mejores datos. La recopilación de datos sobre las

Relevant for Financial Stability Analysis, IMF Working Paper, núm. 13/6.

⁹ “Declaration Summit on Financial Markets and the World Economy, November 15, 2008”, <https://g20.org/wp-content/uploads/2014/12/Washington_Declaration_0.pdf>.

¹⁰ Los líderes del G20 continúa reafirmando la importancia de su compromiso. Por ejemplo, en la cumbre de Antalya en noviembre de 2015, los líderes declararon que “Mirando hacia adelante, nos comprometemos a la ejecución completa y congruente del marco regulatorio financiero mundial en línea con los plazos acordados...”

Cuadro 1

RECOMENDACIONES DE LA DGI

<i>Recomendaciones de la DGI-I</i>	<i>Recomendaciones de la DGI-II</i>
<i>I.1 Mandato</i>	<i>II.1 Mandato</i>
<i>Acumulación de riesgo en el sector financiero</i>	<i>Vigilancia de los riesgos en el sector financiero</i>
I.2: Indicadores de solidez financiera (ISF) 🟡 I.3: Riesgo de evento excepcional 🟢 I.4: Apalancamiento agregado y descalces de vencimientos 🟢 I.5: Permutas de incumplimiento crediticio (CDS) 🟡 I.6: Productos estructurados 🟢 I.7: Datos de valores 🟢	II.2: Indicadores de solidez financiera 🔴 II.3: Medidas de concentración y distribución de los ISF II.4: Datos de las instituciones financieras de importancia sistémica mundial (G-SIFI) II.5: Banca paralela II.6: Derivados II.7: Estadísticas de valores 🔴
<i>Vinculaciones financieras transfronterizas</i>	<i>Vulnerabilidades, interconexiones y diseminaciones</i>
I.8 y I.9: Datos de las instituciones financieras de importancia sistémica mundial (G-SIFI por sus siglas en inglés) 🟢 I.10 y I.11: Estudio Coordinado sobre Inversión de Cartera (ECIC) 🟡 y participación 🟡 y mejora 🟢 en las estadísticas bancarias internacionales (EBI) I.12: Posición de inversión internacional (PII) 🟡 I.13 y I.14: Exposiciones transfronterizas de las corporaciones financieras y no financieras 🟢	II.8: Cuentas sectoriales 🔴 II.9: Información de distribución del sector de hogares II.10: Posición de inversión internacional (PII) II.11: Estadísticas bancarias internacionales (EBI) II.12: Estudio Coordinado sobre Inversión de Cartera (ECIC) II.13: Estudio Coordinado de Inversión Directa (ECID) II.14: Exposiciones transfronterizas de corporaciones no bancarias II.15: Estadísticas financieras del gobierno 🔴 II.16: Base de datos de deuda del sector público (PSDS, por sus siglas en inglés) 🔴 II.17: Precios de inmuebles residenciales (PIR) II.18: Precios de inmuebles comerciales (PIC)
<i>Vulnerabilidad de las economías nacionales a los choques</i>	
I.15: Cuentas sectoriales 🟢 I.16: Información de distribución 🟢 I.17: Estadísticas financieras del gobierno (GFS, por sus siglas en inglés) 🟡 I.18: Deuda del sector público 🟡 I.19: Precios de bienes raíces 🟡	
<i>Comunicación de estadísticas oficiales</i>	<i>Comunicación de estadísticas oficiales</i>
I.20 Principales indicadores mundiales 🟢	II.19: Cooperación y comunicación de datos internacionales II.20: Promoción del intercambio de datos
🟢 Recomendaciones que se completan con base en las metas introducidas en 2014. 🟡 Recomendaciones donde se hizo un progreso significativo y están próximas a concretarse, dependiendo de la participación de todo el G20. 🟠 Recomendaciones donde el progreso fue lento. 🔴 Indica áreas prioritarias identificadas por las economías del G20 y por las instituciones internacionales en 2015.	

exposiciones de los bancos de importancia sistémica mundial (G-SIB, por sus siglas en inglés) y de sus dependencias de financiamiento es uno de los pasos para abordar el problema de los bancos *demasiado grandes para quebrar* al reducir la probabilidad y el efecto de la quiebra de un G-SIB. El CEF trabaja en la fijación de estándares y procesos para la recopilación de datos en el mundo. La agregación de las transacciones financieras con valores tiene como meta transparentar la titularización con el objetivo central de reducir los riesgos vinculados al sistema bancario paralelo. Los mercados de derivados OTC o extrabursátiles, incluidos los mercados de permutas de incumplimiento crediticio (CDF), pasaron a ser objeto de mayor escrutinio después de la crisis mundial con el fin de hacer más seguros los mercados de derivados. La DGI respaldó esta meta al mejorar la información sobre los mercados de CDS. Hay otras iniciativas del G20 que se relacionan estrechamente con la DGI tales como el trabajo del CEF en el fortalecimiento de la supervisión y regulación de la banca paralela, y el trabajo sobre los identificadores de entidades legales mundiales (LEI, por sus siglas en inglés)¹¹ que contribuyen a la solidez de las estructuras de datos con un enfoque más micro. Las cambiantes reformas regulatorias en escala mundial, en particular la aplicación de Basilea III, se tomaron también en cuenta para la redacción de la DGI.

2.3.2 Agenda de supervisión

Se destacó también la importancia de cerrar las deficiencias de datos que obstaculizan la supervisión de los sistemas financieros en el Triennial Surveillance Review (TSR) de 2014 del FMI.¹² El TSR de 2014 destacaba que debido a la creciente

¹¹ Un sistema de LEI mundial identificaría únicamente las partes involucradas en las transacciones financieras.

¹² Los documentos que han contribuido a la revisión, entre ellos el documento general, están disponibles en <<http://www.imf.org/external/np/spr/triennial/2014/>>. El TSR se realizó con base en consultas exhaustivas entre los países miembros del FMI, la academia y el sector privado.

interconectividad entre fronteras, los choques del mercado financiero continuarán teniendo importantes diseminaciones mediante los flujos de capital y de los cambios en las posiciones de riesgo. También, seguirán surgiendo nuevas dimensiones de interconectividad como las potenciales diseminaciones de corto plazo que generen las reformas regulatorias financieras. A estos efectos, el TSR recomendó mejorar la información sobre los balances generales y enriquecer los datos sobre flujos de fondos.

El FMI ha ajustado su supervisión con el objetivo de que se base más en el riesgo. Para ello, el Plan de Acción de la Directora Gerente del FMI para fortalecer la supervisión después del TSR 2014¹³ destacó que el FMI revivirá y adaptará el enfoque de balance general (BSA, por sus siglas en inglés) con el fin de facilitar el análisis más profundo de los efectos de los choques y su transmisión entre sectores; posiblemente, esto sea el inicio para que los flujos de fondos mundiales reflejen mejor la interconexión mundial (recuadro 1). Este trabajo requiere de datos obtenidos de la DGI ya que ayudará a respaldar el trabajo macrofinanciero del FMI incluido en los ejercicios e informes clave (por ejemplo, el ejercicio de advertencia temprana, el programa de evaluación del sector financiero y el informe sobre la estabilidad financiera mundial).

2.3.3 El proyecto de la DGI

El proyecto de la DGI ha permitido que una amplia gama de necesidades de los usuarios se incorpore en la elaboración de estadísticas económicas y financieras. En 2015, como parte del proceso de trabajo de la DGI, se llevaron a cabo consultas exhaustivas¹⁴ que incluyeron la información del Segundo Foro de Estadísticas del FMI; el mismo constituye un espacio anual de alcance mundial donde los usuarios de datos, los proveedores de

¹³ Ver <<http://www.imf.org/external/np/pp/eng/2014/112114.pdf>>.

¹⁴ Se realizaron cuatro conferencias regionales así como reuniones con los participantes del sector privado.

RECUADRO 1

FLUJOS DE FONDOS MUNDIALES

Mediante el uso de estándares estadísticos acordados internacionalmente, los datos sobre las exposiciones financieras transfronterizas (EBI, ECIC, y la Estudio Coordinado de Inversión Directa, ECID) pueden relacionarse con los datos de las cuentas sectoriales nacionales para construir una imagen integral de las interconexiones financieras nacionales y transfronterizas, vinculándolas nuevamente con la economía real por medio de las cuentas sectoriales. Este trabajo se conoce como Flujos de Fondos Mundiales (GFF, por sus siglas en inglés).¹ El proyecto GFF está orientado principalmente a la elaboración de una matriz que identifica las interconexiones entre los sectores nacionales con los países contraparte (y posiblemente entre los sectores de los países contraparte) para crear un cuadro de las exposiciones financieras bilaterales y

para respaldar el análisis de las fuentes potenciales de contagio.

El concepto de los GFF se esbozó por primera vez en el segundo informe de progreso de la Iniciativa sobre Deficiencias de los Datos del G20 y comenzó en 2013 como parte de una iniciativa más amplia del FMI con el fin de fortalecer el análisis de la interconexión a través de las fronteras, de los flujos mundiales de liquidez y de las interdependencias financieras mundiales. A más largo plazo, la matriz de GFF tiene como meta respaldar la vigilancia regular de las posiciones financieras transfronterizas bilaterales mediante un marco que destaca los riesgos para la estabilidad nacional e internacional. Los funcionarios del FMI están trabajando en la elaboración de una matriz de GFF que comienza con las economías más grandes del mundo.

¹ Ver "Mapping the Shadow Banking System through a Global Flow of Funds Analysis", <<http://www.imf.org/external/pubs/ft/wp/2014/wp1410.pdf>>

datos y los responsables de políticas económicas pueden reunirse a debatir las necesidades emergentes de información estadística para comunicar a los responsables de políticas económicas.¹⁵

Las consultas con los usuarios indicaron la necesidad de garantizar la amplitud e integridad de los datos que respaldan el análisis de las interconexiones de las economías. Se subrayó la importancia del enfoque de balance general para comprender las interconexiones sectoriales con la

economía nacional. Se convino en que eran esenciales los datos clave para evaluar la sustentabilidad fiscal, pero también se destacaron los retos de su aplicación. Como resultado de esto, se identificaron áreas clave del enfoque que son comunes para las economías del G20: divulgar indicadores compatibles y comparables de solidez financiera; asegurar la recolección regular de las EBI y ECIC; proporcionar estadísticas coherentes de títulos valores; mejorar la disponibilidad de los datos de cuentas sectoriales, y divulgar datos generales de las operaciones y deudas de los gobiernos de manera puntual y comparable.

El Segundo Foro de Estadísticas del FMI con el tema principal de Estadísticas para la Elaboración

¹⁵ El Segundo Foro de Estadísticas del FMI se llevó a cabo en las oficinas centrales del FMI en Washington, del 18 al 19 de noviembre de 2014. Las memorias del foro están disponibles en <<http://www.imf.org/external/np/seminars/eng/2014/statsforum/>>.

de Políticas: Identificación de Vulnerabilidades Macroeconómicas y Financieras, destacó la importancia de la calidad y comparabilidad de los datos; la necesidad de vigilar las interconexiones y por ende la importancia de los datos de las cuentas sectoriales, de los balances generales y de las posiciones de inversión internacionales (PII); la necesidad de mejorar la información sobre las corporaciones no financieras, el sector de hogares y los mercados de bienes raíces.

En consonancia con estos sucesos, la DGI-2 ha adoptado un enfoque en: 1) la vigilancia del riesgo en el sector financiero y 2) las vulnerabilidades, interconexiones y diseminaciones. Como se ilustra en la gráfica 1, las recomendaciones en la DGI-2 pueden agruparse principalmente en dos grandes títulos generales. Además, las recomendaciones se presentan como un paquete coherente cuya aplicación genera externalidades positivas tanto para la compilación de datos como para su análisis. La categoría de vulnerabilidades, interconexiones y diseminaciones se basa en el sistema de cuentas nacionales dominante mientras que las recomendaciones relativas a la vigilancia del riesgo en el sistema financiero, que comprende a las instituciones financieras y a los mercados financieros, y a la banca paralela que abarca tanto a las instituciones como a los mercados, como se explicará más adelante. Todas las recomendaciones cobran sentido para brindar un panorama general de la economía y del sector financiero.

2.4 ¿CÓMO AYUDA LA DGI A VIGILAR LOS RIESGOS EN EL SECTOR FINANCIERO?

2.4.1 Evaluación de la solidez del sistema bancario

Desde hace tiempo se reconoce que los microanálisis de las instituciones financieras necesitan complementarse con un enfoque macro. A estos efectos, se crearon los indicadores de solidez financiera (ISF) a principios de este siglo y, aunque con visión retrospectiva, son un importante

componente del marco macroprudencial para vigilar y evaluar la salud y solidez del sector financiero en general.¹⁶ A la fecha, el foco principal de los ISF ha sido el sector bancario con indicadores adicionales para los clientes de los bancos así como para los mercados en los cuales operan.

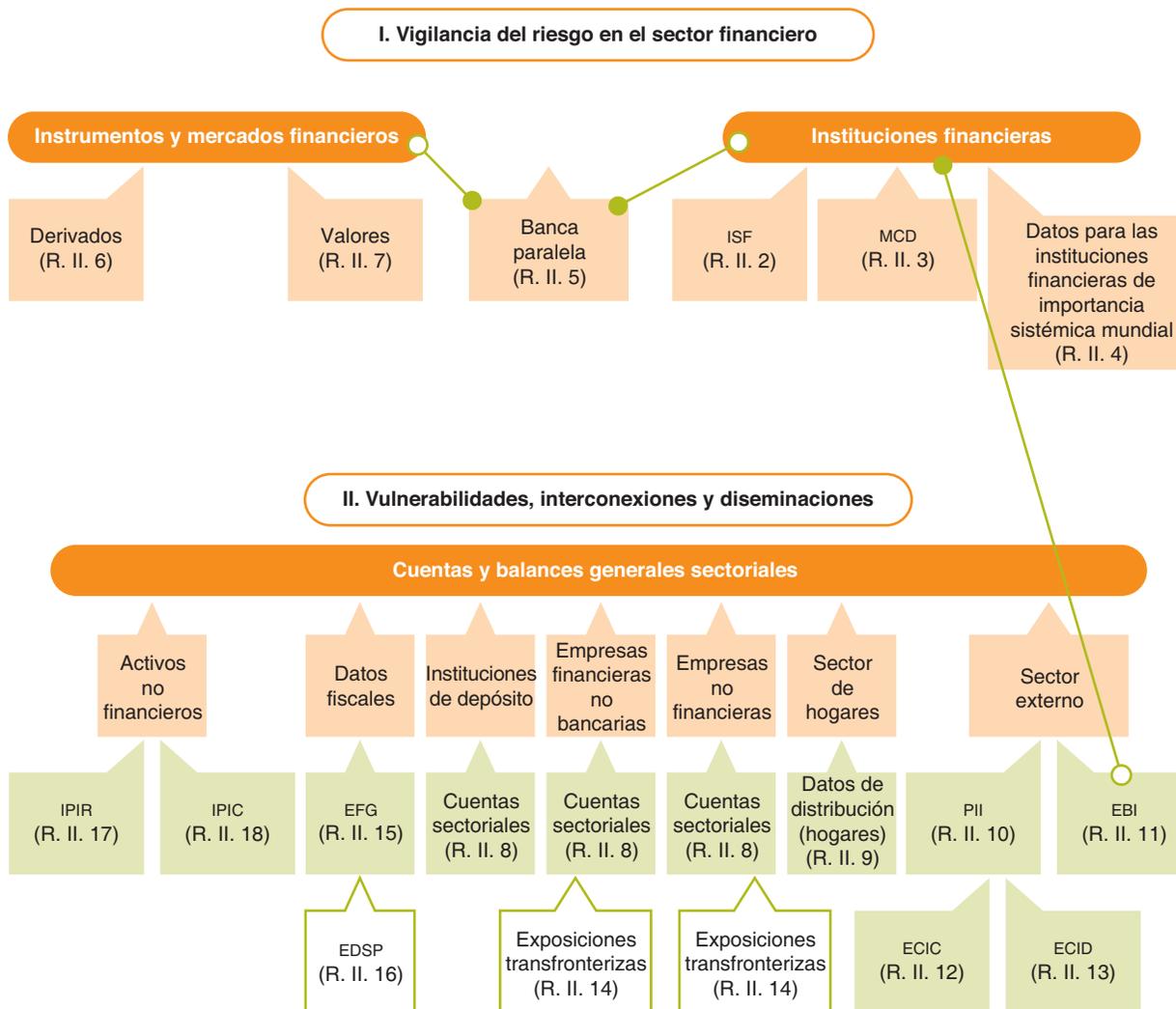
A partir de las consultas con compiladores y usuarios realizadas en 2015, queda claro que los ISF son cada vez más usados por las autoridades nacionales para establecer los parámetros nacionales, para realizar análisis entre países y para generar indicadores de alerta temprana. Tales análisis contribuyen con los informes de estabilidad financiera para nutrir la elaboración de la política económica. Además, el FMI incluye los datos de los ISF en los informes de consulta del artículo IV de las economías individuales y en el anexo estadístico del informe sobre la estabilidad financiera mundial. Este enfoque vinculado a las políticas de los ISF tanto en escala nacional como internacional ha ayudado a fomentar un incremento significativo en la cantidad de países que proporcionaron sus indicadores de solidez financiera al FMI durante la DGI-1. A fines de 2015 más de 100 economías proveían datos de ISF al FMI, incluyendo a todas las economías del G20, un incremento respecto de las 45 economías de 2009.

Sin embargo, a fin de mantener la utilidad de los ISF como herramienta para la evaluación de la estabilidad financiera, la lista de indicadores se actualizó en 2013 para reflejar los cambios en el entorno financiero, principalmente la mayor prominencia de las instituciones financieras no bancarias, y las reformas regulatorias mundiales, en particular la aplicación de Basilea III.¹⁷ Esta última revisó las definiciones de capital e introdujo nuevas medidas

¹⁶ Ver por ejemplo *Financial Soundness Indicators and Banking Crises*, IMF Working Paper, núm. 13/263, <<https://www.imf.org/external/pubs/ft/wp/2013/wp13263.pdf>>.

¹⁷ Ver el documento del Directorio Ejecutivo, "Modifications to the Current List of Financial Soundness Indicators", noviembre, 2013, <<https://www.imf.org/external/np/pp/eng/2013/111313.pdf>>.

VÍNCULOS DENTRO DE LAS RECOMENDACIONES DE LA DGI-2



EBI: Estadísticas de la banca internacional.
 ECIC: Estudio Coordinado de Inversión de Cartera.
 ECID: Estudio Coordinado de Inversión Directa.
 EDSP: Estadísticas de deuda del sector público.
 EFG: Estadísticas de las finanzas gubernamentales.

IPIC: Índice de precios de inmuebles comerciales.
 IPIR: Índice de precios de inmuebles residenciales.
 ISF: Indicadores de solidez financiera.
 MCD: Medidas de concentración y distribución.
 PII: Posición de inversión internacional.

de apalancamiento, liquidez y financiamiento, todo lo cual se refleja en la lista actualizada de los ISF.

La lista actualizada de los ISF para las corporaciones no bancarias tiene como meta contribuir con el análisis y la evaluación de los efectos potenciales del sector de la banca paralela en la estabilidad del sistema financiero. Si bien la lista anterior sólo analizaba el subsector como un todo, el cual incluía un conjunto muy heterogéneo de instituciones, la nueva lista incluye ISF separados para los fondos de mercados de dinero, aseguradoras, fondos de pensión y otras instituciones financieras no bancarias. Se introdujeron también nuevos ISF para las empresas no financieras y los hogares.¹⁸ La DGI-2 dará mayor importancia al aumento de la frecuencia y cobertura con las cuales se informan los ISF, en especial para las instituciones financieras no bancarias (anexo, recomendación II.2).

La crisis también destacó la necesidad de tomar en cuenta los riesgos de un evento excepcional como complemento para la evaluación general de los riesgos del sector financiero mediante mediciones agregadas. Con este fin, para registrar las perturbaciones en todo el sistema que podrían ser causadas por instituciones que están en la cola de las distribuciones, se mejoraron las mediciones de los ISF agregados con un estudio piloto sobre las medidas de concentración y distribución (CDM, por sus siglas en inglés). Se considera que la expansión de los ISF para el sistema financiero con las CDM permitiría que los responsables de políticas y los funcionarios del Fondo registraran mejor el desempeño del sector financiero con mayor detalle y de manera prospectiva.

El proyecto piloto se completó en 2015 con la participación de 35 países diferentes.¹⁹ Las CDM se compilaron para seis ISF de los tomadores de depósitos: capital regulatorio nivel 1 a activos ponderados por riesgo, préstamos en mora a préstamos totales brutos, rentabilidad sobre activos,

rentabilidad sobre capital, activos líquidos a pasivos a corto plazo, y capital a activos totales.²⁰ Los datos proporcionaban información importante que los promedios no revelaban. Por ejemplo, las distribuciones de valores mínimos de las medidas de concentración y distribución²¹ que representan a las instituciones con los mayores riesgos para cualquier variable mostraron una variación significativa entre los países y entre distintos momentos dentro de los países. El proyecto piloto señaló que la emisión regular de informes de las CDM podría ser factible y podría ser una herramienta útil para vigilar las vulnerabilidades del sector financiero. En la DGI-2, el FMI expondrá la posibilidad de recopilar con regularidad los datos de las CDM (anexo 1, recomendación II.3)

En cuanto al sector bancario, el Banco de Pagos Internacionales llevó a cabo un trabajo conceptual enfocado en mediciones que abarcan el sistema de los descortes de plazos (brechas de financiamiento) en los balances generales internacionales de los bancos, con base en las estadísticas bancarias internacionales del BPI.²² El trabajo del BPI señaló que el análisis de los riesgos de financiamiento en todo el sistema bancario y la transmisión de choques entre países requiere de datos desagregados geográficamente en los balances generales de los bancos para registrar los patrones de financiamiento que son específicos de cada lugar, y facilita las evaluaciones orientadas de vulnerabilidades que se exponen en los datos agregados. Este trabajo ayudó a informar sobre las mejoras de las EBI del BPI que se adoptaron durante 2012-2015 mejorando de esta manera la utilidad de este conjunto de datos para la construcción de descortes de plazos y medidas de apalancamiento. El fortalecimiento de las

¹⁸ Ver <<<http://www.imf.org/external/np/sta/ISF/eng/ISF.htm>>.

¹⁹ Ver *Pilot Project on Concentration and Distribution Measures for a Selected Set of Financial Soundness Indicators*, IMF Working Paper, núm. WP/16/26, 2016.

²⁰ Las CDM incluían los siguientes indicadores: 1) mínimo, máximo y media; 2) desviaciones estándar ponderadas y asimetrías; y 3) cuartiles, y la distribución de activos en el cuartil inferior.

²¹ Valores máximos en el caso de los indicadores de solidez financiera sobre préstamos en mora.

²² Ver BPI, *Bank Structure, Funding Risk and the Transmission of Shocks Across Countries: Concepts and Measurement*, 2010, <http://www.bis.org/publ/qtrpdf/r_qt1009h.htm>.

EBI incluyó información mejorada sobre la contraparte, los plazos residuales y el desglose por divisa de las posiciones internacionales de los bancos, con mejoras a las estadísticas según la residencia y la consolidación (usando conceptos de supervisión conforme a la nacionalidad). La DGI-2 continuará mejorando la comunicación de las EBI a las economías del G20 (anexo, recomendación II.11).

2.4.2 Vigilancia de la banca paralela

Según el Consejo de Estabilidad Financiera, el sistema de la banca paralela se define como “la intermediación crediticia que abarca a entidades y actividades fuera del sistema regular de la banca”. Tales instituciones podrían brindar fuentes alternativas de financiamiento para los participantes de mercados complementando a la banca tradicional, pero podrían también acarrear riesgos similares a los que implican los bancos. Esos riesgos podrían fácilmente esparcirse al resto del sistema debido a las complejas relaciones entre estas instituciones y los bancos, y por lo tanto necesitan ser vigilados.

Por lo general, estas instituciones están muy apalancadas y dependen del financiamiento a corto plazo al mismo tiempo que invierten en activos ilíquidos a largo plazo; por lo tanto están expuestas a riesgos de liquidez y de plazo. Durante la crisis, cuando se materializaron estos riesgos, todo el sistema financiero sufrió las consecuencias, y de este modo quedó clara la importancia de vigilar estos riesgos.

En la Cumbre de 2011 en Cannes, los líderes del G20 pidieron que el CEF abordara las dudas sobre estabilidad financiera relacionadas con la banca paralela. La estrategia del CEF tiene dos elementos.²³

- Primero, el CEF inició un ejercicio anual de vigilancia de la banca paralela en escala mundial.

- Segundo, el CEF está trabajando para formular políticas que fortalezcan la supervisión y regulación del sistema de la banca paralela.²⁴

El informe anual del CEF abarca a 26 jurisdicciones las cuales, a partir de 2014,²⁵ constituyen el 80% del producto interno bruto (PIB) mundial y el 90% de los activos del sistema financiero mundial, y se basa en los datos de balances de las cuentas financieras nacionales. Los informes anuales son una recopilación coordinada de datos agregados en escala planetaria para permitir el análisis de tendencias y riesgos mundiales en el sistema de la banca paralela. Por primera vez, el informe de 2015 introdujo una nueva medida de la banca paralela basada en las funciones económicas de las entidades financieras no bancarias y enfocadas solamente en aquellas instituciones financieras no bancarias involucradas en significativas transformaciones de madurez/liquidez o apalancamiento y que son parte de una cadena de intermediación crediticia. Esto permite que los responsables de la política se enfoquen más en los riesgos potenciales que podría plantear la banca paralela. Con base en esta medida, se incrementaron los activos mundiales de las entidades financieras clasificadas como banca paralela en comparación con 2014, alcanzando 36 billones de dólares estadounidenses teniendo por telón de fondo una leve caída de los activos del sistema bancario. Sin embargo, esta es una estimación inicial sujeta a cambio a medida que madura la aplicación de la metodología.

En cuanto a la supervisión y la regulación del sistema de la banca paralela, entre los temas tratados están los riesgos en los mercados de préstamos de valores y de reportos. Esta crisis dejó en claro

²³ Ver CEF, *Strengthening Oversight and Regulation of the Shadow Banking: An Overview of Policy Recommendations*, agosto, 2013.

²⁴ El progreso del trabajo del CEF sobre la banca paralela quedó expuesto en el informe del CEF de 2015: *Transforming Shadow Banking into Resilient Market-based Finance. Overview of Progress*, <http://www.financialstabilityboard.org/wp-content/uploads/shadow_banking_overview_of_progress_2015.pdf>.

²⁵ CEF, *Global Shadow Banking Monitoring Report*, noviembre, 2015, <<http://www.financialstabilityboard.org/wp-content/uploads/global-shadow-banking-monitoring-report-2015.pdf>>.

que el financiamiento basado en figuras semejantes a depósitos de corto plazo en las entidades no bancarias puede fácilmente llevar a corridas en el mercado si se pierde la confianza. El uso de estas técnicas de financiamiento garantizado puede exacerbar estas corridas y fomentar el apalancamiento, en especial cuando los precios de los activos están en bonanza y son bajos los márgenes de seguridad sobre el financiamiento asegurado. Por lo tanto, el CEF inició un trabajo para recabar y acumular datos sobre los mercados de financiamiento de valores que actualmente está incorporado en la DGI-2. (anexo, recomendación II.5). Se intenta iniciar la recopilación de datos agregados para los mercados de financiamiento de valores para 2017, con el apoyo operativo del BPI.

2.4.3 Vigilancia de las instituciones financieras de importancia sistémica mundial (G-SIFI)

Dada la importancia de las G-SIFI en la diseminación de choques entre países y los efectos potenciales de su quiebra para el sistema financiero mundial, se tomaron diversas medidas para mejorar la capacidad de recuperación de estas instituciones y así limitar los efectos del riesgo moral. Entre estas medidas estaba la identificación de los G-SIB en 2011 por parte del Comité de Supervisión Bancaria de Basilea y la introducción de medidas adicionales de absorción de pérdidas para estas instituciones. Contar con mejores datos sobre las vinculaciones bilaterales de estas instituciones, así como sus exposiciones y dependencias de financiamiento de los sistemas financieros nacionales, se consideró requisito importante para entender los riesgos asociados con estas instituciones. A tal fin, el trabajo para la elaboración de una plantilla de datos de las G-SIFI, como se recomendaba en la DGI, se enfocó inicialmente en los bancos mundiales de importancia sistémica.

El producto final de este ejercicio, liderado por el CEF, con el asesoramiento directo del FMI, es un conjunto de plantillas de datos únicos que reúnen

información detallada y sólida sobre los G-SIB que es útil tanto para el análisis micro como macroprudencial. La recopilación de datos empezó con información sobre las conexiones bilaterales entre los G-SIB, así como con alguna información agregada basada en los datos por instituciones que están en las EBI consolidadas y continuarán con la recopilación de información sobre las exposiciones de los G-SIB y sus préstamos en 35 economías clave con el detalle de una combinación de sector, instrumento, moneda y plazo. Los datos se recopilan en el International Data Hub establecido en el BPI y actualmente se comparten con las autoridades nacionales que aportan datos.²⁶ Este proceso en sí ha reforzado el intercambio de información y la coordinación entre las autoridades nacionales de supervisión. Sin embargo, dado el detalle del conjunto de datos, se plantean temas de confidencialidad que necesitan ser abordados a largo plazo para hacer mejor uso de esta información crítica.

El objetivo de las planillas es aportar a las autoridades una visión más clara sobre las redes financieras mundiales y ayudarlas en sus responsabilidades de supervisión y macroprudenciales.²⁷ Los datos de los G-SIB respaldarían el trabajo del FMI para salvaguardar la estabilidad financiera mundial incluso con la eficaz supervisión multilateral y bilateral y el estímulo a respuestas de política coherentes entre los países miembro.²⁸ Los datos permitirían que la información por banco se use en conjunto con mediciones de exposiciones mundiales, mejorando de manera significativa la capacidad de encontrar vulnerabilidades que podrían originarse por las exposiciones comunes/las posiciones de financiamiento concentrado, de modo que se profundiza en el conocimiento de la fuente potencial de diseminaciones. Los datos de los G-SIB mejorarían también el

²⁶ En principio, se ha acordado compartir los informes de datos de los G-SIB con las instituciones financieras internacionales (BPI, CEF y FMI) en estrictas condiciones de confidencialidad.

²⁷ Ver *FSB Data Gaps Initiative – A Common Data Template for Global Systemically Important Banks*, mayo de 2014.

²⁸ Ver *2014 Triennial Surveillance Review – Managing Director’s Action Plan for Strengthening Surveillance*, página 3.

seguimiento de las actividades de financiamiento en distintas monedas y de transformación de los plazos. Además, los datos ayudarían a mejorar la comprensión de la innovación financiera, la complejidad del mercado y las fuentes emergentes de potenciales riesgos sistémicos.

Más allá de la industria bancaria, el CEF y la Asociación Internacional de Supervisores de Seguros (IAIS, por sus siglas en inglés) han identificado también a las compañías de seguros de importancia sistémica mundial gracias a una metodología desarrollada por la IAIS.²⁹ La metodología de evaluación se relaciona con la metodología elaborada por el Comité de Supervisión Bancaria de Basilea para los G-SIB pero también toma en cuenta la naturaleza específica del sector asegurador. En particular, los grupos aseguradores que se involucran en actividades no tradicionales o distintas del aseguramiento pueden ser vulnerables a los riesgos de precios del mercado y de liquidez, amplificando o contribuyendo al riesgo sistémico.³⁰ Por lo tanto, tales actividades no tradicionales se incluyen como indicador en la metodología de evaluación. Como consecuencia de estos avances regulatorios, en la DGI-2 se investigará la posibilidad de elaborar una plantilla común de datos para las instituciones financieras no bancarias de importancia sistémica en el plano mundial empezando con las compañías de seguros (anexo, recomendación II.4).

2.4.4 Comprensión de los mercados financieros

Si bien son un canal importante para el financiamiento de la economía real, los mercados de valores han sido también un canal clave para la transmisión del riesgo, en especial debido a la creciente dependencia en el financiamiento basado en el mercado. Por lo tanto, hay consenso sobre la importancia de contar con mejor información sobre

estos mercados para entender la diversificación de fuentes de financiamiento y las exposiciones de los emisores y de los acreedores, incluyendo en el sector no financiero. Aunque hace mucho tiempo que se consideran de importancia en las economías avanzadas, hay pruebas de una creciente emisión de valores en las economías de mercado emergentes (EME), ya que la composición de la deuda empresarial ha abandonado los préstamos y pasado a los bonos;³¹ se calcula que durante el último decenio, los mercados de títulos de deuda nacional en las economías de mercado emergentes han aumentado de aproximadamente un tercio del PIB de la EME a casi la mitad.³²

La DGI ha abordado este interés creciente de política con respecto a los mercados de valores brindando asesoramiento conceptual mediante la publicación del *Handbook on Securities Statistics*³³ preparado conjuntamente por el BPI, el Banco Central Europeo y el FMI; y promoviendo mejoras en las estadísticas de valores y títulos, al alentar que las economías del G20 entreguen sus estadísticas sobre valores a la base de datos del BPI.

Desde 2009, la cantidad de economías que informan de manera regular y consistente sus estadísticas de valores al BPI ha aumentado de manera significativa. Además, si bien son diversos los niveles de complejidad de la estructura estadística nacional entre las economías del G20, cada vez son más las que reconocen la importancia de tener una información pormenorizada sobre estos mercados, por lo que están considerando construir bases de datos por cada título.³⁴ Los datos sobre

²⁹ Disponible en <<http://iaisweb.org>>.

³⁰ IAIS 2011, *Insurance and Financial Stability*, <<http://iaisweb.org/index.cfm?event=openFile&nodeId=34041>>.

³¹ FMI, *Informe sobre la estabilidad financiera mundial*, capítulo 3, octubre, 2015.

³² Ver Masazumi Hattori y Előd Takáts, *The Role of Debt Securities Markets*, noviembre, 2015, <https://www.bis.org/publ/bppdf/bispap83c_rh.pdf>.

³³ Ver *Handbook on Securities Statistics*, <<<http://www.imf.org/external/np/sta/wgsd/hbook.htm>>>.

³⁴ La base de datos del BPI sobre títulos de deuda internacionales, que recopila información pormenorizada que ha permitido la identificación paralela de la residencia y de la nacionalidad de los emisores de títulos de deuda, es un ejemplo de una base de datos por título elaborada en escala internacional.

la emisión (y los acervos) de valores son también un insumo de las cuentas nacionales, las balanzas de pagos y las estadísticas de finanzas gubernamentales. El enfoque inicial de la DGI-2 es mejorar los datos sobre la emisión de títulos de deuda con información clave sobre los mercados, los sectores, la moneda, el plazo y la tasa de interés. La información sólida sobre los acervos de títulos de deuda y los datos de las partes involucradas se consideran una meta a más largo plazo (anexo, recomendación II.7).

La necesidad de aclarar la opacidad de los mercados de derivados OTC es también un enfoque de la DGI. En la DGI-1 los datos de las permutas de incumplimiento crediticio (CDS) se expandieron tanto en el detalle como en la cobertura del país y se comenzaron a elaborar informes periódicos sobre este conjunto de datos expandido. Todas las economías con importantes mercados de CDS informan los datos de las CDS a la encuesta del BPI, la cual incluye más detalles sobre el tipo y la geografía de las contrapartes así como el instrumento subyacente.³⁵

En la DGI-2, se reconoce la necesidad de mejorar los datos de los mercados de derivados OTC de manera más general (anexo, recomendación II.6). En septiembre de 2009, los líderes del G20 acordaron una agenda de reforma integral para mejorar la transparencia de los mercados de derivados OTC, mitigar el riesgo sistémico y proteger contra los abusos de los mercados; solicitaron también al CEF y a sus miembros relevantes que evaluaran regularmente su aplicación. Los objetivos de esta reforma incluyen la entrega de información sobre los contratos de derivados OTC al registro de operaciones, las operaciones de todos los contratos estandarizados de intercambios o de plataformas de operaciones electrónicas, donde corresponda, con liquidación a través de contrapartes centrales (CCP). Los contratos no compensados a través de una contraparte central están sujetos a mayores

requerimientos de capital. Esta iniciativa regulatoria para compensar los derivados OTC a través de una cámara de compensación permite una mayor estandarización de la forma y agregación para los datos regulatorios y financieros. A la vez, estos avances también están incrementando el interés en la calidad de la información que se proporciona, así como en la congruencia del conjunto de datos ya disponibles.

2.5 ¿CÓMO ABORDA LA DGI LA AGENDA DE SUPERVISIÓN?

Como se señaló antes, como consecuencia del ETS 2014 la dirección del FMI publicó un Plan de Acción para Reforzar la Supervisión. Entre las acciones por llevar a cabo estaba la del Fondo como responsable de “revivir y adaptar el enfoque del balance general y así facilitar un análisis más profundo del efecto de los choques y su transmisión entre sectores”. Fue la respuesta a un pedido de los expertos externos David Li y Paul Tucker en su estudio externo para el ETS 2014 sobre los riesgos y las diseminaciones.³⁶

2.5.1 Análisis sectorial

Si bien la crisis de 2007-2008 brotó en el sector financiero, debido a sus funciones de intermediario, los problemas en el sector financiero también afectaron a otros sectores de la economía. Para entenderlo, es esencial el análisis de las exposiciones del balance ante una economía mundial cada vez más conectada. Tal como se señala en el ETS 2014 del FMI, el uso de balances generales para identificar las fuentes de vulnerabilidad y la transmisión de choques podría haber ayudado a identificar los riesgos asociados con la dependencia de los bancos europeos respecto el financiamiento mayorista de EUA para financiar productos estructurados.

³⁵ BPI, estadística semestral de derivados OTC, <<http://www.bis.org/statistics/derstats.htm>>.

³⁶ Ver *Triennial Surveillance Review –External Study –Risks and Spillovers*, julio, 2014, <<https://www.imf.org/external/np/pp/eng/2014/073014e.pdf>>.

En junio de 2015, el FMI estableció el rumbo en un documento del Directorio Ejecutivo del FMI sobre Análisis de los Balances Generales en la Supervisión.³⁷ Las cuentas sectoriales y los datos de los balances generales son esenciales, entre ellos los datos sobre las partes involucradas, para brindar el marco para una evaluación de las relaciones entre la economía real y los sectores financieros. Los balances generales sectoriales del sistema de cuentas nacionales se consideran como el marco dominante para el análisis de balances generales como aclara el documento del Directorio Ejecutivo del FMI. Además, el artículo establece un marco referencial de datos para este tipo de análisis.³⁸

Si se pone el balance general sectorial de las cuentas nacionales en un contexto de política, el FMI ha elaborado un enfoque de balance (BSA, por sus siglas en inglés), el cual recaba todos los balances principales en una economía usando datos agregados por sector. El BSA se basa en los mismos principios conceptuales que las cuentas sectoriales, brindando información sobre las partes involucradas con un enfoque adicional en las vulnerabilidades que surgen de los descalces de plazos y de moneda, así como la estructura de capital de los sectores económicos.

Si bien actualmente no son muchas las economías que recaban datos de los balances generales con las partes involucradas, los datos del enfoque del balance pueden recabarse de las formas estandarizadas de reporte del FMI, de la PII, y de los datos de balances generales de los gobiernos, conjunto de datos más limitado que los que se necesitan para recabar información sobre las cuentas sectoriales.

Las recomendaciones de la DGI-2 abordan las deficiencias clave de datos que actúan como restricción para un análisis cabal de los balances. La DGI recomienda abordar tales deficiencias

mejorando la divulgación de las cuentas sectoriales de las economías del G20 y de sus balances generales con base en la cuentas nacionales de 2008, incluyendo a los sectores empresarial no financiero y de hogares (anexo, recomendación II.8) Dado el carácter multifacético de los conjuntos de datos, la aplicación de esta recomendación es todo un reto y su progreso ha sido lento. Sin embargo, todas las economías del G20 coinciden en la importancia de tener esta información y tienen planes en marcha para lograrla.

En un mundo de flujos de capital libres y de menos limitaciones para el crédito, la ampliación de las distribuciones del ingreso, el consumo, los ahorros y la riqueza puede llevar a potenciales vulnerabilidades financieras aun si los datos agregados lucen tranquilizadores. Ciertamente, en años recientes, se ha vuelto cada vez más evidente la importancia de los buenos datos de distribución para el sector de hogares, ya que en los últimos decenios ha aumentado el interés de las políticas por la desigualdad tanto en las economías avanzadas como en las en desarrollo.³⁹ La DGI-2 se enfoca en la recopilación de información distributiva (tal como la información por quintiles de ingreso) para complementar las cifras agregadas, en consonancia con las cuentas nacionales (anexo, recomendación II.9). A estos efectos, la OCDE ha emprendido un trabajo conceptual sobre la información distributiva que se enfoca en: 1) vincular las cuentas nacionales con la información distributiva (micro y macrodatos), resumida en dos documentos de trabajo de la OCDE,⁴⁰ y 2) la provisión de un alineamiento conceptual mejorado del ingreso, el consumo y la riqueza mediante microencuestas, que incluyen un fortalecimiento adicional de las definiciones de riqueza.

³⁷ FMI, *Balance Sheet Analysis in Fund Surveillance*, junio, 2015, <<http://www.imf.org/external/np/pp/eng/2015/061215.pdf>>.

³⁸ FMI, *Balance Sheet Analysis in Surveillance*, junio, 2015, párrafo 27, página 23.

³⁹ Ver el trabajo del FMI sobre desigualdad de ingresos: <<http://www.imf.org/external/np/fad/inequality/index.htm>>.

⁴⁰ <<http://www.oecd.org/statistics/icw-framework.htm>> y <<http://www.oecd.org/statistics/guidelines-for-micro-statistics-on-household-wealth.htm>>.

2.5.2 Análisis de las condiciones fiscales

También hay significativas deficiencias de datos en el área de las estadísticas de finanzas públicas. El apoyo brindando por muchas autoridades nacionales al sector financiero luego de la crisis financiera mundial, junto con el comienzo de la recesión y de los programas de estímulo fiscal para apoyar la demanda, llevó a aumentos de los déficits fiscales y de las deudas de los gobiernos. Sin embargo, se carecía de datos fiables y comparables entre las economías del G20, obstaculizando así los análisis transfronterizos. Además, la vigilancia de las tendencias en la posición fiscal de los gobiernos a menudo estuvo limitada por la falta de datos frecuentes y oportunamente armonizados, incluso por la falta de datos basados en lo que se ha devengado.

En escala internacional, ha habido un progreso significativo a favor de la compilación de estadísticas de finanzas públicas. El trabajo conceptual ha incluido la publicación del *Manual de estadísticas de finanzas públicas* (MEFP) de 2014. Además, el Directorio Ejecutivo del FMI comenzó a abordar el problema de las estadísticas de finanzas públicas en 2010, reafirmado en 2013, requiriendo la inclusión en los informes del personal de los elementos clave de la presentación del *Manual de estadísticas de finanzas públicas*.⁴¹ El documento al cual se hace referencia en la nota al pie anterior confirmó también la intención de establecer un Comité Asesor de Estadísticas de Finanzas Públicas (GFSAC, por sus siglas en inglés) con el fin de respaldar la aplicación del MEFP y de asesorar en temas de datos fiscales emergentes. El GFSAC se reunió por primera vez en marzo de 2015 y acordó respaldar la aplicación del MEFP 2014 con una serie de recomendaciones prácticas.⁴²

A pesar de estos acontecimientos en el ámbito internacional, el progreso ha quedado rezagado con

respecto a otras recomendaciones. Esto se debe a factores tales como la falta de cobertura de los gobiernos estatales y locales, al hecho de que las estadísticas de buen gobierno financiero no están bien establecidas en algunos países desde el punto de vista institucional y, en algunas instancias, a la renuencia de las autoridades a usar técnicas estadísticas para zanjar la deficiencia de datos.⁴³ A estos efectos, la DGI-2 sigue teniendo como meta cubrir las deficiencias en las estadísticas financieras de los gobiernos (anexo, recomendación II.15).

En el mismo contexto, la información sobre los niveles de deuda del sector público, en particular del gobierno general, es crucial para evaluar la solidez fiscal del gobierno. Bajo la DGI, en 2010 el Banco Mundial, la OCDE y el FMI establecieron una base de datos de estadísticas de la deuda del sector público para promover que los países informen de modo estandarizado. Sin embargo, el alcance del sector y la cobertura de instrumentos pueden diferir significativamente entre países. Esto es muy relevante debido al interés de política y analítico en razones tales como la de deuda bruta a PIB. Si un país cubre *solamente* los títulos de deuda y los préstamos al gobierno, la comparación de estos datos con los de un país que incluye todos sus pasivos de deuda (incluidas las cuentas por pagar y las obligaciones por pensión del gobierno general) claramente no será una comparación pareja.⁴⁴ Por lo tanto, respaldada por la DGI-2, la base de datos de deuda pública del Banco Mundial está pasando a una presentación de cobertura de instrumentos y sectorial con base en una matriz mediante la presentación de distintos niveles de sectores y de

⁴¹ Ver *Review of the Implementation of Government Finance Statistics to Strengthen Fiscal Analysis*, noviembre de 2013.

⁴² Ver *Proceedings of the Meeting of the IMF Government Finance Statistics Advisory Committee (GFSAC)*, <<http://www.imf.org/external/pubs/ft/gfs/gfsac/meetings/2015/>>

⁴³ Además, las estadísticas de finanzas públicas no siempre son compatibles con los datos relevantes en las cuentas nacionales a pesar de la armonización de los estándares internacionales entre los dominios estadísticos.

⁴⁴ Ver, *What Lies Beneath: The Statistical Definition of Public Sector Debt*, FMI Staff Discussion Notes, núm. 12/09.

cobertura de instrumentos, de la más estrecha a la más extensa⁴⁵ (anexo, recomendación II.16).⁴⁶

2.5.3 Entendimiento de las interconexiones transfronterizas

La crisis puso en la palestra el hecho de que no es posible aislar los problemas de un solo sistema financiero, ya que los choques se propagan con rapidez entre los distintos sistemas financieros. Ciertamente, desde 2010, el FMI ha estado definiendo jurisdicciones con sectores financieros de importancia sistémica con base en un conjunto de criterios relevantes y transparentes que incluyen el tamaño y la conectividad. Dentro de este marco de identificación, las interconexiones transfronterizas se consideran una importante medida complementaria del tamaño de la economía: registran el riesgo sistémico que puede surgir por las vinculaciones directas e indirectas entre los sectores financieros del sistema financiero mundial (esto es, el riesgo de que el quiebre o el mal funcionamiento de un sistema financiero pueda tener repercusiones profundas en otros países o en la estabilidad sistémica general).⁴⁷

El ETS 2014 resumió este tema de manera sucinta en su resumen ejecutivo: “Los riesgos y las diseminaciones siguen siendo asuntos de primer orden para la economía mundial y deberían ser cruciales para la supervisión del Fondo. Las reformas recientes han hecho que la supervisión se base más en el riesgo con el fin de registrar mejor las interconexiones mundiales. La experiencia hasta ahora apunta

a la necesidad de obtener un mejor entendimiento de dónde se ubican los riesgos en los países, y la manera en que las diseminaciones pueden extenderse con rapidez entre los sectores dejando expuestas las vulnerabilidades nacionales.”⁴⁸

Cuatro de los conjuntos de datos disponibles que incluyen información clave sobre vinculaciones financieras transfronterizas son PII, EBI del BPI, ECIC del FMI y ECID del FMI. En total, estos conjuntos de datos brindan un panorama integral de las interconexiones financieras transfronterizas. Este panorama es especialmente relevante para quienes formulan la política, ya que las conexiones financieras se fortalecen y las condiciones nacionales se ven afectadas por los sucesos financieros en otras economías con las cuales tienen estrechos vínculos financieros. La DGI-2 se enfoca en la mejora de la disponibilidad y la comparabilidad de estas bases de datos entre países (anexo 1, recomendaciones II.10, 11, 12 y 13).

Las posiciones de inversión internacional son una fuente clave de datos para entender los vínculos entre la economía nacional y el resto del mundo al proporcionar información sobre los activos y pasivos externos de la economía en un desglose detallado de instrumentos. Sin embargo, la crisis reveló la necesidad de desgloses por moneda y desgloses más detallados por sector, en especial para el sector de otras corporaciones financieras (OCF). Por lo tanto, como parte de la DGI, se mejoró la PII para respaldar estas necesidades de política. También ha sido significativo el avance en garantizar la generación regular de informes de la PII junto con el incremento en la frecuencia de anual a trimestral. Para fines de 2015, casi todas las economías del G20 elaboraban informes de datos de PII cada tres meses.

Durante decenios, las EBI han sido una fuente clave de datos de información trimestral sobre activos y pasivos agregados de los sistemas bancarios con actividad internacional. Los datos de las

⁴⁵ Ver, “Progress with Globally Comparable Public Sector Debt Statistics”, documento para el GFSAC, disponible en <<http://www.imf.org/external/pubs/ft/gfs/gfsac/meetings/2015/pdf/1515a.pdf>>.

⁴⁶ El BPI ha publicado un conjunto de datos sobre créditos para el sector de gobierno general de 26 economías avanzadas y para 14 emergentes, *BIS Quarterly Review*, septiembre, 2015, <http://www.bis.org/publ/qtrpdf/r_qt1509g.htm>.

⁴⁷ Ver FMI, *Expanding Surveillance to Require Mandatory Financial Stability Assessments of Countries with Systemically Important Financial Sectors*, 2010, <<https://www.imf.org/external/np/sec/pr/2010/pr10357.htm>>.

⁴⁸ Ver *IMF Triennial Surveillance Review –Overview Paper*, 2014, <<http://www.imf.org/external/np/pp/eng/2014/073014.pdf>>.

ECIC, si bien tienen frecuencia anual, ofrecen información relevante sobre los activos de inversiones de cartera. Dicho esto, ambos conjuntos de datos tenían limitaciones en términos de cobertura por país y por detalle. Las ECIC necesitaban también mejorarse en cuanto a su frecuencia y oportunidad. A tal fin, la DGI respaldó las mejoras en estos conjuntos de datos.

Las EBI se fortalecieron mediante 1) la cobertura ampliada de los balances bancarios para incluir también las posiciones nacionales como complemento a las actividades transfronterizas, 2) detalle mejorado en los datos recabando más información por país y por sector de la contraparte de los bancos, en particular para las instituciones financieras no bancarias. Las mejoras a las EBI ofrecerán a los usuarios un cuadro más integral del tamaño y alcance de las actividades de los bancos activos en escala internacional. Así, permitirá realizar un análisis mejorado de las fuentes y usos de los fondos y de la importancia de los negocios internacionales para los bancos de distintas nacionalidades.⁴⁹

La ECIC comenzó a realizarse con frecuencia semestral desde junio de 2013 con un rezago de divulgación de menos de nueve meses.⁵⁰ Hacia fines de 2015, todas las economías del G20 proporcionaron datos de las ECIC al FMI; 18 de las economías lo hicieron de manera semestral. Además, hay un interés creciente en comprender la distribución del sector de tenedores y emisores, a fin de combinar el análisis sectorial con las cuentas nacionales. A fines de 2015, 15 economías del G20 proporcionaron datos del sector de tenedores, al mismo tiempo que en las mejoras realizadas en la DGI había nuevos cuadros mejorados para recabar información sobre el sector de emisores de títulos, cruzado también con el sector de tenedores

de valores. La DGI-2 sigue enfocada en mejorar la elaboración de informes de las ECIC en las economías del G20.

La ECID del FMI complementa a las ECIC y las EBI para el análisis de interconectividad transfronteriza, ya que ofrece información sobre las posiciones de inversión directa desglosadas por capital neto y por deuda neta. La ECID quedó bajo la sombra de la DGI en su segunda fase con el fin de mejorar la calidad de las estadísticas de las posiciones de inversión directa de las economías del G20, tanto para las inversiones entrantes como para las salientes.

El riesgo de divisa es un elemento importante en el análisis de las interconexiones transfronterizas. A estos efectos, se prestó particular atención, incluso por parte de los FMCBG del G20,⁵¹ a la mejora de la información sobre la exposición de divisas dados los potenciales efectos de diseminación de las transferencias de la riqueza desatadas por los bruscos movimientos en los tipos de cambio. En este contexto, el FMI se enfocó en mejorar la compilación de datos de exposiciones de divisas en sus dominios estadísticos, particularmente por medio de las PII. El BPI contribuye al análisis de exposiciones de divisas mediante sus títulos de deuda internacional y su EBI mejorada, la cual ofrece las bases para obtener un panorama más detallado de los balances de los bancos activos en el plano internacional y así medir con más exactitud los potenciales descalces de moneda.

La crisis financiera mundial reveló también la necesidad de entender mejor las exposiciones transfronterizas de divisas de las empresas no financieras. El CEF/Comité sobre el Sistema Financiero Mundial (CSFG)⁵² identificó las deficiencias de información para las exposiciones en divisas en un taller conjunto realizado en 2014 y entregó los resultados a los FMCBG del G20. Además, en

⁴⁹ "Enhanced Data to Analyse International Banking", *BIS Quarterly Review*, septiembre, 2015, <http://www.bis.org/publ/qtrpdf/r_qt1509f.pdf>.

⁵⁰ Ver Twenty-fourth Meeting of the IMF Committee on Balance of Payments Statistics "Enhancements to the Coordinated Portfolio Investment Survey", 2011, <<https://www.imf.org/external/pubs/ft/bop/2011/11-12.pdf>>.

⁵¹ Ver 2015 *BIS/FSB/IMF Reports on Foreign Currency Exposures to the G20*, <<http://www.financialstabilityboard.org/wp-content/uploads/Work-on-Foreign-Currency-Exposures.pdf>>.

⁵² El CSFG es el comité de bancos centrales del BPI que supervisa el conjunto de estadísticas de las EBI.

2014 el CEF llevó a cabo una revisión realizada por colegas del acervo de información de operaciones que incluye todo tipo de derivados OTC, entre ellos derivados de divisas y otros instrumentos que crean exposiciones de divisas.⁵³ También se informó de este trabajo al G20. En términos generales, la DGI-2 respalda este trabajo sobre exposiciones de divisas por medio de la incorporación más explícita de datos de exposiciones de divisas en las recomendaciones (esto es, PII, exposiciones transfronterizas de empresas no bancarias y estadísticas de títulos y valores).

El incremento en las exposiciones en derivados de divisas de las empresas no financieras por medio de entidades fuera de plaza ha sido un tema de preocupación, particularmente para las economías de mercado emergentes, ya que las autoridades no estaban al tanto de las transacciones registradas fueran de sus jurisdicciones. La estructura de la DGI está contribuyendo a aclarar este tema más general de las exposiciones transfronterizas y del financiamiento intragrupal por parte de las empresas no financieras por medio de sus filiales externas (anexo, recomendación II.14). Se brindaron lineamientos conceptuales para aclarar los conceptos de nacionalidad, grupo y consolidación en la DGI-1⁵⁴ y a futuro, el BPI y el FMI continuarán mejorando la información sobre las exposiciones transfronterizas de las empresas no financieras recurriendo principalmente a la compilación de los datos disponibles. La OCDE colaborará también mediante la formulación de un marco de referencia para vincular sus datos sobre empresas transnacionales con sus datos de inversión extranjera directa.

2.5.4 Vigilancia de los mercados de propiedades

Los índices de precios de las propiedades residenciales y comerciales son importantes para la identificación y vigilancia de burbujas de precios de los activos, para la compilación de estimaciones de riqueza del sector de hogares y del empresarial y de formación de capital, y para la evaluación de las implicaciones más generales sobre la estabilidad financiera. La relevancia de los precios de las propiedades quedó de manifiesto en el segundo Foro de Estadísticas del FMI,⁵⁵ mientras que el trabajo del BPI ha destacado la importancia de la evolución de los precios de los activos, en especial los precios de las propiedades, para impulsar el llamado ciclo económico.⁵⁶

Sin embargo, antes de la crisis financiera mundial, su disponibilidad y comparabilidad internacional eran limitadas. Como parte de la DGI-1, se proporcionaron lineamientos conceptuales por medio del *Handbook on Residential Property Price Indices* (RPPI),⁵⁷ y en 2010 el BPI comenzó a divulgar las estadísticas de precios de bienes raíces en su página web. En la actualidad, la mayoría de los países del G20 proporcionan datos si bien estos a menudo están en desarrollo y se necesita más trabajo para garantizar la solidez y la comparabilidad internacional de los mismos.

En los últimos años ha quedado cada vez más clara la importancia de que los responsables de la política cuenten con buenas estadísticas de precios de los inmuebles dado el vínculo con el consumo

⁵³ El CEF profundizó más sobre las estructuras de financiamiento y los incentivos de las empresas no financieras en su informe al G20 sobre *Corporate Funding Structures and Incentives*, de febrero de 2015, <<http://www.financialstabilityboard.org/2015/09/corporate-funding-structures-and-incentives/>>.

⁵⁴ *Consolidation and Corporate Groups: Overview of Methodological and Practical Issues*, IAG Reference Document, <<http://www.bis.org/ifc/publ/iagrefdoc-oct15.pdf>>.

⁵⁵ Ver *Real Estate Prices –Availability, Importance, and New Developments*, <<http://www.imf.org/external/np/seminars/eng/2014/statsforum/>>.

⁵⁶ Ver Borio y Drehmann, “Assessing the Risk of Banking Crises –Revisited”, *BIS Quarterly Review*, marzo, 2009, y Christian Dembiermont, *Measuring Property Prices: The BIS Contribution*, septiembre, 2015, <https://www.bis.org/ifc/events/ifc_isi_2015/432_dembiermont_paper.pdf>.

⁵⁷ Ver *Handbook on Residential Property Prices Indices*, <<http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-12-022>>. Este manual tuvo el apoyo conjunto del Eurostat, la Organización Internacional del Trabajo (OIT), el FMI, la OCDE, Naciones Unidas y el Banco Mundial.

del sector de hogares y la necesidad de vigilar los precios de los activos en un entorno de políticas monetarias expansivas. Por lo tanto, la mayoría de las economías del G20 ha estado realizando mayores esfuerzos para elaborar buenas estadísticas sobre los precios de bienes raíces. El RPPI es uno de los indicadores de ISF que se prescriben para los adherentes a las NEDD Plus.

Los buenos índices de precios de inmuebles pueden también respaldar la medición de activos no financieros en las cuentas sectoriales.

Los índices de precios de las propiedades comerciales son un estadio menos desarrollado, tanto conceptualmente como en términos de disponibilidad de datos. Sin embargo, es importante para estabilidad financiera la divulgación de los datos del índice de precios de las propiedades comerciales para vigilar las burbujas de activos, ya que las propiedades comerciales se usan para los préstamos garantizados de los bancos; y tales datos son importantes para la evaluación de los activos titulizados. Por lo tanto, se necesita trabajar para mejorar el manual metodológico que se está formulando y para fomentar la provisión de datos disponibles al BPI para su divulgación.

En las recomendaciones II.17 y 18 del DGI-2 se aborda la elaboración de los índices de propiedades.

2.6 LA NECESIDAD DE COMUNICAR MEJOR LAS ESTADÍSTICAS

La DGI proporcionó un nuevo nivel de estándares de datos del FMI (NEDD Plus), principalmente destinado a las economías con sistemas financieros de importancia sistémica, para guiar a los países miembro del FMI en la provisión de datos económicos y financieros para el público en apoyo de la estabilidad financiera nacional e internacional. Se espera que las economías que se adhieren a las NEDD Plus divulguen los datos en nueve categorías que abarcan cuatro sectores macroeconómicos: el sector real, el sector fiscal, el sector financiero y el sector externo, tomados en su mayor parte de

las recomendaciones de la DGI-1. Dadas las áreas comunes de enfoque, la adhesión a las NEDD Plus y la aplicación de las recomendaciones de la DGI se apoyarían entre sí. Al momento de la redacción del presente documento, ocho países se habían adherido a las NEDD Plus.

El FMI también continúa considerando las necesidades de los mercados emergentes y de los países de bajos ingresos. Las NEDD tienen como meta mejorar la disponibilidad de estadísticas oportunas y completas, y por lo tanto contribuir en la búsqueda de políticas macroeconómicas acertadas, y al mejor funcionamiento de los mercados financieros. En mayo de 2015, se mejoró el Sistema General de Divulgación de Datos (e-GDDS) para auxiliar a los países con capacidad estadística relativamente menos avanzada. La atención en la divulgación de datos en el e-GDDS respaldará la transparencia, fomentará el desarrollo estadístico y ayudará a generar fuertes sinergias entre la divulgación de datos y la supervisión. En el futuro, el FMI continuará trabajando con las economías miembro, por ejemplo en actividades de desarrollo de capacidades, para garantizar la disponibilidad y la divulgación de la información.

Hubo también una necesidad de mejorar la comunicación de las estadísticas oficiales, ya que en algunas instancias los usuarios no tuvieron acceso pleno a las series de datos disponibles para abordar temas críticos de política. Como parte de la DGI, en 2009 se lanzó la página web Principal Global Indicators (PGI),⁵⁸ creada por el FMI como un proyecto conjunto con el IAG cuya meta era facilitar la vigilancia de sucesos económicos y financieros. La página web PGI incluye datos de las economías del G20 y de 14 países ajenos al Grupo que tienen sectores financieros de importancia sistémica y son objeto del mandato de cinco años del Programa de Evaluación del Sector Financiero (PESF).⁵⁹ Se realizaron mejoras importantes a la página web

⁵⁸ <<http://www.principalglobalindicators.org>>.

⁵⁹ Actualmente, 29 países miembro del FMI están sujetos a evaluaciones obligatorias de estabilidad financiera, mientras que para el resto de la membresía el PESF es voluntario.

PGI como parte de la DGI en términos de cobertura y puntualidad. En la actualidad ofrece acceso a una base de datos en línea con datos históricos de periodos más largos seleccionados por el usuario y presentados en unidades comparables de medida (tasas de crecimiento, valores de índice o porcentaje del PIB). El trabajo para reforzar más la PGI continuará en la DGI-2 (anexo, recomendación 19).

A medida que surgen nuevos riesgos y que las relaciones entre instituciones, sectores y países se tornan más complejas, se vuelve relevante que los datos sean pormenorizados para evaluar tales riesgos. Esto trae aparejada la necesidad de que los compiladores de datos recaben información en escala micro para ayudar a satisfacer las necesidades de los usuarios. Las recomendaciones de la DGI (esto es, datos de bancos de importancia sistémica mundial, EBI, ECIC) respaldan la necesidad de datos más detallados.⁶⁰ Por otra parte, que los datos sean cada vez más pormenorizados eleva también los retos de compartir esta información ya sea dentro de las economías, entre fronteras o con organismos internacionales debido a preocupaciones sobre la confidencialidad. Esto potencialmente limita los beneficios más generales de recabar datos nuevos o existentes incluyendo algunos según la DGI. Si bien esto no es un problema fácil de resolver, la DGI-2 se enfoca en el tema de la confidencialidad alentando a las economías del G20 para que aumenten la distribución y accesibilidad de datos pormenorizados y que, de ser necesario, revisen los límites vigentes de confidencialidad (anexo, recomendación 20). Sería un avance positivo poder abordar los límites de confidencialidad y buscar cómo superarlos como parte de la DGI-2.

Cabe señalar que el sector privado está trabajando para mejorar la divulgación de datos de los bancos. A estos efectos, el Grupo de Trabajo para

la Mejora de la Divulgación (EDTF, por sus siglas en inglés), un grupo del sector privado de instituciones financieras establecido por el CEF, publicó un informe en 2012 que incluyó siete principios fundamentales para mejorar la divulgación de riesgos de los bancos.⁶¹

La necesidad de datos pormenorizados y en tiempo real para entender la interconexión y las diseminaciones entre países e instituciones quedó también de manifiesto en el Tercer Foro de Estadística del FMI realizado en Francfort en noviembre de 2015.⁶² Los participantes exigieron a las oficinas de estadísticas y a los responsables de política económica que establecieran nuevos marcos legales para respaldar el acceso a microdatos y al mismo tiempo preservar la confidencialidad.

2.7 IMPLICACIONES MÁS GENERALES DE LA DGI

Si bien el foco de la DGI han sido las economías del G20, abarca a un rango más amplio de economías, ya que la iniciativa se basa en marcos referenciales de estadísticas internacionales de aceptación general. En particular, el lineamiento metodológico resultado de la DGI es para todas las economías miembro del FMI. Además, la mayoría de las recomendaciones de la DGI se basa en diversas iniciativas sobre estadísticas que involucran a un grupo más grande de economías. Además, a medida que las economías ajenas al G20 ven las ventajas de esta iniciativa para su propio trabajo de política, aun sin tener ninguna presión de política como los FMCBG del G20, se esfuerzan por aplicar las recomendaciones de la DGI.

Es evidente que la calidad mejorada de la información en todo el mundo es esencial para garantizar una evaluación completa de las vinculaciones

⁶⁰ Las fuentes de datos micro, pormenorizados, pueden también mejorar la exactitud y el nivel de detalles de las macroestadísticas tradicionales. Bruno Tissot, *Closing Information Gaps at the Global Level –What Micro Data Can Bring*, <https://www.bis.org/ifc/events/ws_micro_macro/tissot_paper.pdf> .

⁶¹ El informe se enfocó en las divulgaciones de datos sobre el buen gobierno y la gestión de riesgos, suficiencia de capital, liquidez y financiamiento, riesgo de mercado, riesgo crediticio y otros riesgos. Ver *EDTF Principles and Recommendations for Enhancing the Risk Disclosures of Banks*, octubre de 2012, <http://www.fsb.org/wp-content/uploads/pr_121029.pdf> .

⁶² <<http://www.imf.org/external/np/seminars/eng/2015/statsforum/>> .

macrofinancieras en el ámbito mundial. Con esto en mente, una nota de referencia al documento del FMI, *Balance Sheet Analysis in Surveillance*,⁶³ proporcionó un listado completo de conjuntos de datos macro disponibles, relacionados con los balances, incluyendo su relevancia para la supervisión, y un resumen de datos disponibles para cada miembro del Fondo. Muchos de los conjuntos de datos referenciados fueron los abarcados por la DGI-2. Además, los funcionarios del FMI ofrecen cada vez más apoyo por medio de su asistencia técnica y capacitación para que los países miembro compilen estos conjuntos de datos.⁶⁴

3. CONCLUSIÓN

La DGI podría considerarse como una iniciativa general para abarcar una amplia variedad de marcos operativos estadísticos que están conectados entre sí, en apoyo a la meta común de un mejor entendimiento de los mercados e instrumentos financieros y de arrojar luz sobre la interconectividad.

No es tarea fácil proporcionar una amplia cantidad de información, como es la intención de la DGI, de manera estandarizada, frecuente y oportuna, y al mismo tiempo garantizar que los datos sean confiables, de alta calidad y que reflejen adecuadamente las circunstancias económicas cambiantes. Esto no puede lograrse de un día al otro, pero sí es posible con el paso del tiempo y con un esfuerzo en escala mundial.

Se requiere también de un apoyo de alto nivel, que incluye garantizar los recursos. Desde la crisis financiera mundial, todas las partes relevantes han hecho importantes esfuerzos para garantizar que los responsables de políticas económicas tengan acceso a la información relacionada con la DGI como componente clave de su herramienta. A futuro, es importante tener la capacidad de cosechar los frutos de las inversiones hechas en la DGI y así mantener el ritmo de trabajo y continua coordinación entre todos los jugadores de la economía mundial.

⁶³ FMI, *Balance Sheet Analysis in Fund Surveillance –Reference Note*, julio de 2015, <<http://www.imf.org/external/np/pp/eng/2015/071315.pdf>>.

⁶⁴ Ver *IMF Statistics Department at a Glance*, 2015, <<http://www.imf.org/external/np/sta/pdf/aglance.pdf>>.

RECOMENDACIONES DE LA DGI-2

Recomendación II.1: Mandato de la DGI

Las economías del G20 han de recabar regularmente estadísticas financieras y económicas comparables y de alta calidad de acuerdo con los estándares internacionales y divulgar estas estadísticas de manera oportuna. La IAG ha de coordinar y vigilar la aplicación de las recomendaciones de la DGI, y promover el sitio web de los principales indicadores mundiales (PGI) como una base de datos de referencial mundial. El personal del CEF y del FMI han de brindar actualizaciones anuales sobre el progreso a los gobernadores de banco centrales y ministros de finanzas del G20.

VIGILANCIA DE LOS RIESGOS EN EL SECTOR FINANCIERO

Recomendación II.2: Indicadores de solidez financiera

Las economías del G20 han de informar los siete indicadores de solidez financiera (ISF) que se esperan de las economías que se adhieren al NEDD Plus, con frecuencia trimestral. Se insta a las economías del G20 a informar la lista central y ampliada de ISF, con especial atención en las empresas financieras (no bancarias). El FMI coordinará el trabajo y vigilará el progreso.

Recomendación II.3: Medidas de concentración y distribución (CDM, por sus siglas en inglés)

El FMI ha de investigar la posibilidad de la recolectar de manera regular las medidas de concentración y distribución para los ISF. Las economías del G20 respaldarán el trabajo del FMI.

Recomendación II.4: Datos para las instituciones financieras de importancia sistémica en el ámbito mundial (G-SIFI)

Las economías del G20 han de respaldar el International Data Hub del BPI para garantizar la recolección regular y la distribución apropiada de los datos sobre bancos de importancia sistémica en el ámbito mundial (G-SIB). Además, corresponde al CEF, tras previa consulta con el FMI y órganos supervisores relevantes, investigar la posibilidad de una plantilla de datos comunes para las instituciones financieras no bancarias de importancia sistémica empezando por las aseguradoras. El trabajo tomará debida cuenta de la confidencialidad y de los aspectos legales.

Recomendación II.5: Banca paralela

Las economías del G20 mejorarán la recolección de datos del sistema de la banca paralela contribuyendo al proceso de seguimiento del CEF, incluso mediante la provisión de datos de cuentas sectoriales. El CEF ha de trabajar en mejoras adicionales del marco conceptual y el desarrollo de estándares y procesos para recabar y conjuntar datos fiables en escala mundial.

Recomendación II.6: Derivados

El BPI revisará los datos de derivados recabados para las estadísticas bancarias internacionales (EBI) y la encuesta semestral de estadísticas de derivados extrabursátiles (OTC), y el CEF ha de formular un mecanismo para agregar y compartir en escala mundial los datos de derivados OTC de los registros de operaciones. Las economías del G20 respaldarán este trabajo según corresponda.

Recomendación II.7: Estadísticas de títulos

A las economías del G20 corresponde aportar con frecuencia trimestral datos de emisión de títulos de deuda al BPI conforme al *Handbook on Security Statistics*, comenzando por el sector, la moneda, el tipo de tasa de interés, el plazo original y, si fuera posible, el mercado de emisiones. El informe sobre acervos de títulos de deuda y los datos de contrapartida sectorial prescritos para las economías que se adhieren las NEDD Plus sería una meta a más largo plazo. El BPI, con la asistencia del Grupo de Trabajo sobre Bases de Datos de Títulos, vigilará la recolección regular y la fiabilidad de los datos de títulos de deuda.

VULNERABILIDADES, INTERCONEXIONES
Y DISEMINACIONES

Recomendación II.8: Cuentas sectoriales

A las economías del G20 corresponde recabar y divulgar, con frecuencias trimestral y anual, los

flujos de cuentas sectoriales y datos de balances generales, con base en la plantilla acordada internacionalmente, incluyendo los datos para el otro sector de empresas financieras (no bancarias), y desarrollar matrices de contrapartida tanto para las transacciones como para los acervos para respaldar el análisis de los balances. El IAG, en colaboración con el Grupo de Trabajo Intersecretarial sobre Cuentas Nacionales, fomentará y vigilará el progreso de las economías del G20.

Recomendación II.9: Información sobre la distribución entre hogares

El IAG, en colaboración con las economías del G20, ha de fomentar la producción y divulgación de información sobre la distribución del ingreso, el consumo, el ahorro y la riqueza, para el sector de hogares. A la OCDE corresponde coordinar el trabajo en estrecha colaboración con el Eurostat y el Banco Central Europeo.

Recomendación II.10: Posición de inversión internacional (PII)

A las economías del G20 corresponde proveer datos trimestrales de la PII al FMI, que concuerde con el Manual de Posición de Inversión Internacional y Balance de Pagos, sexta edición (BPM6) y que incluya las mejoras tales como la composición por moneda y la identificación separada de otras empresas financieras (no bancarias), introducidas en dicho Manual. El FMI vigilará la generación de informes y la fiabilidad de datos sobre la PII, y considerará la identificación separada de empresas no financieras, en colaboración con el Comité de Estadísticas de la Balanza de Pagos del FMI (BOP-COM, por sus siglas en inglés).

Recomendación II.11: Estadísticas bancarias internacionales (EBI)

Las economías del G20 deberán proveer las estadísticas bancarias internacionales mejoradas del

BPI. El BPI trabajará con todos los países que producen informes para acabar con las deficiencias en la información de las EBI, revisará las opciones para mejorar la fiabilidad entre las EBI consolidadas y los datos de supervisión, y apoyará los esfuerzos para hacer que los datos tengan mayor disponibilidad.

Recomendación II.12: Estudio Coordinado sobre la Inversión de Cartera (ECIC)

Las economías del G20 proporcionarán, con frecuencia semestral, datos del Estudio Coordinado sobre Inversión de Cartera del FMI, incluidos el cuadro del sector de tenedores y, de preferencia, también el del sector de emisores no residentes. Al FMI corresponde vigilar la generación regular de informes y la solidez de los datos, continuar mejorando la cobertura de centros financieros significativos e investigar la posibilidad de generación de informes trimestrales.

Recomendación II.13: Estudio Coordinado sobre la Inversión Directa (ECID)

Las economías del G20 deberán participar y mejorar su generación de informes para el Estudio Coordinado sobre la Inversión Directa del FMI, tanto de la inversión directa entrante como de la saliente. Al FMI corresponde vigilar el progreso.

Recomendación II.14: Exposiciones transfronterizas de las empresas no bancarias

El IAG mejorará la fiabilidad y divulgación de datos de las exposiciones transfronterizas de las empresas no bancarias, incluyendo las de las filiales extranjeras y los financiamientos intragrupal, para un mejor análisis de los riesgos y las vulnerabilidades provenientes de tales exposiciones, incluyendo los descalces de divisas. El trabajo se concentrará en la recolección de datos por parte del BPI y del FMI, y en la elaboración del marco de

referencia de la OCDE para la inversión extranjera directa. A las economías del G20 toca apoyar el trabajo del IAG.

Recomendación II.15: Estadísticas de finanzas gubernamentales

Las economías del G20 divulgarán datos trimestrales del gobierno federal conforme al *Manual de estadísticas de finanzas públicas 2014* (MEFP, 2014). Se exhorta a que se adopte la contabilidad de lo devengando en las economías del G20. Al FMI corresponde la elaboración regular de informes y la divulgación de datos oportunos, comparables y de alta calidad sobre las finanzas públicas.

Recomendación II.16: Estadísticas de deuda del sector público

Las economías del G20 proporcionarán datos de deuda pública general con amplia cobertura de instrumentos de las bases de datos de deuda del sector público del Banco Mundial/FMI/OCDE. El Banco Mundial coordinará el trabajo.

Recomendación II.17: Precios de las propiedades residenciales

Las economías del G20 publicarán índices de precios de las propiedades residenciales acordes con el *Manual del índice de precios de inmuebles residenciales* y aportará estos datos a las organizaciones internacionales relevantes, entre ellas al BPI, Eurostat y la OCDE. El IAG en colaboración con el Grupo Intersecretarial de Estadísticas de Precios (IWGPS, por sus siglas en inglés) trabajará en un conjunto de índices de precios finales de propiedades residenciales; fomentará la producción de largas series temporales; elaborará una lista de otros indicadores relacionados con la vivienda; y divulgará datos finales sobre los precios de las propiedades residenciales en la página web PGI.

Recomendación II.18: Precios de las propiedades comerciales

Al IAG, en colaboración con el Grupo Intersecretarial de Estadísticas de Precios, toca mejorar los lineamientos metodológicos para compilar índices de precios de las propiedades comerciales y fomentar la divulgación de datos sobre los precios de las propiedades comerciales por medio de la página web del BPI.

COMUNICACIÓN DE ESTADÍSTICAS OFICIALES

Recomendación II.19: Cooperación y comunicación de datos internacionales

El IAG promoverá la cooperación entre organizaciones internacionales en favor de datos internacionales mejorados y apoyará la transmisión oportuna y estandarizada de datos mediante los formatos acordados internacionalmente (por ejemplo,

SDMX), para reducir así la carga sobre las economías que informan datos y promover el alcance de usuarios. Al IAG corresponde continuar trabajando con las economías del G20 para presentar datos nacionales oportunos y sólidos en la página PGI y en las páginas de las organizaciones internacionales participantes.

Recomendación II.20: Promoción de la distribución de datos

El IAG y las economías del G20 promoverán y fomentarán el intercambio de datos y metadatos entre las economías del G20 y dentro de ellas, y con organismo internacionales, para mejorar la calidad (esto es, la fiabilidad) de los datos y la disponibilidad para el uso de política. Asimismo, se insta a las economías del G20 a mejorar la distribución y accesibilidad de datos pormenorizados, si fuera necesario revisando las limitaciones vigentes de confidencialidad.

BIBLIOGRAFÍA

- Avdjiev, McGuire y Wooldridge (2015), “Enhanced Data to Analyse International Banking”, *BIS Quarterly Review*, septiembre, <http://www.bis.org/publ/qtrpdf/r_qt1509f.pdf>.
- Borio y Drehmann (2009), “Assessing the Risk of Banking Crises –Revisited”, *BIS Quarterly Review*, marzo, <http://www.bis.org/publ/qtrpdf/r_qt0903e.htm>.
- BPI (2015), *Semiannual OTC Derivatives Statistics*, <<http://www.bis.org/statistics/derstats.htm>>.
- BPI, BCE y FMI (2015), *Handbook on Securities Statistics*, <<http://www.imf.org/external/np/sta/wgsd/hbook.htm>>.
- BPI/CEF/FMI, (2015), *Reports on Foreign Currency Exposures to the G20*, <<http://www.financialstabilityboard.org/wp-content/uploads/Work-on-Foreign-Currency-Exposures.pdf>>.
- CEF (2012), *EDTF Principles and Recommendations for Enhancing the Risk Disclosures of Banks*, <http://www.fsb.org/wp-content/uploads/pr_121029.pdf>.

- CEF (2015), *Corporate Funding Structures and Incentives*, <<http://www.financialstabilityboard.org/2015/09/corporate-funding-structures-and-incentives/>>
- CEF (2014), *FSB Data Gaps Initiative –A Common Data Template for Global Systemically Important Banks*, <http://www.financialstabilityboard.org/wp-content/uploads/r_140506.pdf>.
- CEF (2015), *Global Shadow Banking Monitoring Report*, <<http://www.financialstabilityboard.org/2015/11/global-shadow-banking-monitoring-report-2015/>>.
- CEF (2013), *Strengthening Oversight and Regulation of the Shadow Banking: An Overview of Policy Recommendations, Progress Report to G20 Ministers and Governors*, agosto, <http://www.financialstabilityboard.org/wp-content/uploads/r_120420c.pdf?page_moved=1>.
- CEF (2015), *Transforming Shadow Banking into Resilient Market-based Finance*, <<http://www.financialstabilityboard.org/wp-content/uploads/FSB-Standards-for-Global-Securities-Financing-Data-Collection.pdf>>.
- CEF/FMI (varios años), *Progress Reports of the G20 Data Gaps Initiative*, <<http://www.imf.org/external/ns/cs.aspx?id=290>>.
- CEF/FMI (2015), *Sixth Progress Report on the Implementation of the G20 Data Gaps Initiative*, <<http://www.imf.org/external/np/g20/pdf/2015/6thprogressrep.pdf>>.
- CEF/FMI (2009), *The Financial Crisis and Information Gaps: Report to the G20 Finance Ministers and Central Bank Governors*, <<http://www.imf.org/external/np/g20/pdf/102909.pdf>>.
- Costa Navajas, Matias, y Thegeya Aaron (2013), *Financial Soundness Indicators and Banking Crises*, núm. WP/13/263, <<https://www.imf.org/external/pubs/ft/wp/2013/wp13263.pdf>>.
- Dembiermont, Scatigna, Szemere y Tissot (2015), “A New Database on General Government Debt”, *BIS Quarterly Review*, septiembre, <http://www.bis.org/publ/qtrpdf/r_qt1509g.pdf>.
- Dembiermont, Christian (2015), “Measuring Property Prices: The BIS Contribution”, <https://www.bis.org/ifc/events/ifc_isi_2015/432_dembiermont_paper.pdf>.
- Dippelsman Robert, Claudia Dziobek y Carlos A. Gutiérrez Mangas (2012), *What Lies Beneath: The Statistical Definition of Public Sector Debt*, IMF Staff Discussion Notes, núm. 12/09, <<https://www.imf.org/external/pubs/ft/sdn/2012/sdn1209.pdf>>.
- Errico, Harutyunyan, Loukoianova, Walton, Korniyenko, Amidžić, AbuShanab y Hyun Song Shin (2014), *Mapping the Shadow Banking System through a Global Flow of Funds Analysis*, IMF Working Paper, núm. 14/10, <<http://www.imf.org/external/pubs/ft/wp/2014/wp1410.pdf>>.
- Eurostat (2013), *Handbook on Residential Property Prices Indices*, <<http://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-RA-12-022>>.
- Fender, Ingo, y Patrick McGuire (2010), *Bank Structure, Funding Risk and the Transmission of Shocks across Countries: Concepts and Measurement*, <http://www.bis.org/publ/qtrpdf/r_qt1009h.htm>.
- FMI (2010), *Expanding Surveillance to Require Mandatory Financial Stability Assessments of Countries with Systemically Important Financial Sectors*, <<https://www.imf.org/external/np/sec/pr/2010/pr10357.htm>>.
- FMI (2011), *Twenty-fourth Meeting of the IMF Committee on Balance of Payments Statistics: Enhancements to the Coordinated Portfolio Investment Survey*, <<https://www.imf.org/external/pubs/ft/bop/2011/11-12.pdf>>.
- FMI (2013), *Modifications to the Current List of Financial Soundness Indicators*, <<https://www.imf.org/external/np/pp/eng/2013/111313.pdf>>.
- FMI (2013), *Review of the Implementation of Government Finance Statistics to Strengthen Fiscal Analysis*, <<https://www.imf.org/external/pubs/ft/gfs/manual/pdf/021214.pdf>>.
- FMI (2014), *Proceedings of the Second Statistical Forum*, <<http://www.imf.org/external/np/seminars/eng/2014/statsforum/>>.
- FMI (2014), *Triennial Surveillance Review –Overview Paper*, <<http://www.imf.org/external/np/pp/eng/2014/073014.pdf>>.

- FMI (2014), *Triennial Surveillance Review—External Study, Risks and Spillovers*, <<https://www.imf.org/external/np/pp/eng/2014/073014e.pdf>>.
- FMI (2014), *Triennial Surveillance Review—Managing Director’s Action Plan for Strengthening Surveillance*, <<https://www.imf.org/external/np/pp/eng/2014/112114.pdf>>.
- FMI (2015), *Balance Sheet Analysis in Fund Surveillance*, <<http://www.imf.org/external/np/pp/eng/2015/061215.pdf>>.
- FMI (2015), *Balance Sheet Analysis in Fund Surveillance—Reference Note*, <<http://www.imf.org/external/np/pp/eng/2015/071315.pdf>>.
- FMI (2015), *Proceedings of the Third Statistical Forum*, <<http://www.imf.org/external/np/seminars/eng/2015/statsforum/>>.
- FMI (2015), *Global Financial Stability Report*, <http://www.imf.org/External/Pubs/FT/GFSR/2015/02/pdf/c3_v2.pdf>.
- FMI (2015), *Proceedings of the Meeting of the IMF Government Finance Statistics Advisory Committee (GFSAC)*, <<http://www.imf.org/external/pubs/ft/gfs/gfsac/meetings/2015>>.
- FMI (2015), “Progress with Globally Comparable Public Sector Debt Statistics”, document para el Government Finance Statistics Advisory Committee (GFSAC), <<http://www.imf.org/external/pubs/ft/gfs/gfsac/meetings/2015/pdf/1515a.pdf>>.
- FMI (2015), *Statistics Department at a Glance*, <<http://www.imf.org/external/np/sta/pdf/aglance.pdf>>.
- FMI (varios años), *IMF’s Work on Income Inequality*, <<http://www.imf.org/external/np/fad/in-equality/index.htm>>.
- G20 (2008), *Declaration Summit on Financial Markets and the World Economy*, 15 de noviembre, <https://g20.org/wp-content/uploads/2014/12/Washington_Declaration_0.pdf>, <<http://www.imf.org/external/np/g20/pdf/2014/5thprogressrep.pdf>>.
- Hattori, Masazumi, y Takáts Előd (2015), *The Role of Debt Securities Markets*, <https://www.bis.org/publ/bppdf/bispap83c_rh.pdf>.
- Heath, Robert (2013), *Why Are the G20 Data Gaps Initiative and the NEDD Plus Relevant for Financial Stability Analysis?*, núm. WP/13/6, <<http://www.imf.org/external/pubs/ft/wp/2013/wp1306.pdf>>.
- Heath, Robert (2015), *What Has Capital Flow Liberalization Meant for Economic and Financial Statistics?*, núm. WP/15/88, <<http://www.imf.org/external/pubs/ft/wp/2015/wp15/88.pdf>>.
- IAG (2015), *Consolidation and Corporate Groups: An Overview of Methodological and Practical Issues*, IAG Reference Document, <<http://www.bis.org/ifc/publ/iagrefdoc-oct15.pdf>>.
- IAIS (2011), *Insurance and Financial Stability*, <<http://iaisweb.org/index.cfm?event=openFile&nodeId=34041>>.
- IAIS (2013), “Global Systemically Important Insurers: Initial Assessment Methodology”, <<http://iaisweb.org/index.cfm?event=showHomePage&persistId=90B48F72155D89A406A8E62C456775A9>>.
- OCDE (2013), *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*, <<http://www.oecd.org/statistics/icw-framework.htm>>.
- OCDE (2013), *OECD Guidelines for Micro Statistics on Household Wealth*, <<http://www.oecd.org/statistics/guidelines-for-micro-statistics-on-household-wealth.htm>>.
- Tissot, Bruno (2015), *Closing Information Gaps at the Global Level—What Micro Data Can Bring*, <https://www.bis.org/ifc/events/ws_micro_macro/tissot_paper.pdf>.

MODELOS DEL RIESGO DE CRÉDITO

Somnath Chatterjee

INTRODUCCIÓN

El crédito es dinero proporcionado por un acreedor a un prestatario (conocido también como obligado, ya que tiene una obligación de pago). El riesgo de crédito es el asociado al incumplimiento de un pago contratado. Se da por sentado que los mercados le ponen un precio a este riesgo. Por ende, tal precio se incluye en el precio de compra del mercado para un pago contratado. La parte del precio que corresponde al riesgo de crédito es el diferencial de crédito. La función de un modelo típico de riesgo de crédito es tomar como insumos las condiciones de la economía en general y de una empresa en particular, y generar un resultado que es el diferencial de crédito.

El motivo para crear modelos de riesgo de crédito radica en la necesidad de calcular cuánto capital económico es necesario para sustentar las actividades de toma de riesgo de un banco. Los requerimientos de capital mínimo se han coordinado en el plano internacional desde el Acuerdo de Basilea de 1998. Conforme a Basilea I, los activos de un banco se asignaban por regla general a una de cuatro categorías amplias de riesgo, cada una con una *ponderación del riesgo* que oscilaba entre el 0% y el 100%. Una cartera de créditos a empresas, por ejemplo, recibía una ponderación de riesgo del 100%, mientras que las hipotecas, consideradas más seguras, recibirían una ponderación de riesgo más favorable (un 50%). Así, el capital mínimo se establecía en proporción a la suma ponderada de estos activos.

Requerimiento de capital mínimo = $8\% \times \sum$ activos ponderados por riesgo

Traduce y publica el CEMLA, con la debida autorización, el Handbook, núm. 34, del Centre for Central Banking Studies, Banco de Inglaterra. El autor agradece a Abbie McGillivray por el diseño del formato de este manual. Las opiniones expresadas en este son las del autor y no necesariamente las del Banco de Inglaterra. <Somnath.Chatterjee@bankofengland.co.uk>.

Con el tiempo, este cálculo fue objeto de críticas por no ser tan detallado como para consignar la distribución transversal del riesgo. A todos los créditos hipotecarios, por ejemplo, se les asignaba el mismo requerimiento de capital, sin tomar en cuenta el perfil de riesgo subyacente del prestatario (como el coeficiente de préstamo a valor o de deuda a ingreso). Esto despertó la preocupación de que la normatividad estimulara el *traslado de riesgo*. Como el riesgo no se estaba justipreciando, entonces—se argumentaba— los bancos tenían un incentivo para retener en su balance únicamente las exposiciones más elevadas al riesgo, pues esta probablemente ofrecería el rendimiento esperado más alto.

En respuesta, Basilea II adoptó un enfoque mucho más minucioso en la ponderación del riesgo. Las técnicas para gestión del riesgo de crédito utilizadas conforme a Basilea II son de dos tipos:

- estandarizadas, que sencillamente implican distribuir a los obligados en categorías, sin considerar sus riesgos reales de crédito, y recurren a calificaciones externas del riesgo;
- basadas en calificaciones internas, es decir, se permite a los bancos utilizar sus *modelos internos* para calcular el requerimiento de capital reglamentario para el riesgo de crédito.

Estas técnicas se formularon para llegar a los activos ponderados por riesgo (APR), el denominador de cuatro coeficientes de capitalización fundamentales: el capital total, el capital de nivel 1, el capital básico de nivel 1 y el capital accionario de nivel 1. Conforme a Basilea II, los bancos que siguen la modalidad de APR pueden computar los requerimientos de capital con base en una fórmula que se acerca al modelo de Vasicek para el riesgo de crédito de una cartera. El procedimiento de Vasicek se describe en la siguiente sección.

El capital mínimo obligatorio no cambió con Basilea III, pero se introdujeron reglas más estrictas para garantizar que su calidad fuera suficiente. Ahora el requisito para el capital accionario de

categoría 1 es un 4.5%. También incrementó la cantidad de capital al introducir colchones de capital utilizable en vez de un capital mínimo (ver Comité de Basilea, 2010). Basilea III aclaró la definición de capital, es decir, el numerador del coeficiente de capital. Sin embargo, no buscó modificar radicalmente el método de Basilea II, basado en el riesgo, para medir los activos ponderados por riesgo, es decir, el denominador del coeficiente de capital; por lo tanto, la arquitectura del régimen de capital ponderado por riesgo prácticamente no fue modificada. Basilea III busca mejorar el método estandarizado de administrar el riesgo de crédito de varias maneras, entre otras, fortaleciendo el vínculo entre el método estandarizado y el basado en calificaciones internas.

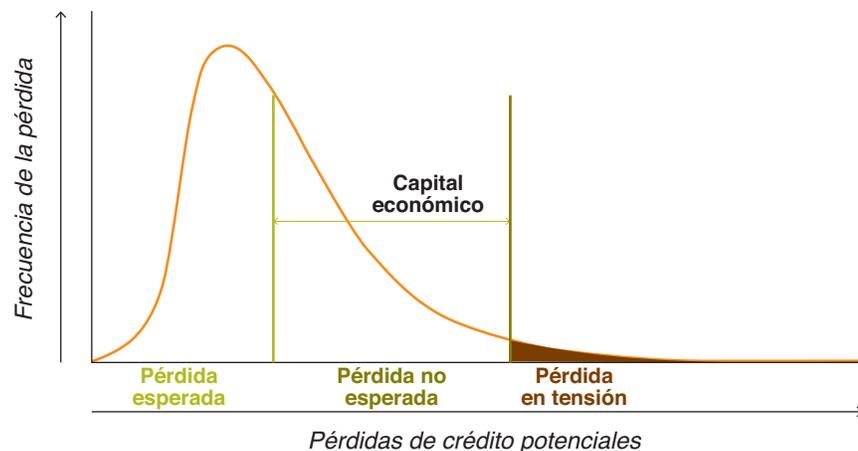
Basilea III busca mejorar el método estandarizado de administrar el riesgo de crédito.

1. ASIGNACIÓN DEL CAPITAL ECONÓMICO

Los bancos emplean un procedimiento analítico cuando calculan la cantidad de capital económico necesaria para sustentar las actividades con riesgo de crédito. El procedimiento que relaciona el capital económico que en general se requiere para el riesgo de crédito con la función de densidad de probabilidad (FDP) de pérdidas crediticias de la cartera, lo que también se conoce como distribución de pérdidas de una cartera de créditos. La gráfica 1 muestra esta relación. Aunque los distintos métodos de elaboración de modelos varían, todos ellos consideran la estimación como una función de densidad de probabilidad.

Gráfica 1

DISTRIBUCIÓN DE LA PÉRDIDA DE UNA CARTERA DE CRÉDITO



1.1 FUNCIÓN DE DENSIDAD DE LA PROBABILIDAD DE PÉRDIDAS CREDITICIAS

Los mecanismos para asignar capital económico contra el riesgo de crédito por lo general suponen que es posible obtener la forma de la FDP mediante distribuciones que podrían parametrizarse con la media y la desviación estándar de las pérdidas de cartera. La gráfica 1 muestra que el riesgo de crédito tiene dos componentes. Primero, la pérdida esperada (PE), la cantidad de pérdida crediticia que el banco esperaría sufrir en su cartera de créditos durante cierto horizonte elegido. Esto podría verse como el costo normal de hacer negocios, cubierto mediante políticas de provisiones y precios. Segundo, los bancos expresan el riesgo de la

cartera con una medida de la pérdida no esperada (PNE). Se aparta capital para compensar la PNE y, conforme al método basado en calificaciones internas, la cantidad de capital reglamentario depende únicamente de la PNE. La desviación estándar, que muestra la desviación promedio de las pérdidas esperadas, es una medida de la pérdida inesperada que se usa con frecuencia. El área bajo la curva en la gráfica 1 es igual a 100%. La curva muestra que las pérdidas pequeñas en torno a la PE o ligeramente inferiores a esta ocurren con más frecuencia que las pérdidas cuantiosas.

La probabilidad de que las pérdidas superen la suma de la PE y la PNE —es decir, la probabilidad de que el banco no pueda cumplir con sus obligaciones crediticias mediante sus utilidades y capital— equivale al área sombreada en el lado derecho de la curva, que representa la pérdida por tensión financiera. Un 100% menos esta probabilidad es lo que se conoce como el valor en riesgo (VaR) en este nivel de confianza. Si el capital se fija de acuerdo con la brecha entre la PE y el VaR, y si la PE está cubierta por reservas o ingresos, entonces la probabilidad de que el banco permanezca solvente en un horizonte de un año es igual al nivel de confianza. Conforme a Basilea II, el capital se fija para mantener un nivel de confianza fijo con fines de supervisión. Este nivel de confianza se fija en un 99.9%, es decir, se espera que una institución sufra pérdidas que rebasen su capital una vez cada 1,000 años. Las lecciones aprendidas de la crisis financiera mundial de 2007-2009 apuntarían a que la pérdida por tensión financiera es la pérdida inesperada posible contra la cual se considera demasiado caro mantener capital. A los organismos supervisores les preocupa particularmente la cola de la distribución de pérdidas y el punto donde los bancos deberían fijar el límite entre la pérdida inesperada y la pérdida por tensión financiera. Para más detalles sobre las distribuciones de pérdidas en escenarios de tensión, ver Haldane *et al.* (2007).

Un banco debe decidir en qué horizonte temporal va a evaluar el riesgo de crédito. En los acuerdos de Basilea, el horizonte para todas las categorías

de activos es un año. La pérdida esperada de una cartera se presume es igual a la proporción de obligados que podrían incumplir en un marco temporal dado, multiplicada por la exposición pendiente en el momento de incumplimiento y, de nuevo, por la pérdida dado el incumplimiento, que representa la proporción de la exposición que no se recuperará después del incumplimiento. Conforme al método basado en calificaciones internas de Basilea II, la probabilidad de incumplimiento (PI) por grado de calificación es el porcentaje promedio de obligados que incumplirán en el plazo de un año. La exposición en el momento del incumplimiento (EAI) es el monto pendiente si el prestatario incumple. La pérdida dado el incumplimiento (PDI) representa la proporción de la EAI que no se recuperará tras el incumplimiento. Suponiendo un valor uniforme de PDI para una cartera dada, la pérdida esperada puede calcularse como la suma de las PE individuales en la cartera (ecuación 1.1).

$$1.1 \quad PE = \sum_{i=1}^N PI_i PDI_i EAI_i$$

A diferencia de la PE, la pérdida no esperada total no es un agregado de las PNE individuales; sino depende, más bien, de las correlaciones de pérdidas entre todos los préstamos de la cartera. La desviación de las pérdidas con respecto a la PE normalmente se mide mediante la desviación estándar de la variable de pérdidas (ecuación 1.2). La PNE, o la desviación estándar de pérdidas por créditos de la cartera puede descomponerse en la contribución de cada uno de los servicios de financiamiento individuales:

$$1.2 \quad PNE = \sum_{i=1}^N \sigma_i \rho_i$$

donde σ_i denota una desviación estándar independiente de las pérdidas crediticias por el i -ésimo financiamiento y ρ_i denota la correlación entre las pérdidas crediticias en el i -ésimo financiamiento y aquellas en la cartera general. El parámetro ρ_i refleja los efectos de correlación/diversificación del i -ésimo financiamiento con los otros instrumentos

en la cartera de créditos del banco. Si lo demás no cambia, correlaciones más elevadas entre los instrumentos de crédito –representadas por un ρ_i más alto– lleva a una desviación estándar más elevada de las pérdidas crediticias para la cartera en su conjunto.

El acuerdo de Basilea II ha especificado los valores de correlación entre activos para distintas categorías de activos (Comité de Basilea, 2006). Pero la base teórica para calcular la PNE conforme al método basado en calificaciones internas de Basilea II proviene del modelo de Vasicek (2002) para el valor de la cartera de préstamos. Ver Comité de Basilea (2005) para una explicación más detallada de la formulación de dicho método. Un problema con el método basado en calificaciones internas es que implica una dependencia excesiva respecto de los propios modelos internos de los bancos para calcular los requerimientos de capital, dado que el método estandarizado no brinda otra alternativa creíble que consigne los riesgos en las carteras de operación de los bancos. Sin embargo, se ha descubierto que los modelos internos de los bancos dan por resultado ponderaciones del riesgo muy diferentes para carteras comunes de activos bancarios. Parte del problema de evaluar los cálculos de APR de los bancos radica en distinguir las diferencias que se derivan del riesgo de la cartera y de la calidad de los activos de aquellas que provienen de diferencias en los modelos. Para identificar las diferencias entre los modelos internos de los bancos, los organismos de regulación han realizado varios ejercicios en los que los bancos aplicaron modelos internos para calcular los principales parámetros de riesgo a los activos de una cartera hipotética. Esto aseguró que las diferencias en las ponderaciones calculadas para el riesgo se redujeran a qué modelo utilizaban los bancos y no a diferencias en el riesgo de las carteras evaluadas. La sección a continuación comenta la metodología de Vasicek (2002) para calcular la distribución conjunta de las pérdidas para una cartera de exposiciones bancarias.

1.2 CÁLCULO DE LA DISTRIBUCIÓN CONJUNTA DE PÉRDIDAS UTILIZANDO EL MODELO DE VASICEK

El modelo de Vasicek (2002) supone que el valor del activo de un determinado obligado lo proporciona el efecto combinado de un factor sistemático y uno idiosincrásico. Supone una estructura de incumplimiento gaussiana equicorrelacionada. Es decir, cada obligado i incumple si cierta variable aleatoria X_i cae por debajo de un umbral, y estas X_i son todas normales y equicorrelacionadas. El valor de activo del obligado i -ésimo en el tiempo t está entonces proporcionado por:

$$1.3 \quad X_{it} = S_t \sqrt{\rho} + Z_{it} \sqrt{1-\rho},$$

donde S y Z son, respectivamente, el componente sistemático y el idiosincrásico y puede demostrarse que ρ es la correlación de activos entre dos obligados diferentes. Ver el recuadro 1 para más detalles sobre el modelo de Vasicek de cartera de préstamos. Aquí X, Z_1, Z_2, \dots, Z_n son variables normales estándar mutuamente independientes. El modelo de Vasicek utiliza tres insumos para calcular la probabilidad de incumplimiento (PI) de una categoría de activos. Una variable es la PI durante el ciclo (PI_DEC) específica para esa categoría. Otros insumos son un factor común de la cartera, como un índice económico en el intervalo $(0, T)$ proporcionado por S . La tercera variable de entrada es la correlación de activos, ρ . Luego el término $S_t \sqrt{\rho}$ es la exposición de la compañía al factor sistemático, mientras que el término $Z_{it} \sqrt{1-\rho}$ representa su riesgo idiosincrásico.

Una sencilla condición de umbral determina si el obligado i incumple o no.

Incumple si $X_i < c$; donde c aparecerá como una función de PI_DEC.

RECUADRO 1

MODELO DE VASICEK DEL VALOR DE UNA CARTERA DE PRÉSTAMOS

Vasicek aplicó a los valores de activos de las compañías lo que se había convertido en el modelo estándar de movimiento browniano geométrico. Expresado como una ecuación diferencial estocástica,

$$dA_i = \mu_i A_i dt + \sigma_i A_i dx_i;$$

donde A_i es el valor de los activos de la i -ésima empresa, μ_i y σ_i son el índice de deriva y la volatilidad de ese valor, y x_i es un proceso de Wiener o movimiento browniano, es decir, un paseo aleatorio en tiempo continuo en el cual el cambio sobre cualquier periodo temporal finito se distribuye normalmente con una media de cero y una varianza igual a la longitud del periodo, y los cambios en periodos separados son independientes uno del otro. Al resolver esta ecuación diferencial estocástica se obtiene el valor de los activos de la i -ésima empresa en el tiempo T como:

$$1 \quad A_i(T) = e^{A(0) + \mu_i T - \frac{1}{2} \sigma_i^2 T + \sigma_i \sqrt{T} X_i}$$

La i -ésima empresa incumple cuando $A_i(T) < B$, por lo que la probabilidad de tal suceso es:

$$2 \quad P[A_i(T) < B_i] = P[X_i < c_i] = N(c_i) = p^*,$$

donde c_i se deriva con facilidad de la ecuación 1 y N es la FDP normal acumulativa. Es decir, el incumplimiento de un único obligado ocurre si el valor de una variable aleatoria normal cae por debajo de cierto c_i ,

La correlación entre incumplimientos se introduce al suponer correlación en los procesos. La correlación entre incumplimientos se introduce suponiendo una correlación en los procesos A_i y, por lo tanto, en los valores terminales $A_i(T)$. En particular, se supone que las X_i en la ecuación 1 están correlacionados en pares de acuerdo con el factor ρ .

Por ser normales y equicorrelacionadas, cada variable aleatoria X_i puede entonces representarse como la suma de dos otras variables aleatorias: una común a las empresas y otra idiosincrásica:

$$X_i = S\sqrt{\rho} + Z_i\sqrt{1-\rho},$$

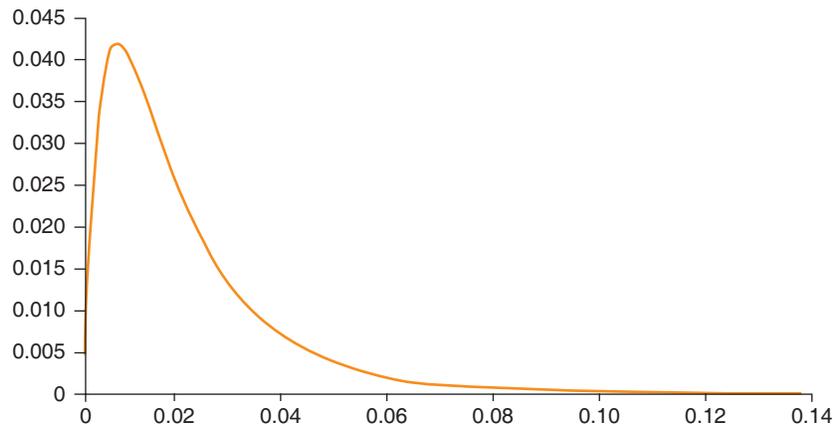
con $S \sim N(0,1)$, $Z_i \sim N(0,1)$. Por lo tanto, la probabilidad de incumplimiento del obligado i también puede escribirse como:

$$3 \quad \begin{aligned} P[A_i(T) < B_i] &= P[X_i < c_i] \\ &= P[X_i < N^{-1}(p^*)] \\ &= P\left[(S\sqrt{\rho} + Z_i\sqrt{1-\rho}) < N^{-1}(p^*)\right] \\ &= P\left[Z_i < \frac{c_i - S\sqrt{\rho}}{\sqrt{1-\rho}}\right] = N\left(\frac{N^{-1}(p^*) - S\sqrt{\rho}}{\sqrt{1-\rho}}\right). \end{aligned}$$

p^* en la ecuación 2 es la pérdida promedio durante el ciclo. $P[\dots]$ en la ecuación 3 es la pérdida sujeta a las condiciones crediticias S . La proporción de créditos en la cartera que caen en incumplimiento la proporciona la siguiente FDP:

$$4 \quad \begin{aligned} P[p(S) \leq x] &= P[S \geq p^{-1}(x)] \\ &= N(-p^{-1}(x)) = N\left(\frac{\sqrt{1-\rho}N^{-1}(x) - N^{-1}(\rho)}{\sqrt{\rho}}\right). \end{aligned}$$

DISTRIBUCIÓN DE LA PÉRDIDA NO CONDICIONAL



El modelo de Vasicek puede interpretarse en el contexto de un mecanismo desencadenante que resulta útil para elaborar modelos del riesgo de crédito. Una sencilla condición de umbral determina si el obligado i incumple o no.

La integración de S a la ecuación 1.3 denota la probabilidad incondicional de impago por parte de p^* (esta es la PI_DEC):

$$P_r(X_i < c) = N^{-1}(c) = p^*.$$

La probabilidad de incumplimiento condicionada a S_t puede escribirse como:

$$\begin{aligned} P_r(X_i < c | S) &= P_r(S\sqrt{\rho} + Z_i\sqrt{1-\rho} < c | S) \\ &= P_r\left(Z_i < \frac{c - S\sqrt{\rho}}{\sqrt{1-\rho}} | S\right). \end{aligned}$$

Entonces, la probabilidad de incumplimiento condicionada a S es igual a:

$$\begin{aligned} 1.4 \quad P_r(X_i < c | S) &= N\left(\frac{c - S\sqrt{\rho}}{\sqrt{1-\rho}}\right) \\ &= N\left(\frac{N^{-1}(p^*) - S\sqrt{\rho}}{\sqrt{1-\rho}}\right). \end{aligned}$$

La función de distribución de la proporción de pérdidas por incumplimiento es determinada por dos parámetros: la probabilidad de incumplimiento, p , y la correlación de activos ρ (rho). La gráfica 2 muestra la distribución de pérdidas de la cartera con probabilidad de incumplimiento ($p=0.02$ o 2%) y correlación de activos ($\rho = 0.1$ o 10%). Esta es la probabilidad incondicional de incumplimiento.

En el modelo de Vasicek, dos procesos rigen el nivel cíclico de la tasa de pérdidas de una cartera: el factor común estocástico S y las correlaciones de activos ρ . Lo que sigue es una interpretación económica de estos dos procesos comenzando con el factor común S .

Dado un escenario macroeconómico, un S puede computarse y luego utilizarse en el modelo de Vasicek para calcular la tasa de pérdidas condicionada a ese escenario específico. El componente común S pudiera verse como una representación de las condiciones macrofinancieras agregadas que pueden extraerse de los datos económicos observables. El riesgo de crédito agregado depende del factor común estocástico S porque, cuando vivimos buenos tiempos económicos, la tasa de pérdidas esperada tiende a ser inferior al promedio de largo plazo; cuando vivimos épocas malas, la tasa de pérdidas esperada tiende a ser superior

al promedio de largo plazo. En este modelo, S es inobservable. A pesar de su naturaleza latente, muchas de las variables macroeconómicas y financieras que se recopilan regularmente contienen información relevante sobre el estado de las condiciones económicas y financieras. Si de cada una de estas variables observables podemos extraer la parte común de información, que representa el estado de las condiciones agregadas, entonces podemos usar esa medida como el factor S en el modelo de Vasicek y computar la tasa de pérdida condicional. Es mediante la S estimada que un escenario macroeconómico específico se toma en cuenta para calcular la tasa de incumplimiento. Por lo tanto, S pudiera considerarse como la parte de incumplimiento macro a micro del modelo, mediante la cual las condiciones macroeconómicas y crediticias se traducen en tasas de incumplimiento aplicables. El algoritmo del filtro de Kalman puede utilizarse para computar S . La principal ventaja de esta técnica es que permite que las variables de estado sean magnitudes no observadas.

La gráfica 3 muestra la pérdida esperada condicionada al factor común S , donde este ha sido calculado utilizando el filtro de Kalman. En la figura puede verse que S muestra una fuerte persistencia. Por lo tanto, una mala ejecución de S por lo general va seguida de una mala ejecución de S y viceversa. S es una variable normal estándar con media 0 y desviación estándar 1. En épocas normales no deberían observarse valores negativos elevados de S . Pero con tensión financiera, S podría entrar más de lleno en territorio negativo.

En el apéndice al final de este documento se demuestra cómo puede calcularse S de manera empírica utilizando el algoritmo del filtro de Kalman. En Harvey (1989) y en Durbin y Koopman (2012) se encuentra una explicación detallada del filtro de Kalman.

Las correlaciones de activos ρ son una manera de medir la probabilidad de incumplimiento simultáneo de dos obligados que pertenecen a la misma cartera y, por lo tanto, son factores determinantes del riesgo de crédito. Es necesario aclarar el papel

que desempeñan las correlaciones en el modelo de Vasicek. Una cartera con correlaciones elevadas produce mayores oscilaciones de incumplimiento durante el ciclo S , en comparación con una cartera con correlaciones más bajas. Las correlaciones no afectan en qué momento ocurre el incumplimiento; una correlación más elevada no implica que el incumplimiento se presentará antes o después que en otras carteras. Por lo tanto, durante las épocas de auge, una cartera con correlaciones elevadas producirá menos incumplimientos que una cartera con correlaciones bajas. Aunque lo opuesto es cierto en las malas épocas, las correlaciones elevadas están creando más incumplimientos.

Algunos valores de referencia de ρ pueden obtenerse de las autoridades regulatorias. Las fórmulas de riesgo ponderado de Basilea II basadas en calificaciones internas, las cuales emplean el modelo de Vasicek, prescriben, para las exposiciones de compañías, correlaciones de un 12% a un 24%, en las que el número real se computa como una probabilidad de incumplimiento promedio ponderado.

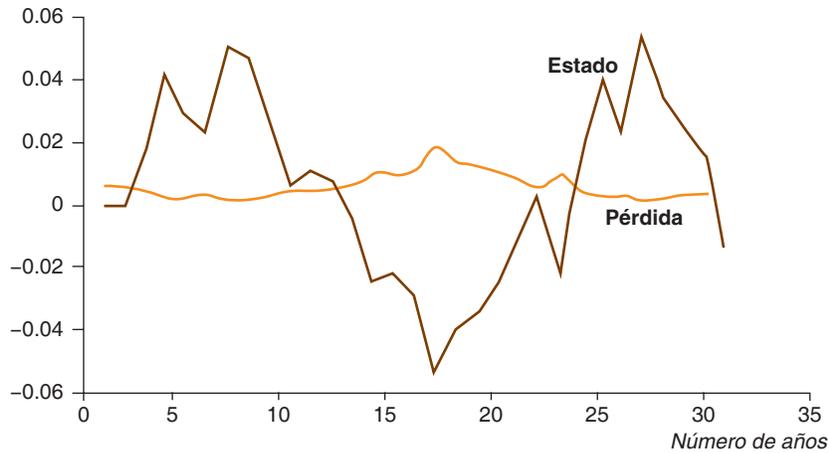
Conforme al modelo de Vasicek, dos prestatarios están correlacionados porque ambos están vinculados al factor común S . Claramente, esta es una simplificación de la estructura verídica de correlación. Sin embargo, permite al modelo de Vasicek proporcionar un cálculo directo del riesgo de incumplimiento de una cartera.

1.3 EL MODELO DE VASICEK Y LA INVARIANCIA DE LA CARTERA

Los bancos emplean sus modelos internos para calcular los parámetros de entrada del modelo de Vasicek. Con el fin de que el método basado en calificaciones internas sea aplicable en todas las jurisdicciones, el Comité de Basilea dispuso que se utilizaran *ponderaciones de riesgo* para determinar los requerimientos de capital. Estas ponderaciones de riesgo serían válidas si un cierto activo tiene la misma ponderación de riesgo sin importar

Gráfica 3

**PÉRDIDA ESPERADA CONDICIONAL
SOBRE EL FACTOR DE ESTADO COMÚN**



a qué cartera se incorpora. Esta propiedad se conoce como invariancia de la cartera.

En teoría, las ponderaciones de riesgo de una cartera sin variación pueden utilizarse para limitar la probabilidad de que las pérdidas excedan el capital total si se cumplen dos supuestos. Primero, las carteras deben ser *desmenuzadas asintóticamente*, lo que significa que el tamaño de cada préstamo debe ser despreciable. Segundo, se debe establecer un supuesto *factor único asintótico de riesgo* (FUAR). El supuesto del FUAR significa que, si bien cada préstamo está expuesto al riesgo idiosincrásico, sólo hay una fuente de choques comunes. Ver Gordy (2003). Podemos considerar el FUAR como la representación de las condiciones microfinancieras agregadas, S , descritas antes. Cada crédito pudiera tener una correlación diferente con el FUAR, pero las correlaciones entre ellos se rigen por su vínculo con ese factor único S . La violación de cualquiera de estos supuestos produciría una valoración incorrecta del riesgo de crédito de la cartera.

El supuesto de *desmenuzado asintóticamente* trata la cartera como si estuviera conformada por un número infinito de exposiciones despreciables.

Sin embargo, en la práctica, las carteras reales suelen ser nodulares. La nodularidad de hecho disminuye la diversificación de la cartera y, por lo tanto, aumenta la varianza en la distribución de pérdidas. De manera que, para impedir que la misma cartera contenga pérdidas, se requiere más capital para una cartera nodular que para una desmenuzada asintóticamente. El FUAR implica que los efectos de diversificación particulares de una institución no se toman en cuenta al calcular los APR. Más bien, los APR se calibran conforme a un banco *ideal* que está bien diversificado y es internacional.

Los valores específicos en las fórmulas basadas en calificaciones internas son dependientes de las categorías de activos, dado que los distintos prestatarios y las distintas categorías de activos muestran distintos grados de dependencia con respecto de las condiciones macroeconómicas agregadas. Para el método basado en calificaciones internas, los bancos deben clasificar la exposición en libros en cinco categorías generales de activos: empresas, deuda soberana, bancos, cartera minorista y capital accionario. Como parte del método basado en calificaciones internas, la cantidad de capital regulador depende únicamente de la PNE, es decir,

deben calcularse niveles de capital mínimos que sean suficientes para cubrir la pérdida no esperada en la cartera.

Para la exposición a empresas, deuda soberana y bancos, la pérdida no esperada se define como: $PNE = (\text{pérdida total} - PE) \times \text{ajuste del vencimiento}$

$$K = PNE = \left[PDI.N \left[\frac{\sqrt{\rho}N^{-1}(0.999) + N^{-1}(PI)}{\sqrt{1-\rho}} \right] - PI.PDI \right] \times \frac{1 + (M - 2.5) \times b}{1 - 1.5 \times b},$$

donde:

- N y N^{-1} representan la función de distribución normal e inversa, respectivamente;
- correlación de activos

$$\rho = 0.12 \times \frac{1 - e^{-50 \times PI}}{1 - e^{-50}} + 0.24 \times \left[1 - \frac{1 - e^{-50 \times PI}}{1 - e^{-50}} \right];$$

- ρ tiene un rango permitido del 12%-24%;
- M es el plazo efectivo promedio de la cartera; ajuste del plazo

$$b = (0.11852 - 0.05478 \times \ln(PI))^2.$$

Se pueden hacer varias inferencias a partir de esta caracterización de la pérdida no esperada. Primero, la correlación de activos se modela enteramente como una función sólo de la PI. Segundo, la fórmula fija un requerimiento de capital mínimo tal que las pérdidas no esperadas no rebasen el capital del banco hasta un coeficiente de confianza del 99.9%. Tercero, el plazo promedio de la cartera se supone que es de 2.5 años. Las exposiciones con plazos mayores necesitarán mantener más capital.

Para la exposición a carteras minoristas, los bancos deben proporcionar su propio cálculo de PI, PDI y EAI. Asimismo, para las categorías de activos minoristas, no es aplicable ningún ajuste de plazo. Así, el modelo de Vacisek se ve simplificado.

$$K = PNE = \left(PDI.N \left(\frac{\sqrt{\rho}N^{-1}(0.999) + N^{-1}(PI)}{\sqrt{1-\rho}} \right) - PI.PDI \right)$$

Para las exposiciones a hipotecas residenciales, $\rho = 0.15$. Para cuantificar la exposición a los créditos revolventes de las carteras minoristas (tarjetas de crédito), $\rho = 0.04$.

2. MODELOS ESTRUCTURALES DEL RIESGO DE CRÉDITO

El banco utiliza un modelo del riesgo de crédito para calcular la FDP de una cartera de créditos. En este sentido, los modelos de riesgo de crédito pueden dividirse en dos categorías principales: los estructurales y los de forma reducida. Los modelos estructurales se utilizan para calcular la probabilidad de incumplimiento de una empresa con base en el valor de sus activos y sus pasivos. La idea básica es que una empresa (con responsabilidad limitada) deja de pagar si el valor de sus activos es inferior a su deuda.

Los modelos de forma reducida por lo general suponen una causa de incumplimiento exógena. Modelan el incumplimiento como un evento aleatorio sin concentrarse en el balance de la empresa. Este evento aleatorio de incumplimiento se describe como una distribución de Poisson. Dado

que los modelos de Poisson consideran la tasa de llegada, o la intensidad, de un evento específico, esta manera de modelar el riesgo de crédito también se conoce como modelo de intensidad de incumplimiento. En esta sección se analiza el enfoque estructural del riesgo de crédito. En la sección 3 se consideran los modelos de forma reducida.

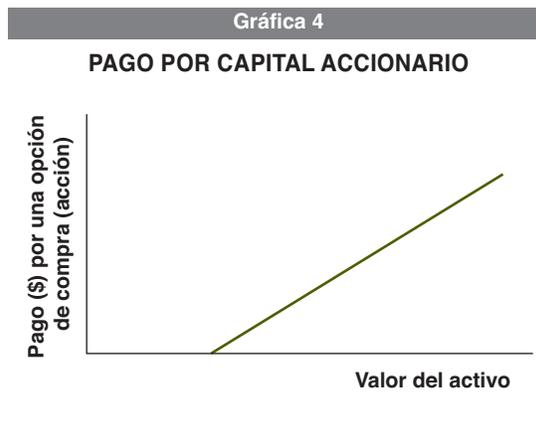
Los modelos estructurales se iniciaron con Merton (1974) y utilizan el marco de fijación de precios de opciones de Black-Scholes para caracterizar el comportamiento de incumplimiento. Se emplean para calcular la probabilidad de incumplimiento de una empresa con base en el valor de sus activos y sus pasivos. El principal problema de estos modelos es que no se observa el valor de mercado de los activos de una compañía. El informe anual de un banco sólo proporciona una versión contable de sus activos. Pero para un banco que cotiza en bolsa, el valor de mercado de su capital accionario es observable, como lo es su deuda. El análisis que sigue se conoce como análisis de derechos contingentes (CCA, por sus siglas en inglés) y utiliza precios del capital accionario e información contable para medir el riesgo de crédito de las instituciones que cotizan en bolsa.

2.1 EL CAPITAL ACCIONARIO Y LA DEUDA COMO DERECHOS CONTINGENTES

Los modelos estructurales del riesgo de crédito consideran los pasivos de una empresa (su capital accionario y su deuda) como *derechos contingentes* contra los activos subyacentes de la compañía. Como cabe la posibilidad de impago de la deuda, la deuda es riesgosa. La deuda riesgosa es el valor sin riesgo de incumplimiento de la deuda menos la pérdida esperada. El valor de la deuda riesgosa, entonces, se deriva del valor de los activos inciertos. Como la deuda riesgosa es un derecho sobre los activos inciertos, tales derechos se conocen como contingentes.

El capital accionario es un derecho subsidiario sobre los activos después de que la deuda ha sido pagada. Esto significa que el tenedor del capital accionario tiene una opción de compra sobre el

valor de los activos de la compañía en el tiempo T , V_T , cuyo pago es cero o el valor de los activos menos los pasivos, D , lo que sea mayor. El precio de ejecución es el valor nominal de la deuda vigente, D . Esto se muestra en la gráfica 4.



El valor de la opción de compra al plazo E_T depende del valor final del subyacente, V_T :

$$E_T = \max(V_T - D, 0).$$

El valor de la deuda riesgosa es el valor sin riesgo de incumplimiento de la deuda menos la garantía de la deuda.

Deuda riesgosa \equiv deuda sin riesgo – garantía contra incumplimiento.

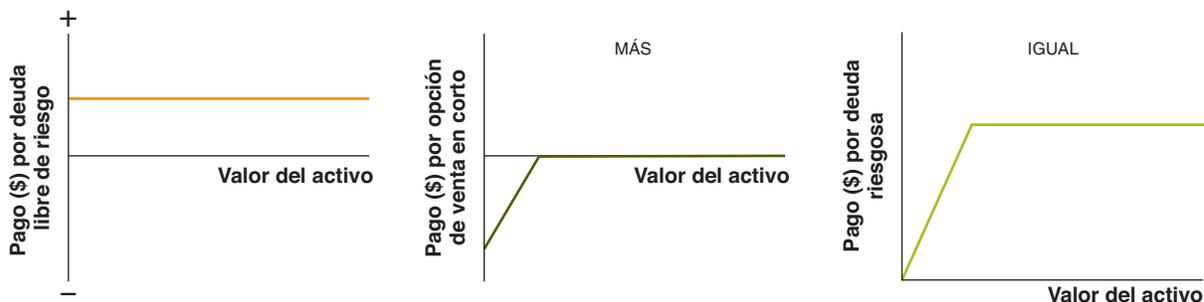
Si la deuda está garantizada con un activo específico, entonces la garantía contra el incumplimiento puede modelarse como una opción de venta sobre los activos cuyo precio de ejecución es igual al valor nominal de la deuda. El tenedor de la deuda está ofreciendo una garantía implícita, pues está obligado a absorber las pérdidas si ocurriera un incumplimiento.

Deuda riesgosa = deuda sin riesgo – opción de venta implícita

Garantía financiera = opción de venta implícita

Gráfica 5

DIAGRAMAS DE PAGO POR DEUDA RIESGOSA: DEUDA LIBRE DE INCUMPLIMIENTO MÁS OPCIÓN DE VENTA EN CORTO



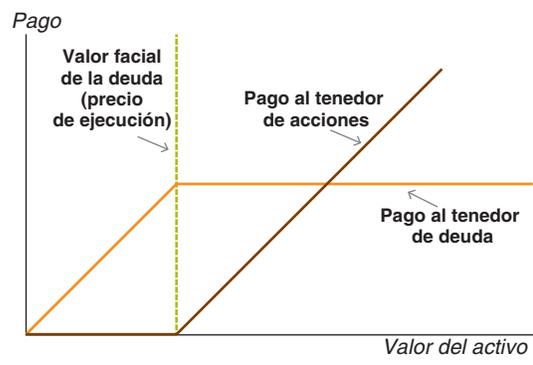
No hay ningún diagrama de pagos simple que corresponda a la perfección con el pago de la deuda. Es necesario replicar la liquidación de la deuda con una combinación de pagos por opciones y de otros valores también. Para empezar, consideremos la compra de un bono cupón cero del Tesoro del Reino Unido, que se considera como deuda sin riesgo. El pago por la deuda sin riesgo en comparación con el valor del activo es una línea horizontal plana, como se muestra en el primer diagrama de la gráfica 5. Sin importar cuánto cambie el valor del activo del banco, el tenedor del bono sólo recibe el valor nominal de este a su vencimiento. Al combinar el pago de un bono del Tesoro con el pago de una opción de venta en corto, obtenemos el pago por la deuda riesgosa como se muestra en el tercer diagrama de la gráfica 5. Invertir en deuda riesgosa es lo mismo que comprar un bono del Tesoro y suscribir una opción de venta sobre los activos de la compañía. Los tenedores de deuda efectivamente están dando a los tenedores de acciones el derecho de vender los activos de la empresa.

Las crisis financieras han elevado la incertidumbre de los tenedores de capital accionario y de deuda de un banco respecto a cómo cambiará el valor de los activos de la institución. Los acreedores de un banco tienen parte en el riesgo a la baja, pero

reciben un pago máximo equivalente al valor nominal de la deuda. Los tenedores de capital accionario se benefician de las alzas cuando no ocurre un incumplimiento, pero su responsabilidad es limitada cuando hay caídas. Por lo tanto, la incertidumbre en el valor del activo del banco tiene un efecto asimétrico en el valor de mercado de la deuda y de las acciones. La asimetría en el pago a los tenedores de deuda y a los de acciones se muestra en la gráfica 6.

Gráfica 6

PAGO A LOS TENEDORES DE DEUDA Y DE ACCIONES DE LA FIRMA



2.2 INCERTIDUMBRE EN EL VALOR DEL ACTIVO

La idea fundamental detrás del CCA es que el riesgo de incumplimiento de una institución depende de la incertidumbre en el valor de sus activos respecto de los pagos prometidos por sus obligaciones de deuda. Los activos de un banco son inciertos y cambian debido a factores como los flujos de utilidades y las exposiciones al riesgo. El riesgo de incumplimiento en un horizonte dado depende de los cambios inciertos en el valor futuro de los activos en relación con los pagos prometidos sobre la deuda; a estos pagos se les suele denominar *barreras al incumplimiento*. Habiendo identificado la naturaleza de los pagos para los tenedores tanto de deuda como de acciones, el siguiente paso sería analizar cómo evolucionaría el valor de los activos de un banco en relación con una barrera al incumplimiento. Los activos estocásticos que evolucionan en relación con una barrera de pánico pueden utilizarse para determinar el valor de los pasivos con opciones implícitas. La probabilidad de que estos activos se ubiquen por debajo de la barrera de pánico es la probabilidad de incumplimiento, que es necesario calcular con el fin de cuantificar el riesgo de crédito.

Un proceso estocástico es un proceso aleatorio indexado por el tiempo. Digamos que $V(t)$ es el valor de los activos en el tiempo t . Los cambios entre cualesquiera dos puntos en el tiempo pueden ser explicados por un componente cierto (el término de deriva) y un componente de incertidumbre (el término aleatorio o estocástico). La deriva representa la tasa de crecimiento esperada (promedio) del valor del activo. El término estocástico es un paseo aleatorio donde la varianza es proporcional al tiempo y la desviación estándar es proporcional a la raíz cuadrada del tiempo. Representa la incertidumbre respecto a la evolución del valor del activo.

La dinámica de los activos que son inciertos sigue este proceso de *difusión*, con deriva y volatilidad, proporcionado por

$$\frac{dV}{V} = \mu dt + \sigma_v dZ .$$

Tal proceso en un tiempo continuo es un movimiento browniano geométrico (MBG) y dz es un proceso de Wiener, que normalmente se distribuye con una media de cero y con varianza unitaria. Una descripción más a fondo se encuentra en Baz y Chacko (2004).

Para un MBG, el valor del activo en el tiempo t puede calcularse a partir del valor del activo en el tiempo 0 utilizando la siguiente relación:

$$V_t = V_0 \exp \left[\left(\mu - \frac{\sigma_v^2}{2} \right) t + \sigma_v \varepsilon \sqrt{t} \right] .$$

Aquí, ε es la ejecución de una variable aleatoria normal con media cero y varianza unitaria. El término de deriva se ajusta conforme al término $\sigma_v^2/2$, que debe incluirse si hay incertidumbre en la evolución de los activos. En un mundo de certidumbre, $\sigma_v = 0$ y, en este caso, $V_t = V_0 \exp(\mu_v t)$.

El activo en el tiempo t pudiera ser mayor o menor a una barrera D_t , la cual representa la cantidad de pagos prometidos sobre la deuda. Dado que los activos pueden caer por debajo de la barrera, podemos calcular la probabilidad de que $V_t < D_t$.

Utilizando la ecuación de arriba para V_t ,

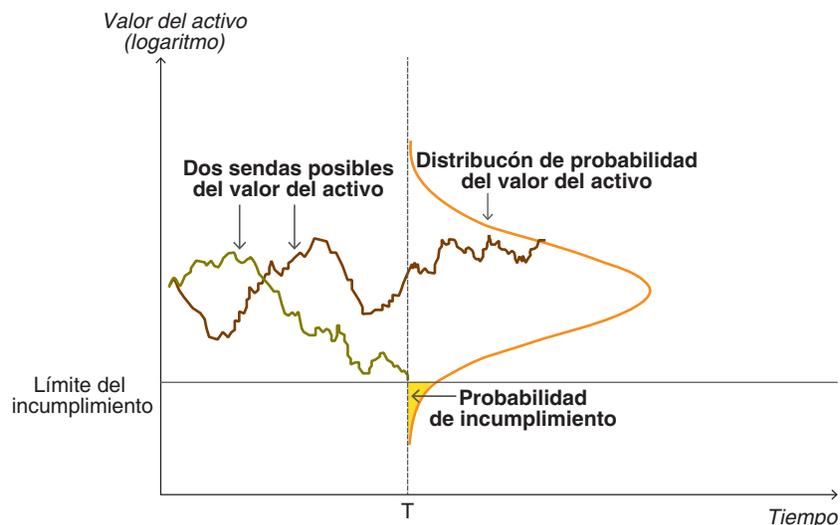
$$\text{Prob}(V_t \leq D_t) = \text{Prob} \left(V_0 \exp \left[\left(\mu - \frac{\sigma_v^2}{2} \right) t + \sigma_v \varepsilon \sqrt{t} \right] \leq D_t \right) .$$

La reorganización de la probabilidad de que los activos sean inferiores o iguales que la barrera equivale a la probabilidad de que el componente aleatorio del rendimiento del activo, ε , sea inferior a $-d_2$.

El término d_2 se denomina *distancia hasta el incumplimiento* y es el número de desviaciones estándar que el valor actual del activo se aleja de la barrera al incumplimiento, D_t . Es un concepto propuesto por KMV Corporation y explicado en Kealhofer (2003). Dado que $\varepsilon \sim N(0, 1)$ está distribuido normalmente,

$$\text{Prob}(V_t \leq D_t) = \text{Prob}(\varepsilon \leq -d_2) \sim N(-d_2),$$

PROBABILIDAD DE INCUMPLIMIENTO



es decir, la probabilidad de incumplimiento de la deuda es la distribución normal estándar acumulativa de la distancia hasta el incumplimiento negativa, d_2 . La gráfica 7 ilustra dos rutas posibles para el valor del activo de la empresa. Un mayor rendimiento del activo aumenta más rápidamente el valor del activo de la empresa, lo que reduce la probabilidad de incumplimiento si las otras cosas permanecen igual. Pero también hay incertidumbre acerca del crecimiento del valor del activo. La incertidumbre respecto al crecimiento del valor del activo significa que el rango de valores posibles para los activos de la compañía se amplía con el tiempo. La distribución de probabilidad del valor del activo en el tiempo T se desarrolla sobre el supuesto de que los activos financieros siguen una distribución logarítmica normal. Por lo tanto, el logaritmo del valor del activo sigue una distribución normal en un tiempo T . Si el valor del activo de la compañía cae por debajo de la línea horizontal (el límite del incumplimiento), se da un incumplimiento. La probabilidad de incumplimiento es la zona debajo de la barrera al incumplimiento en la gráfica 7. Para llegar a la probabilidad de incumplimiento,

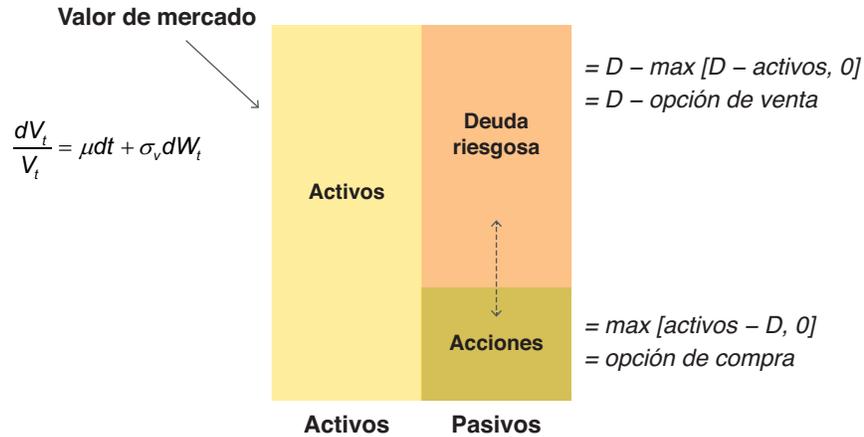
necesitamos calcular la media y la varianza de la distribución de probabilidad.

2.3 CÁLCULO DE LA PROBABILIDAD DE INCUMPLIMIENTO

La gráfica 8 muestra una identidad del balance que siempre se mantiene: los activos equivalen al valor de la deuda en riesgo más el capital. El valor del activo es estocástico y pudiera caer por debajo de los pasivos por pagar que constituyen el punto de quiebre (la *barrera al incumplimiento*), D . D se define como el valor presente de los pagos prometidos sobre la deuda, descontados a la tasa sin riesgo.

Los pasivos pendientes de la compañía constituyen el punto de quiebre cuya densidad estándar normal define la *distancia hasta el incumplimiento* en relación con el valor de la empresa. El valor del capital accionario es el valor de una opción de compra implícita sobre los activos cuyo precio de ejecución es igual a la barrera al incumplimiento. El valor del capital accionario puede computarse como el valor de una opción de compra como se muestra en la ecuación 2.1

EVOLUCIÓN DEL BALANCE



$$2.1 \quad E(t) = V(t)N(d_1) - De^{-rt}N(d_2),$$

donde los factores d_1 y d_2 los proporcionan:

$$d_1 = \frac{\ln\left(\frac{V_0}{D}\right) + \left(r + \frac{\sigma_v^2}{2}\right)T}{\sigma_v \sqrt{T}}$$

$$d_2 = d_1 - \sigma_A \sqrt{T} = \frac{\ln\left(\frac{V_0}{D}\right) + \left(r - \frac{\sigma_v^2}{2}\right)T}{\sigma_v \sqrt{T}},$$

donde r es la tasa sin riesgo, σ es la volatilidad del valor del activo y $N(d)$ es la probabilidad de que la función de densidad normal estándar sea inferior a d .

El valor presente de las pérdidas esperadas implícitas en el mercado asociadas a las obligaciones pendientes puede valorarse como una opción de venta implícita, que se calcula con el umbral de incumplimiento D como precio de ejecución sobre el valor del activo V . La opción de venta implícita es proporcionada por:

$$2.2 \quad P(t) = De^{-rt}(N(-d_2)) - V_0N(-d_1).$$

El valor de la deuda riesgosa B es, por lo tanto, el valor sin riesgo de incumplimiento menos la pérdida esperada, como lo proporciona la opción de venta implícita:

$$2.3 \quad B = De^{-rt} - P(t).$$

El valor de mercado de los activos de los bancos no puede observarse directamente, pero puede inferirse utilizando los precios de los activos financieros. A partir de los precios observados y de las volatilidades de los valores cotizados en bolsa, es posible estimar los valores implícitos y las volatilidades de los activos subyacentes en los bancos. Usando técnicas numéricas, también pueden estimarse directamente los activos y su volatilidad para calibrar el modelo de Merton. En la ecuación 2.1, ni V ni σ_v son observables directamente. Sin embargo, si la compañía cotiza en bolsa, entonces observamos E . Esto significa que la ecuación 2.1 proporciona una condición que

debe ser satisfecha por V y σ_v . σ_E también puede calcularse a partir de los datos históricos. Con el fin de calibrar el modelo de Merton, necesitamos encontrar una segunda ecuación en estas dos variables desconocidas, V y σ_v . Para hacerlo, invocamos el lema de Ito como sigue:

$$dE_t = \frac{\delta E}{\delta V} dV_t + \frac{1}{2} \sigma_v^2 \frac{\delta E^2}{\delta V^2} (dV_t)^2.$$

Utilizando el lema de Ito también podemos enunciar que:

$$\sigma_E E = \frac{\delta E}{\delta V} \sigma_v V,$$

donde $\frac{\delta E}{\delta V}$ es la delta del capital accionario. Puede probarse que esta delta es:

$$\sigma_E E = N(d_1) \sigma_v V.$$

Significativamente, a partir de lo anterior podemos relacionar la volatilidad desconocida de los valores de los activos a la volatilidad observable del capital accionario:

2.4
$$\sigma_v = \left(\frac{V}{E} \frac{\delta E}{\delta V} \right)^{-1} \sigma_E.$$

Esto proporciona otra ecuación que debe ser satisfecha por V_0 y δ_v .

Por lo tanto, calibrar el modelo de Merton requiere conocer el valor del capital accionario E , la volatilidad del capital accionario, σ_E y la barrera de pánico como variables de entrada para las ecuaciones $E(t) = V(t)N(d_1) - De^{-rt}N(d_2)$ y $E\sigma_E = V\sigma_v N(d_1)$, con el fin de calcular el valor implícito del activo V y la volatilidad implícita del activo σ_v .

2.4 APLICACIÓN DEL MODELO DE MERTON

Ilustremos el procedimiento del modelo de Merton, descrito antes, con un ejemplo. Para hacerlo,

inicializamos los parámetros del modelo de Merton con los siguientes valores:

- $V = 100$; valor del activo.
- $D = 90$; valor de la deuda sin riesgo de incumplimiento o *barrera al incumplimiento*.
- $r = 0.05$ (5%); tasa de interés sin riesgo.
- $\sigma_v = 0.10$ (10%); incertidumbre del rendimiento del valor del activo.
- $T = 1$; tiempo para el vencimiento.

La solución del modelo arroja el valor del capital accionario, E , y de la deuda riesgosa, B . Utilizando un procedimiento iterativo, el resultado del modelo de Merton da 14.63 como valor de E y 85.37 como valor de B . La probabilidad neutral al riesgo de que la compañía deje de pagar su deuda es $N(-d_2)$. La probabilidad neutral al riesgo de incumplimiento describe la posibilidad de que una empresa caiga en situación de incumplimiento, si esta opera en una economía neutral al riesgo, es decir, en la que los inversionistas no obtienen una prima por tomar un riesgo de incumplimiento.

La probabilidad neutral al riesgo de incumplimiento $N(-d_2)$ es 6.63% para un año.

Como estamos modelando el riesgo de crédito, deseamos estimar el diferencial crediticio s . Esta es la prima de riesgo requerida para compensar la pérdida esperada (PE). El diferencial crediticio S es el diferencial entre la tasa hasta el vencimiento, y , sobre la tasa de interés sin riesgo r .

El rendimiento al vencimiento para la deuda riesgosa B , cuya notación es y , se obtiene de la siguiente manera:

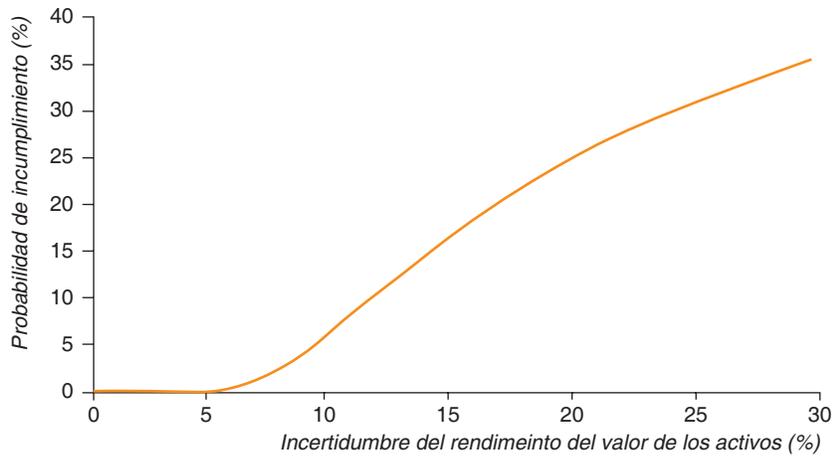
$$B = De^{-yT}$$

$$y = \frac{\ln\left(\frac{D}{B}\right)}{T}$$

$$s = y - r = 0.0028.$$

Gráfica 9

**VARIACIÓN DE LA PROBABILIDAD DE INCUMPLIMIENTO
CON INCERTIDUMBRE DE ACTIVOS**



Por lo tanto, el diferencial crediticio para la deuda riesgosa es igual a 28 puntos básicos (0.28%). Para más detalles sobre este método con el fin de estimar el riesgo de crédito soberano, ver Gray *et al.* (2008).

Utilizando los mismos parámetros del modelo, podemos hacer un análisis de sensibilidad variando del 0% al 30% la incertidumbre del valor del activo. La gráfica 9 muestra cómo va aumentando la probabilidad de incumplimiento conforme aumenta la volatilidad en el valor del activo. Esto implica que, si el valor de los activos del banco fluctúa con el tiempo, aumenta la probabilidad de que disminuya hasta ser inferior al valor de la deuda al vencimiento.

El método CCA es útil porque proporciona probabilidades prospectivas de incumplimiento que toman en cuenta tanto el grado de apalancamiento como la perspectiva de los participantes del mercado respecto a la calidad crediticia. En el marco de las pruebas de resistencia, proporciona un referente estandarizado del riesgo de crédito (probabilidades de incumplimiento) que facilita las comparaciones entre sectores y entre densidades. Sin embargo, el CCA sólo puede aplicarse a entidades cuyas acciones cotizan en bolsa o con diferenciales de CDS muy líquidos, y no puede consignar la liquidez ni el riesgo de renovación de financiamiento.

En el modelo estructural de riesgo de crédito, el valor del activo subyacente sigue un MBG estándar sin saltos y con una deriva y volatilidad constantes:

$$\frac{dV_t}{V_t} = \mu dt + \sigma dW_t.$$

Como ya dijimos, este es el proceso de difusión del valor del activo en el modelo de Merton (1974). Una variable estocástica V_t puede seguir un MBG como el ya descrito y, además, registrar, en la cima de este, saltos aleatoriamente cuando desciende a un valor más bajo. Desde estos valores posteriores al salto, puede retomar el proceso de difusión original hasta el siguiente salto, y así sucesivamente. Podemos ampliar la ecuación agregando un proceso de salto:

$$\frac{dV_t}{V_t} = \mu dt + \sigma dW_t + (J - 1)dY_t.$$

La ocurrencia del salto se modela utilizando un proceso de Poisson Y_t con intensidad λ :

dY_t es un proceso de salto definido por

$$dY_t = \begin{cases} 0 \\ 1 \end{cases} \text{ con probabilidades } \begin{matrix} (1 - \lambda)dt \\ \lambda dt \end{matrix},$$

el tamaño del salto J se obtiene aleatoriamente a partir de una distribución con función de densidad de probabilidad $P(J)$ que es independiente tanto del movimiento browniano como del proceso de Poisson. Intuitivamente, si hay un salto ($dY = 1$), V inmediatamente asume un valor JV . Por ejemplo, una caída súbita del 10% en el precio del activo podría modelarse fijando $J = 0.9$.

En el método estructural, el término dY está ausente; el valor de la empresa se modela como un proceso continuo y el incumplimiento ocurre cuando el valor alcanza alguna barrera. En los modelos de forma reducida, el hincapié se hace en el proceso de salto dY y el incumplimiento ocurrirá en el primer salto de J .

3.1 INTENSIDAD DEL INCUMPLIMIENTO

En la forma reducida o en los modelos de intensidad del incumplimiento, la herramienta fundamental es el proceso de Poisson. Empezaremos por demostrar sus propiedades. Suponemos que hay extracciones constantes de la distribución de Poisson y que cada extracción produce ya sea un

0 o un 1. La mayoría de las extracciones dan por resultado un 0, pero cuando la extracción da por resultado un 1, este representa un incumplimiento. La distribución de Poisson especifica que el tiempo entre la ocurrencia de este suceso en particular y la ocurrencia previa del mismo suceso tiene una distribución exponencial. El recuadro 2 formaliza el proceso de Poisson.

RECUADRO 2

EL PROCESO DE POISSON Y LAS DISTRIBUCIONES

Un proceso de Poisson es un proceso de llegadas en el que N_t es el número de llegadas del tiempo 0 al tiempo t , y,

- Todas las llegadas tienen un tamaño de 1.
- Para toda $t, s > 0$, $N_{t+s} - N_t$ es independiente de la historia hasta t .
- Para toda $t, s > 0$, $N_{t+s} - N_t$ es independiente de t .

La probabilidad de k llegadas en un tiempo t tiene una distribución de Poisson:

$$\text{Prob}(N_t = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}.$$

Y el número esperado de llegadas entre 0 y t es:

$$E[t_1] = \lambda t.$$

Es importante señalar que λ tiene una dimensión temporal. Por ejemplo, si se refiere a

un año, entonces $t = 1$, arriba da las llegadas esperadas en un año; $t = 1/52$ da las llegadas esperadas en una semana.

El tiempo de espera esperado (t_1) hasta la primera llegada es:

$$E[t_1] = \frac{1}{\lambda};$$

y la probabilidad de ninguna llegada entre 0 y t es:

$$\text{Prob}(N_t = 0) = e^{-\lambda t}.$$

El mismo parámetro, λ , determina todas las magnitudes de arriba: nos da los tiempos de espera y el número esperado de llegadas antes de un tiempo dado. λ tiene varios nombres, como *tasa de riesgo*, *tasa de llegada* o *intensidad de llegada*. Las llegadas N_t pueden utilizarse para representar la llegada de incumplimientos en una cartera de bonos, por ejemplo.

Cuando el proceso de Poisson se utiliza para el riesgo de crédito, la tasa de llegada se conoce como intensidad de incumplimiento y normalmente se representa por λ .

- La probabilidad de incumplimiento entre 0 y t es:

$$\text{Prob}(\text{incumplimiento}) = 1 - e^{-\lambda t}.$$

- La probabilidad de ningún incumplimiento entre 0 y t es:

$$\text{Prob}(\text{no incumplimiento}) = e^{-\lambda t}.$$

- El tiempo esperado hasta el incumplimiento (es decir, el primer y único incumplimiento posible) es:

$$E[t_d] = \frac{1}{\lambda}.$$

3.2 INSTRUMENTOS DE CAPITAL CONVERTIBLE CONTINGENTE

Aplicaremos el modelo de intensidad de incumplimiento, descrito arriba, para determinar el precio de los instrumentos de capital convertible contingente (CoCo). Un CoCo es un bono que se convertirá en acciones o sufrirá un castigo en su valor nominal en cuanto el capital del banco emisor disminuya a menos de cierta cantidad definida. Esta cantidad de activación es el punto en el que el banco es considerado como carente de capital regulatorio suficiente. Una gran lección de la crisis financiera es que los instrumentos de capital regulatorio en el futuro deben ser capaces de absorber las pérdidas con el fin de contribuir a que los bancos sigan siendo *entidades en funcionamiento*. La activación de la conversión del bono a acciones o del castigo en el valor nominal ocurre cuando el banco aún es una entidad en funcionamiento. La conversión debe ocurrir antes de que los bancos tengan que castigar el valor en libros de sus activos y mucho antes de que desencadenen medidas resolutorias. Un suceso desencadenante es una barrera que causa otro suceso, en este caso la conversión del CoCo. El riesgo de conversión puede compararse con un riesgo de incumplimiento.

Un CoCo puede convertirse en un número predefinido de acciones. Otra posibilidad es que el valor nominal de la deuda se reduzca en libros. Este análisis se concentra en la conversión de un CoCo en acciones y en el concepto de la tasa de recuperación. Para un análisis de las características de los CoCo, ver Haldane (2011). Un análisis cuantitativo adicional se encuentra en Spiegeleer y Schoutens (2011).

El número de acciones recibidas por bono convertido es el coeficiente de conversión C_r . El precio de conversión (C_p) de un CoCo con un valor nominal F es el precio de compra implícito de las acciones subyacentes en el suceso desencadenante:

3.1

$$C_p = \frac{F}{C_r}.$$

Si el bono se convierte en acciones, la pérdida para el inversionista (L_{CoCo}) depende del coeficiente de conversión (C_r) y del valor S^* de las acciones cuando ocurre la activación. Por lo tanto, si el CoCo se activa y una conversión ocurre:

3.2

$$L_{CoCo} = F - C_r S^* = F \left(1 - \frac{S^*}{C_p} \right) = F(1 - R_{CoCo}).$$

La ecuación 3.2 plantea la introducción de una tasa de recuperación para un bono CoCo. Entonces, con la notación de arriba, la tasa de recuperación R_{CoCo} se define como:

3.3

$$R_{CoCo} = \frac{S^*}{C_p}.$$

La tasa de recuperación R_{CoCo} estará determinada por el precio de conversión, C_p , y por S^* , el precio de las acciones al momento de conversión. Cuando más cercano esté el precio de conversión C_p al precio de mercado S^* de las acciones en la fecha de activación, más alto será el coeficiente de recuperación.

3.3 DETERMINACIÓN DEL PRECIO DE LOS BONOS COCO

Hay diferentes maneras de fijar el precio de los bonos CoCo. En este análisis, vemos el bono CoCo como un instrumento de crédito y fijamos su precio con un método de forma reducida. Al principio de este capítulo se explicó que, con un método de forma reducida, se utiliza un parámetro de intensidad de incumplimiento λ para modelar el incumplimiento. A esto también se conoce como determinación del precio de derivados de crédito. Los instrumentos de crédito por lo general cotizan con base en su diferencial de crédito con respecto a la tasa de interés sin riesgo. El diferencial de crédito CS está vinculado a la tasa de recuperación $(1-R)$ y a la intensidad de incumplimiento mediante lo que se conoce como el triángulo del crédito:

$$3.4 \quad CS = (1-R) \times \lambda.$$

El diferencial del crédito es el producto de la pérdida $(1-R)$ y la probabilidad instantánea de que tal pérdida ocurra (λ). La aplicación de este principio nos permite ver el suceso desencadenante que convierte un CoCo en acciones como un suceso extremo parecido al que ocurre en el mercado de canjes de riesgo crediticio. Entonces, la activación de la conversión del CoCo puede modelarse como un suceso extremo. La intensidad de incumplimiento λ es reemplazada por una intensidad de activación λ_{activ} cuyo valor es más elevado que la intensidad de incumplimiento correspondiente. A partir de la ecuación 3.4 podemos determinar el valor del diferencial de crédito de un CoCo utilizando el triángulo del crédito.

$$3.5 \quad CS_{CoCo} = (1-R_{CoCo}) \times \lambda_{activ}.$$

Este método puede aplicarse a la fijación del precio de los CoCos después de hacer algunos ajustes. Primero, para prevenir un incumplimiento por parte del banco, la conversión del CoCo debe ocurrir antes del tiempo de incumplimiento. Esto

implica que la intensidad de incumplimiento de la conversión λ_{activ} tiene que ser mayor que la intensidad de incumplimiento de la propia entidad, a saber, el banco. Esto es porque el CoCo cumplirá su propósito si se convierte antes de que el banco incurra en un incumplimiento.

La intensidad de la activación λ_{activ} está vinculada a la probabilidad de que se alcance la activación, P_{activ} , de acuerdo con el proceso de Poisson:

$$3.6 \quad P_{activ} = 1 - e^{-\lambda_{activ}T},$$

donde T es el plazo del bono CoCo. La probabilidad de que se alcance la activación sería equivalente a alcanzar una barrera en un marco de opción de barrera. La ecuación 3.6 da la probabilidad de que el CoCo incumpla en el tiempo T . Al resolver la ecuación 3.6 para λ_{activ} , obtenemos lo siguiente:

$$3.7 \quad \lambda_{activ} = \frac{\log(1 - P_{activ})}{T}.$$

Por ende, la ecuación 3.5 para el diferencial de crédito del CoCo tiene la siguiente solución computable:

$$3.8 \quad CS_{CoCo} = \lambda_{activ} (1 - R_{CoCo}) \\ = -\frac{\log(1 - P_{activ})}{T} \left(1 - \frac{S^*}{C_p}\right).$$

Es difícil determinar el precio de los CoCo debido a su sensibilidad a la probabilidad de activación. Spiegeleer y Schoutens (2011) muestran que en un modelo de Black-Scholes, la probabilidad p^* de alcanzar S^* es proporcionada por:

$$3.9 \quad p^* = N \left[\frac{\log\left(\frac{S^*}{S}\right) - \mu T}{\sigma \sqrt{T}} \right] + \\ + \left(\frac{S^*}{S}\right)^{\frac{2\mu}{\sigma^2}} N \left[\frac{\log\left(\frac{S^*}{S}\right) + \mu T}{\sigma \sqrt{T}} \right].$$

- $\mu = r - q - \frac{\sigma^2}{2}$,
- q = rendimiento del dividendo continuo,
- r = tasa de interés continua,
- σ = volatilidad,
- T = plazo del bono convertible contingente, y
- S = precio actual de la acción.

Esto permite una solución cerrada y también favorece una mejor apreciación de las cualidades de absorción de pérdidas del bono CoCo.

4. RIESGO DE CRÉDITO DE LA CONTRAPARTE

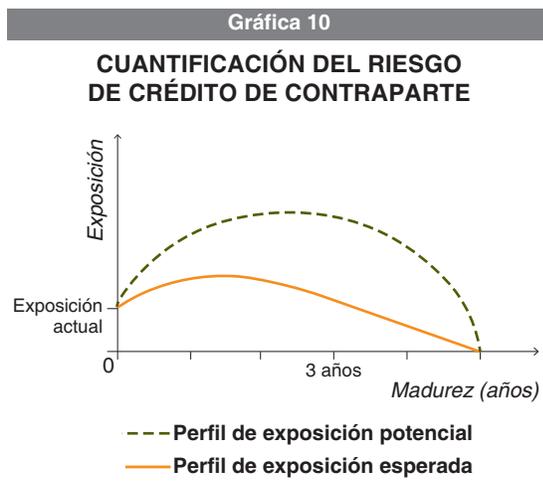
El riesgo de crédito de la contraparte (RCC) es que la otra parte en una transacción incumpla antes de la liquidación de los flujos finales de efectivo. Este riesgo se presenta en los derivados extrabursátiles, valores para el financiamiento de transacciones y las operaciones con liquidación en el largo plazo. El RCC tiene las siguientes características generales:

- es bilateral (es decir, cada parte puede tener una exposición a la contraparte);
- lo que se conoce hoy es sólo la exposición actual;
- es aleatorio y depende de la posible exposición futura.

Estas características diferencian el RCC del riesgo de crédito. A diferencia del riesgo de mercado, el RCC ocurre cuando el valor de mercado de las transacciones está a nuestro favor (es decir, una valoración a precio de mercado positiva) y la contraparte incumple. Por lo general, la cuantificación del RCC implica:

- simular los factores de riesgo en múltiples puntos futuros en el tiempo durante la vida de la cartera;
- recalcular el precio de las posiciones en cada punto en el tiempo;
- agregar posiciones de manera coherente con la trayectoria, tomando en cuenta la posición neta y la garantía.

La gráfica 10 muestra que la cuantificación de la exposición implica lograr un equilibrio entre dos efectos. Primero, la incertidumbre de las variables del mercado y, por lo tanto, que el riesgo aumenta conforme avanzamos en el tiempo. Segundo, los contratos de derivados implican flujos de efectivo que se van pagando con el tiempo y reducen el perfil de riesgo conforme los valores subyacentes se amortizan con el tiempo. Por ejemplo, en un contrato de canje de tasa de interés a cinco años, la exposición máxima para el operador es improbable que ocurra en el primer año porque hay menos incertidumbre respecto a las tasas de interés en



ese periodo. También es improbable que ocurra en el último año, dado que para entonces ya se habrán realizado la mayoría de los pagos del canje. Es más probable que la exposición máxima se dé a la mitad del contrato. Un análisis de los diferentes métodos para cuantificar el RCC se encuentra en Gregory (2011).

4.1 AJUSTES AL VALOR DEL CRÉDITO

El ajuste de valuación crediticia (AVC) suele mencionarse en el contexto del riesgo de mercado y el RCC. Es una corrección que los bancos hacen al

valor de las transacciones para reflejar las posibles pérdidas futuras en que podrían incurrir debido al incumplimiento de la contraparte. El AVC es la diferencia entre el precio de un derivado con riesgo crediticio y el precio de un derivado sin riesgo de incumplimiento que da cuenta de la pérdida esperada por un incumplimiento de la contraparte. Los bancos reconocen el riesgo de contraparte en las operaciones con derivados y hacen ajustes al valor del crédito. El acuerdo de Basilea II informa que dos terceras partes de las pérdidas por riesgo de crédito durante la crisis financiera mundial se debieron a la volatilidad de este tipo de ajustes, más que en los incumplimientos reales. El AVC también forma parte integral del acuerdo de Basilea III; sin embargo, es básicamente una valoración y un precio apegado al mercado y no sustituye la administración tradicional del riesgo de crédito de contraparte.

Cuando se presenta este riesgo, el valor de un derivado puede anotarse como:

$$4.1 \quad X = X_f - AVC,$$

donde X es el valor del derecho. X_f es el valor sin riesgo de crédito del activo y AVC es el ajuste al valor del crédito que varía conforme a la solvencia de la contraparte. AVC es por definición ≥ 0 . De la ecuación 4.1 se tiene que un derivado con riesgo crediticio tiene un precio más bajo que un derivado sin riesgo. Esto se debe a que el comprador del derivado con riesgo crediticio (comúnmente un operador) reduce el precio del derivado para tomar en cuenta el riesgo de crédito de la contraparte (el vendedor del derivado). En particular, si la contraparte incumple, el comprador del derivado no recibirá pago alguno por el derivado. El AVC es un ajuste porque el comprador del derivado corrige (reduce) el precio del derivado debido al riesgo de crédito.

El AVC es proporcionado por:

$$4.2 \quad AVC = PDI \sum_{t=1}^T FD_t EE_t PI_t,$$

donde:

- PDI es la pérdida dado el incumplimiento;
- FD_t es el factor de descuento para el plazo t ;
- EE_t es la exposición esperada en el tiempo t ; y
- PI_t es la probabilidad (condicional) de incumplimiento en el tiempo t .

El valor del AVC es una función creciente tanto de la probabilidad de que la contraparte incumpla, como de la exposición esperada en el tiempo del incumplimiento. Puede verse en la ecuación 4.2 que una PI más elevada, una PDI más elevada y una EE más elevada incrementarían el AVC. El AVC de los bancos aumentó notablemente durante la crisis financiera. Las reformas regulatorias se orientaron a reducir la magnitud del AVC. El punto de partida sería reducir la probabilidad de incumplimiento de los bancos o disminuir las exposiciones esperadas.

4.2 EXPOSICIONES ESPERADAS CON REPOSICIÓN DE GARANTÍA Y SIN ELLA

EE_t es el valor esperado *estando bien de dinero (in the money)* del contrato. Si una contraparte está *in the money* en un contrato de derivados, $X_t > 0$

- Si no hay garantías, entonces:

$$4.3 \quad EE_t = E(\max\{X_t, 0\}).$$

La ecuación 4.3 muestra una exposición sin garantía. Es decir, la exposición esperada cuando la garantía no se intercambia.

Si introducimos la reposición de garantía (RG) que se calcula cada día (o dentro del día) conforme al mercado, la exposición esperada (EE) disminuye. Es así como las contrapartes en una transacción con derivados intercambian pérdidas y ganancias.

- Si hay reposiciones diarias de garantías:

$$4.4 \quad EE_t = E(\max\{X_t - RG, 0\}),$$

donde $RG_t = \max\{X_{t-1}, 0\}$.

La RG se basa en el valor del contrato el día previo. Para la contraparte que permite estar *in the money*, X_{t-1} es positiva. Las reposiciones de garantía están planeadas para garantizar que la exposición esperada nunca se vuelva excesiva.

Los márgenes iniciales (MI) tienen el objeto de cubrir las pérdidas potenciales en los contratos de derivados dentro del precio, en caso de que haya incumplimiento de la contraparte. Sus rangos se basan en un modelo que considera los movimientos del producto y los históricos del mercado. En el caso de las transacciones bilaterales, ambas partes pagan y los márgenes se segregan. Si hay liquidación central, todos los afiliados a la cámara de compensación pagan a un fondo común centralizado.

Con reposición diaria de garantía y un margen inicial,

$$4.5 \quad EE_t = E(\max\{X_{t+m} - RG - MI_t, 0\}),$$

donde el MI se fija como un VaR a n días, es decir, $P(X_n - X_0 > MI_n) = 1 - \alpha$, y suponemos un periodo de riesgo de m días de garantía.

Si hay incumplimiento de la contraparte, el contrato tendrá que ser reemplazado en el mercado. Debido a fluctuaciones en la liquidez del mercado, no siempre será posible reemplazar el contrato inmediatamente después del incumplimiento. El VaR a n días se basa en el número de días que tomaría hacerlo. El periodo de riesgo de m días de margen implicaría que la contraparte está expuesta al valor fluctuante del contrato m días en el futuro. El MI tiene como finalidad mitigar esta pérdida. Supongamos que tomara hasta cinco días encontrar una nueva contraparte para el contrato y que, en condiciones de mercado normales, no se espera que el precio se mueva más de cinco libras a cinco días con probabilidad del 99%. Entonces, si el margen inicial se fija en cinco libras, un operador quedará protegido contra el incumplimiento de una contraparte con una confianza del 99%. Un VaR del 95% implicaría que en 95 de cada 100 casos,

la contraparte contará con garantía suficiente para cubrir esta pérdida. En la práctica, el intervalo de confianza normalmente es del 99.7 por ciento.

La constitución de una garantía para derivados extrabursátiles en el mercado bilateral históricamente ha sido discrecional. Durante la crisis financiera de 2007-2009 se dieron elevadas exposiciones bilaterales que, en muchos casos, no gozaban de garantía suficiente. Una proliferación de contratos redundantes que se traslapaban exacerbó el riesgo de crédito de la contraparte. Desde entonces, las garantías han cobrado protagonismo en las transacciones con derivados extrabursátiles para mitigar el riesgo de crédito de la contraparte. Para fortalecer el mercado de derivados extrabursátiles, el G20 ha dictaminado que todos

los contratos estandarizados se compensen mediante contrapartes centralizadas y que se formulen normas para garantizar las operaciones que no se compensan centralmente. Los requerimientos de garantías más estrictos para las operaciones compensadas bilateralmente también mejorarán la administración del riesgo. Sin embargo, la obligatoriedad de la compensación centralizada para los derivados extrabursátiles, en conjunto con las reformas regulatorias propuestas, como las de Basilea III, posiblemente aumentarán la demanda de garantías en general. Sidanus y Zikes (2012) han hecho una valoración de cómo la reforma a los derivados extrabursátiles incrementará la demanda de activos de alta calidad como garantía.

6. APÉNDICE

CÁLCULO DE LA VARIABLE DE ESTADO NO OBSERVADO UTILIZANDO EL FILTRO DE KALMAN

El filtro de Kalman es una serie de ecuaciones que permite la actualización de un estimador en cuanto se dispone de una nueva observación. Primero forma una variable predictiva óptima del vector S de la variable de estado no observado dados sus valores calculados previamente. Estos cálculos de las variables de estado no observado posteriormente se actualizan utilizando la información proporcionada por las variables observadas.

Digamos que Y_t es una serie de variables observables m en un tiempo t . El modelo puede anotarse como:

$$\begin{aligned}
 1 \quad & Y_t = \beta s_t + \phi Y_{t-1} + \psi \varepsilon_{t-1} + \varepsilon_t \\
 & s_{t+1} = \alpha s_t + \eta_{t+1} \\
 & s_0 \sim N(S_0, p_0) \\
 & \varepsilon_t \sim N(0, I\sigma_\varepsilon^2) \\
 & \eta_t \sim N(0, 1) \\
 & t = 0, 1, \dots, T.
 \end{aligned}$$

Donde s_t es el proceso estocástico escalar que representa el estado de las condiciones macrofinancieras; β es el vector de dimensión m de cargas del factor; ϕ es una matriz diagonal $m \times m$ de los coeficientes AR(1); φ

es una matriz diagonal de los coeficientes MA(1); α es el coeficiente autorregresivo del proceso estocástico no observable s_t ; ε_t es un proceso de ruido blanco multivariado gaussiano con varianza diagonal y matriz de covarianza $I\sigma_\varepsilon^2$; y η_t es un ruido blanco gaussiano escalar con varianza unitaria.

Con el fin de cerrar el modelo, necesitamos especificar las condiciones iniciales del proceso estocástico s_t . Por lo tanto, suponemos que s_0 tendrá una distribución normal con una media S_0 y una varianza p_0 .

El factor común S_t en el tiempo t se obtiene como valor esperado del proceso s_t . Los parámetros del modelo pueden calcularse reescribiendo la ecuación 1 en forma de espacio de estado y usando el filtro de Kalman. Una representación de estado de espacio equivalente de 1 puede obtenerse como

$$2 \quad Y_t = F\theta_t,$$

$$3 \quad \begin{aligned} \theta_{t+1} &= G\theta_t + Ru_{t+1} \\ \theta_0 &\sim N(\Theta_0, P_0). \end{aligned}$$

La ecuación 2 es la ecuación de medición que relaciona las variables macroeconómicas y financieras observables con las variables de estado no observable (θ_t). La ecuación 3 es la ecuación de estado y consigna las dinámicas de la variable de estado latente θ_t . El algoritmo funciona como un proceso de dos pasos. En el paso de predicción, el filtro de Kalman genera estimaciones de las variables de estado actual, así como la incertidumbre correspondiente de las estimaciones. Una vez que la siguiente medición se observa (con algo de ruido), estas estimaciones se actualizan.

Con el modelo anotado en forma de espacio de estado, la estimación de los parámetros del modelo puede obtenerse mediante la maximización de la probabilidad.

El filtro de Kalman permite calcular la media condicional $E(\theta_t | Y_t, Y_{t-1}, \dots, Y_0) = \Theta_t$ y la varianza condicional $Var(\theta_t | Y_t, Y_{t-1}, \dots, Y_0) = P_t$.

Las recursiones de Kalman que nos permiten computar el factor común S_t como el primer elemento del vector Θ_t son:

$$4 \quad v_t = Y_t - F\Theta_t,$$

$$5 \quad Z_t = FP_tF',$$

$$6 \quad K_t = GP_tF'Z_t^{-1},$$

$$7 \quad L_t = G - K_tF,$$

$$8 \quad \Theta_{t+1} = G\Theta_t + K_tv_t,$$

$$9 \quad P_{t+1} = GP_tL_t' + W;$$

que se repiten a partir de $t=0$. La matriz K_t en 7 se conoce como la ganancia de Kalman.

Las recursiones de Kalman se obtienen a partir de la fórmula de la media condicionada y de la varianza condicionada en una distribución multinormal. Una distribución normal se caracteriza plenamente por sus dos primeros momentos y la función de probabilidad exacta se obtiene como subproducto del algoritmo del filtro de Kalman. Con el fin de inicializar el filtro, necesitamos especificar Θ_0 y P_0 . Una posible opción es suponer que estos parámetros son fijos y calcularlos utilizando la probabilidad máxima. Suponiendo que $\theta \sim N(\Theta_0, P_0)$, donde Θ_0 y P_0 son conocidos, y dejando que ψ denote la serie de parámetros que se van a calcular, es decir, $\psi = \{F, G, R\}$, la función de probabilidad se puede escribir como:

$$10 \quad L(\psi) = p(y_0) \prod_{t=1}^T p(y_t | Y_{t-1}).$$

Dado que:

$$E(y_t | Y_{t-1}) = E(F\theta_t | Y_{t-1}) = F\Theta_t$$

$$v_t = y_t - F\Theta_t$$

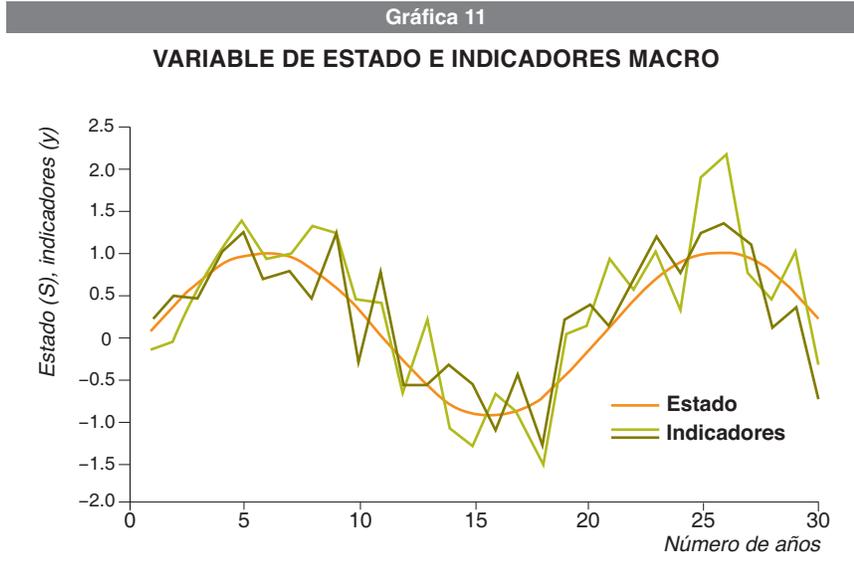
$$Var(y_t | Y_{t-1}) = Var(v_t) = W,$$

Entonces,

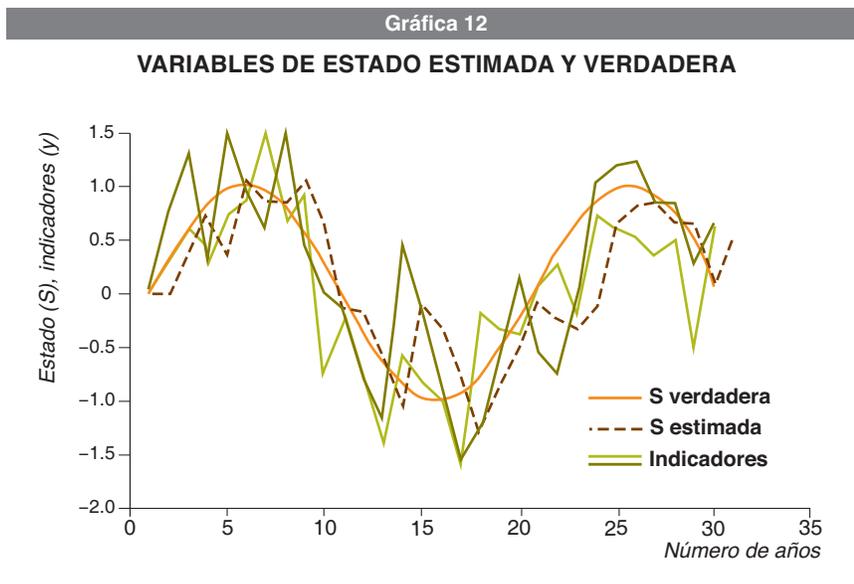
$$L(\psi) = \prod_{t=1}^T p(v_t) \quad v_t \sim N(0, W).$$

Por lo tanto, el valor de la función de probabilidad puede computarse directamente a partir de las recursiones de Kalman en las ecuaciones 4 a 9 y el máximo de $\log(L(\psi))$ puede obtenerse numéricamente.

La gráfica 11 muestra el factor común S de *estado verídico*, en torno al cual construimos los macroindicadores. Hemos considerado dos macroindicadores (las líneas verdes) en un escenario de 30 años. Han sido simulados a partir de una serie de números aleatorios y no se basan en datos reales.



La gráfica 12 muestra la S estimada junto con la S de *estado verídico*. Las ecuaciones recursivas de Kalman, 4 a 9, se han utilizado para computar el factor común estimado S .



BIBLIOGRAFÍA

- Comité de Supervisión Bancaria de Basilea (2005), *An Explanatory Note on the Basel II IRB Risk Weight Functions*, julio.
- Comité de Supervisión Bancaria de Basilea (2006), *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework—Comprehensive Version*, Banco de Pagos Internacionales, junio.
- Comité de Supervisión Bancaria de Basilea (2010), *Calibrating Regulatory Minimum Capital Requirements and Capital Buffers: A Top-down Approach*, octubre.
- Baz, J., y G. Chacko (2004), *Financial Derivatives*, Cambridge University Press.
- Durbin, J., y S. J. Koopman (2012), *Time Series Analysis by State Space Methods*, Oxford University Press.
- Gordy, M. B. (2003), “A Risk-factor Based Model Foundation for Ratings-based Bank Capital Rule”, *Journal of Financial Intermediation*, vol. 12, pp. 199-232.
- Gray, D., R. Merton y Z. Bodie (2008), “A New Framework for Measuring and Managing Macrofinancial Risk and Financial Stability”, Working Paper, Harvard Business School, núm. 9/15.
- Gregory, J. (2011), *Counterparty Credit Risk: The New Challenge for Global Financial Markets*, John Wiley & Sons.
- Haldane, A. (2011), “Capital Discipline”, ponencia en la American Economic Association, Denver, 9 de enero de 2011.
- Haldane, A., S. Hall y S. Pezzini (2007), “A New Approach to Assessing Risks to *financial stability*”, Financial Stability Paper, Banco de Inglaterra, núm. 2.
- Harvey, A. C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.
- Kealhofer, S. (2003), “Quantifying Credit Risk I: Default Prediction”, *Financial Analysts Journal*, vol. 59, núm. 1, pp. 30-44.
- Merton, RC, (1974), “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates”, *Journal of Finance*, vol. 51, pp. 987-1019.
- Sidanus, C., y F. Zikes (2012), *OTC Derivatives Reform and Collateral Demand Impact*, Financial Stability Paper, Banco de Inglaterra, núm. 18.
- Spiegeleer, Jan D., y W. Schoutens (2011), *Pricing CoCos. A Derivatives Approach*, Departamento de Matemáticas, Katholieke Universiteit Leuven.
- Vasicek, O. (2002), “Loan Portfolio Value”, *Risk*, diciembre, pp. 160-162.

Miembros del CEMLA

ASOCIADOS

Banco Central de la República Argentina	Banco Central de Reserva de El Salvador
Centrale Bank van Aruba	Banco de Guatemala
Central Bank of The Bahamas	Bank of Guyana
Central Bank of Barbados	Banque de la République d'Haïti
Central Bank of Belize	Banco Central de Honduras
Banco Central de Bolivia	Bank of Jamaica
Banco Central do Brasil	Banco de México
Eastern Caribbean Central Bank	Banco Central de Nicaragua
Cayman Islands Monetary Authority	Banco Central del Paraguay
Banco Central de Chile	Banco Central de Reserva del Perú
Banco de la República (Colombia)	Banco Central de la República Dominicana
Banco Central de Costa Rica	Centrale Bank van Suriname
Banco Central de Cuba	Central Bank of Trinidad and Tobago
Centrale Bank van Curaçao en Sint Maarten	Banco Central del Uruguay
Banco Central del Ecuador	Banco Central de Venezuela

COLABORADORES

Bancos centrales

Deutsche Bundesbank (Alemania)	Banca d'Italia
Bank of Canada	Bangko Sentral ng Pilipinas
Banco de España	Banco de Portugal
Federal Reserve System (Estados Unidos de América)	Sveriges Riksbank (Suecia)
Banque de France	Swiss National Bank
	European Central Bank

Otras instituciones

Superintendencia de Bancos y Seguros (Ecuador)	Turks and Caicos Islands Financial Services Commission
Superintendencia del Sistema Financiero (El Salvador)	Banco Centroamericano de Integración Económica
Superintendencia de Bancos de Guatemala	Banco Latinoamericano de Comercio Exterior, S. A.
Comisión Nacional de Bancos y Seguros (Honduras)	CAF-Banco de Desarrollo de América Latina
Superintendencia de Bancos de Panamá	Deutscher Genossenschafts- und Raiffeisenverband e.V. (Confederación Alemana de Cooperativas)
Superintendencia de Bancos (República Dominicana)	Fondo Latinoamericano de Reservas

CENTRO DE ESTUDIOS MONETARIOS LATINOAMERICANOS
Asociación Regional de Bancos Centrales

www.cemla.org

