



NOTAS DE INVESTIGACIÓN

Esta sección tiene por objetivo divulgar artículos breves escritos por economistas del Banco Central de Chile sobre temas relevantes para la conducción de las políticas económicas en general y monetarias en particular. Las notas de investigación, de manera frecuente, aunque no exclusiva, responden a solicitudes de las autoridades del Banco.

AGRUPACIÓN DE INSTITUCIONES BANCARIAS A PARTIR DEL ANÁLISIS DE *CLUSTER*: UNA APLICACIÓN AL CASO DE CHILE*

Alejandro Jara R.**
Daniel Oda Z.***

I. INTRODUCCIÓN

En un sistema bancario altamente heterogéneo como el chileno, el análisis agregado de indicadores financieros y de actividad puede generar sesgos importantes al momento de evaluar las potenciales vulnerabilidades y riesgos de la industria bancaria¹. En este contexto, esta nota aborda las siguientes preguntas: ¿Es posible agrupar las instituciones bancarias de forma tal que se reduzca este sesgo?, ¿son estos grupos suficientemente homogéneos, de forma tal que permitan la comparación al interior de cada grupo?, ¿es la agrupación sugerida estable en el tiempo? Para responder estas preguntas, se propone una métrica de agrupación basada en el análisis de *cluster*. Haciendo uso de información mensual de balances para el período 2008-2013, sobre un conjunto de 23 bancos en Chile, se encuentra que la industria bancaria puede agruparse en siete grupos de instituciones homogéneas: (i) multibancos grandes, (ii) multibancos medianos, (iii) banca especializada mediana, (iv) banca de consumo, (v) bancos de tesorería, (vi) bancos de comercio exterior, y (vii) banca dedicada a los servicios financieros.

La nota muestra además, que la agrupación de bancos sugerida es estable en el tiempo. Es decir, si se reestima mensualmente el análisis de *cluster*, las instituciones bancarias tienden a

* Esta nota representa exclusivamente las opiniones de sus autores y no necesariamente refleja la opinión del Consejo del Banco Central de Chile. Se agradecen los comentarios de Rodrigo Alfaro, Juan Francisco Martínez y Nancy Silva.

** Gerencia de Investigación Financiera, Banco Central de Chile: E-mail: ajara@bcentral.cl

*** Gerencia de Regulación e Infraestructura Financiera, Banco Central de Chile: E-mail: dodaze@bcentral.cl

¹ Esta heterogeneidad se manifiesta en varios aspectos (tamaño, exposición a los riesgos y origen de la propiedad, entre otros.) y es el reflejo, en gran medida, de la consolidación que enfrentó la industria bancaria durante la década de los noventa (Ahumada y Marshall, 2001). Este proceso de consolidación se tradujo en un aumento de la concentración bancaria (menor número de bancos), mayor presencia de bancos extranjeros y mayor exposición agregada al segmento de crédito minorista (retail).

permanecer en su gran mayoría dentro de los mismos grupos sugeridos. La estabilidad de los grupos de bancos, sumada a la propia naturaleza del análisis de *cluster*, contribuye a mejorar el monitoreo de riesgos de la industria bancaria, dado que permite comparar el comportamiento de los bancos con instituciones de características similares, tanto en términos de su estructura de balance como en términos de su exposición a los riesgos. Lo anterior se debe a que el análisis de *cluster* determina un agrupamiento natural (*cluster*) de las observaciones, de tal forma que en cada grupo se mantengan ciertas características comunes. La idea central es conseguir grupos que mantengan la mayor similitud entre sus miembros y que tengan la mayor diferencia (disimilitud) con respecto a los pares pertenecientes a otros grupos.

En esta nota aplicamos el análisis de *cluster* sobre un conjunto amplio de indicadores y características de los bancos (análisis multivariado). Estos indicadores son utilizados para identificar patrones específicos o correlaciones entre las variables, de forma tal que los grupos identificados incluyan entidades/objetos relativamente más homogéneos. En particular, utilizamos indicadores que capturan tres dimensiones o criterios del grado de exposición y mitigación a los riesgos bancarios: (i) tamaño, (ii) composición de la cartera y (iii) estructura de financiamiento².

El análisis de *cluster* de la industria bancaria se enmarca dentro de la vasta literatura empírica que aplica distintas técnicas estadísticas para analizar el comportamiento estratégico de los bancos (Koller, 2001). Buena parte de esta literatura se ha enfocado en economías desarrolladas o con sistemas bancarios profundos. Sin embargo, más recientemente, varios estudios han analizado la aplicación de técnicas de aglomeración estadística para identificar grupos de bancos en economías emergentes (Dardac y Boitan, 2009; Terrones y Vargas, 2013). En Chile, la mayoría de los estudios previos agrupan a los bancos de manera *ad hoc* (Wigodski y de la Maza, 2004; Jara y Oda, 2007; Banco Central de Chile, 2007) y, en este sentido, esta nota representa el primer esfuerzo sistemático para aplicar técnicas de aglomeración sobre la industria bancaria en Chile.

II. APLICACIÓN DEL ANÁLISIS DE CLUSTER

1. Metodología

La aplicación del análisis de *cluster* requiere cumplir con varias etapas. En primer lugar, es necesario definir el objetivo principal de la agrupación. La definición del objeto permite seleccionar el tipo adecuado de variables a considerar. En segundo lugar, las variables deben ser estandarizadas de forma tal de hacerlas comparables entre sí y evitar, de este modo, el sesgo generado por la presencia de observaciones extremas o *outliers*³. Luego se debe seleccionar un método específico de agrupación (ej.; jerárquico o particional) y un método específico de búsqueda. Finalmente, se debe elegir la unidad de medida o algoritmo para la distancia o similitud entre entidades.

² Los indicadores de estados de resultado, tales como el nivel de rentabilidad o el gasto operacional, no se incluyen en el análisis por varias razones. Por una parte, suelen ser bastante volátiles y en consecuencia se dificulta su utilización en el análisis de *cluster*. Por otra, estos indicadores incorporan el impacto de variables que no son controladas por los bancos (ej. tasa de política monetaria) y en consecuencia no reflejan necesariamente características estratégicas de los bancos (Amel y Rhoades, 1988).

³ La no estandarización lleva a que aquellas variables que se mueven en un rango mayor pesen más en la agrupación, sesgando el análisis de *cluster*.



Para efectos de esta nota, el principal objetivo de la agrupación propuesta es contribuir a una mejor evaluación de los riesgos que enfrentan los bancos en Chile. En consecuencia, las variables seleccionadas capturan distintas dimensiones de estos riesgos, así como la capacidad de las instituciones para mitigarlos. En cuanto a la estandarización de las variables, estas son normalizadas a media cero y varianza uno, con el objetivo de evitar que las variables que presentan mayor escala dominen a las demás, un problema particularmente importante en la aplicación de la distancia euclidiana.

En relación con la agrupación, se utiliza el método jerárquico en vez del método particional. La principal razón es que los algoritmos jerárquicos obtienen los grupos a partir de clasificaciones previamente establecidas, mientras que los algoritmos particionales dividen simultáneamente las observaciones en grupos no traslapados, que requieren de la especificación del número de grupos a priori. Dado que para nuestro análisis es importante determinar las relaciones segmentadas entre los grupos, optaremos por los métodos jerárquicos⁴. Más específicamente, usamos la aplicación de Ward (1963) para métodos jerárquicos aglomerativos (*bottom-up*) basada en la minimización de la suma de errores al cuadrado⁵. La razón para usar este método por sobre el método jerárquico divisivo (*top-down*) es que el método jerárquico aglomerativo parte con cada observación como un grupo separado y luego las vincula sucesivamente en grupos más grandes. Los métodos jerárquicos divisivos, en cambio, parten con todas las observaciones como un gran grupo y luego lo divide sucesivamente en grupos más pequeños, lo que se traduce en un método más complejo desde el punto de vista computacional⁶.

Con respecto a la medida de distancia, la nota utiliza la distancia euclidiana por ser la más utilizada en la literatura y poseer una interpretación natural de distancia⁷. En particular, considera el set de m indicadores a ser utilizados en el análisis de *cluster*, de tal forma que $x_{m,t}^i$ representa al indicador m normalizado a media uno y varianza cero del banco o grupo de bancos i en el mes t . De esta forma, haciendo uso de la versión cuadrática de la métrica general de Minkowski, la distancia entre el banco o grupo de bancos i y el banco o grupo de bancos j ($dist_{i,j}$) en el mes t , es equivalente a:

$$dist_{i,j,t} = \left\{ \sum_{k=1}^m (x_{k,t}^i - x_{k,t}^j)^2 \right\}^{1/2} \quad (1)$$

Si bien con esta metodología no es posible obtener a priori una métrica que nos indique la importancia relativa de cada variable m , sí es posible determinar la relevancia de la misma en el resultado *ex post*, es decir una vez obtenida la agrupación óptima. De hecho, haciendo

4 Una ventaja de utilizar clusters no jerárquicos es la estabilidad y robustez de los grupos en presencia de outliers (Koller, 2001). Sin embargo, dicha metodología no permite distinguir subgrupos, ni permite establecer el número óptimo de grupos.

5 Esta metodología es recomendable para grupos normales multivariados y esféricos (Kaufman y Rousseeuw, 1990).

6 El método aglomerativo considera $N(N-1)/2$ posibles fusiones para encontrar el par más próximo, y este número crece cuadráticamente con N ; mientras que en el método divisivo, el primer paso para dividir los grupos considera $2(N-1)-1$ comparaciones, cuyo número crece exponencialmente con N (Kaufman y Rousseeuw, 1990).

7 Adicionalmente, la distancia euclidiana se comporta bien cuando el conjunto de datos tiene conglomerados compactos y aislados (Mao y Jain, 1996).

uso de la distancia euclidiana de la ecuación (1), el aporte relativo de cada variable k en el grupo de bancos i en el mes t ($\%x_{k,t}^i$), es equivalente a:

$$\%x_{k,t}^i = \frac{(x_{k,t}^i - x_{k,t}^{-i})^2}{\sum_{k=1}^m (x_{k,t}^i - x_{k,t}^{-i})^2} \quad (2)$$

donde $x_{k,t}^i$ representa el valor promedio de la variable k en el resto de los bancos o grupos de bancos distintos de i .

2. Los datos

El análisis de *cluster* presentado en la siguiente sección se realiza para el período 2008-2013 a un grupo de 23 bancos en Chile sobre la base de indicadores de balances individuales calculados mensualmente. Durante el período de análisis, existen bancos nuevos (entrantes) y otros que dejan de operar o son absorbidos por otros bancos (salientes). Dado que los indicadores contables de este tipo de bancos suelen presentar fuertes distorsiones durante el proceso de entrada y salida del sistema, se excluyen los dos primeros años de observaciones de los bancos entrantes, y los dos últimos años de observaciones de los bancos salientes⁸. Es por este motivo que el sistema queda reducido a 23 instituciones bancarias⁹.

Dado que el objetivo es agrupar las instituciones bancarias de acuerdo con su exposición a los riesgos, se utilizan indicadores de tamaño, estructura de los activos y estructura de los pasivos. El tamaño del banco tiene implicancias importantes en el potencial de diversificación y, por lo tanto, en la capacidad de manejo del riesgo; además, ha sido tradicionalmente la principal medida de clasificación de la banca. Por otra parte, la estructura de los pasivos permite tener una aproximación al riesgo de liquidez y financiamiento, en tanto la estructura de los activos permite capturar el riesgo de crédito y el riesgo de mercado.

En consecuencia, el set de indicadores utilizados para el análisis de *cluster* es el siguiente:

Indicadores de tamaño: (1) activos como porcentaje de los activos totales del sistema (t_{act}), (2) colocaciones como porcentaje de las colocaciones totales del sistema (t_{col}) y (3) patrimonio como porcentaje del patrimonio total del sistema (t_{pat}).

Indicadores de la estructura de pasivos: (1) patrimonio como porcentaje de los activos (f_{lev}), (2) depósitos totales como porcentaje de los pasivos totales (f_{dep}), (3) depósitos vista como porcentaje de los depósitos totales (f_{vis}) y (4) financiamiento externo como porcentaje de los pasivos totales (f_{ext}).

Indicadores de la estructura de los activos: (1) colocaciones totales como porcentaje de los activos totales (c_{cot}), (2) colocaciones comerciales como porcentaje de los activos totales

⁸ Nótese que si un banco absorbe a otro, sus indicadores contables suelen presentar cambios significativos a causa de la absorción o fusión. Estos cambios sí son considerados al momento de construir los grupos o clusters de bancos, ya que representan una de las principales razones de por qué una institución bancaria puede cambiar de grupo en el tiempo.

⁹ Como resultado de lo anterior, cuatro instituciones han sido excluidas del análisis (RBS, DnB Nor, ABN-Amro y Banco del Desarrollo).



(c_com), (3) colocaciones de consumo como porcentaje de los activos totales (c_con), (4) colocaciones de vivienda como porcentaje de los activos totales (c_viv), (5) colocaciones comerciales al exterior como porcentaje de las colocaciones comerciales totales (c_cex), (6) inversión en instrumentos financieros como porcentaje de los activos totales (c_fin), (7) posición larga en derivados como porcentaje de los activos totales (c_der), activos en moneda extranjera como porcentaje de los activos totales (c_me), y (8) activos líquidos disponibles como porcentaje de los activos totales (c_dis).

Correlaciones

Cuando alguna de las variables consideradas está altamente correlacionada con alguna de las otras variables, es recomendable incorporar solo una de ellas para evitar aumentar el peso de dicha característica, ya que ambas variables entregan la misma información. Para evitar este sesgo, calculamos la correlación simple entre el conjunto de indicadores a considerar (cuadro 1). Como se puede observar, los indicadores de tamaño están fuertemente correlacionados entre sí. De igual forma, se observa una alta correlación negativa entre los bancos con alto apalancamiento y aquellos con alto nivel de cuentas corrientes. Además, existe una fuerte correlación negativa entre aquellos que poseen una mayor composición de colocaciones como parte de sus activos y los que mantienen niveles significativos en instrumentos derivados¹⁰.

En consecuencia, el análisis de *cluster* presentado en la siguiente sección solo considera el tamaño medido como participación de activos y excluye las variables de participación de colocaciones y patrimonio. Asimismo, no considera la participación de las colocaciones en los activos, por estar ya considerada en la desagregación por cartera comercial, de consumo y de vivienda.

¹⁰ Las correlaciones mantienen el mismo orden de magnitud si consideramos la muestra entera como si solo analizamos un período en particular. Cabe señalar que nos interesan las diferencias entre bancos, por lo que es la dimensión transversal la que es relevante para el análisis.

Cuadro 1

Correlaciones^a

	t_act	t_col	t_pat	c_cot	c_com	c_con	c_viv	c_cex	c_fin	c_der	c_me	c_dis	f_lev	f_dep	f_vis	f_ext
t_act	1,00															
t_col	0,99	1,00														
t_pat	0,97	0,97	1,00													
c_cot	0,38	0,41	0,36	1,00												
c_com	0,29	0,32	0,28	0,54	1,00											
c_con	-0,07	-0,05	-0,07	0,49	-0,42	1,00										
c_viv	0,61	0,61	0,58	0,67	0,10	0,37	1,00									
c_cex	-0,11	-0,11	-0,09	-0,07	0,33	-0,34	-0,29	1,00								
c_fin	-0,19	-0,21	-0,22	-0,51	-0,28	-0,26	-0,30	-0,26	1,00							
c_der	-0,10	-0,14	-0,08	-0,56	-0,37	-0,23	-0,30	-0,16	0,10	1,00						
c_me	-0,27	-0,27	-0,25	-0,26	0,23	-0,40	-0,50	0,77	-0,19	-0,01	1,00					
c_dis	-0,25	-0,25	-0,22	-0,53	-0,29	-0,23	-0,40	0,31	-0,12	0,05	0,53	1,00				
f_lev	-0,35	-0,34	-0,29	-0,36	-0,13	-0,17	-0,43	0,43	-0,17	-0,07	0,63	0,64	1,00			
f_dep	0,15	0,17	0,11	0,39	0,13	0,30	0,22	0,11	0,08	-0,46	-0,27	-0,38	-0,41	1,00		
f_vis	0,29	0,29	0,28	-0,18	0,02	-0,29	0,05	0,23	-0,13	-0,13	0,09	0,38	0,21	0,01	1,00	
f_ext	-0,08	-0,07	-0,05	-0,05	0,21	-0,22	-0,21	0,20	-0,21	-0,21	0,47	0,45	0,53	-0,57	0,05	1,00

Fuente: Cálculo de los autores a partir de información entre enero del 2008 y mayo del 2013.

a. Las celdas marcadas representan correlaciones mayores que 0,5 o menores que -0,5.

Volatilidades

De igual forma, cuando los indicadores o variables son altamente volátiles distorsionan los resultados de la metodología de *cluster* en el análisis temporal. El cuadro 2 presenta la volatilidad (medida como la desviación estándar) de los indicadores sugeridos durante el período de análisis. Como se puede observar, algunas características de los bancos cambian fuertemente a través del tiempo. Por ejemplo, los descalses de moneda o instrumentos derivados suelen ajustarse rápidamente a las estrategias de inversión del banco y a las variaciones de mercado. De la misma forma, el manejo de la liquidez también es una característica que puede ser poco estable para algunos bancos, principalmente en aquellos dedicados a negocios de tesorería. Adicionalmente, existen algunos bancos cuya línea de negocio es cambiante, con alta variación en el otorgamiento de créditos comerciales, que pueden ser más susceptibles de cambiar de grupo. En consecuencia, el análisis de *cluster* de la siguiente sección toma reparos con las variables inestables como: instrumentos financieros, activos líquidos y depósitos a la vista.



Cuadro 2

Volatilidades^a

Banca tradicional																
	t_act	t_col	t_pat	c_cot	c_com	c_con	c_viv	c_cex	c_fin	c_der	c_me	c_dis	f_lev	f_dep	f_vis	f_ext
Chile	0,003	0,004	0,006	0,022	0,010	0,005	0,018	0,018	0,009	0,013	0,015	0,009	0,005	0,014	0,037	0,009
Internacional	0,002	0,001	0,001	0,030	0,030	0,010	0,006	0,026	0,039	0,009	0,020	0,013	0,005	0,031	0,018	0,024
Estado	0,006	0,008	0,008	0,025	0,021	0,005	0,016	0,017	0,029	0,012	0,022	0,024	0,004	0,038	0,026	0,018
Scotiabank	0,009	0,012	0,020	0,074	0,056	0,012	0,028	0,078	0,019	0,006	0,018	0,009	0,075	0,068	0,036	0,019
BCI	0,003	0,002	0,011	0,020	0,012	0,006	0,007	0,020	0,021	0,010	0,028	0,012	0,009	0,020	0,039	0,010
Corpbanca	0,008	0,005	0,013	0,047	0,032	0,022	0,010	0,019	0,034	0,008	0,021	0,009	0,011	0,022	0,015	0,012
Bice	0,001	0,001	0,001	0,024	0,022	0,002	0,005	0,030	0,024	0,013	0,031	0,021	0,005	0,035	0,046	0,012
Santander	0,008	0,007	0,009	0,026	0,017	0,006	0,008	0,020	0,019	0,011	0,018	0,019	0,004	0,031	0,044	0,023
Itaú	0,004	0,005	0,005	0,027	0,017	0,005	0,015	0,027	0,024	0,010	0,026	0,015	0,009	0,031	0,023	0,017
Security	0,001	0,001	0,001	0,046	0,028	0,010	0,019	0,017	0,034	0,005	0,018	0,017	0,005	0,029	0,016	0,016
Falabella	0,001	0,001	0,000	0,046	0,015	0,040	0,020	0,000	0,028	0,020	0,025	0,024	0,010	0,052	0,033	0,000
Ripley	0,000	0,001	0,000	0,037	0,014	0,026	0,031	0,000	0,046	0,000	0,001	0,013	0,026	0,056	0,008	0,000
Paris	0,000	0,000	0,000	0,029	0,000	0,028	0,008	0,000	0,017	0,000	0,000	0,015	0,015	0,012	0,003	0,000
BBVA	0,002	0,003	0,002	0,024	0,032	0,015	0,016	0,017	0,024	0,023	0,013	0,013	0,005	0,023	0,029	0,011
Volatilidad	Baja	Baja	Baja	Media	Media	Baja	Baja	Media	Media	Baja	Baja	Baja	Baja	Media	Media	Baja
Incluido	✓			✓	✓	✓	✓	✓	✓	✓			✓	✓		✓

Banca de tesorería																
	t_act	t_col	t_pat	c_cot	c_com	c_con	c_viv	c_cex	c_fin	c_der	c_me	c_dis	f_lev	f_dep	f_vis	f_ext
DoBrasil	0,000	0,000	0,000	0,141	0,141	0,000	0,000	0,080	0,000	0,017	0,036	0,042	0,125	0,114	0,127	0,081
HSBC	0,002	0,000	0,002	0,029	0,030	0,001	0,001	0,052	0,088	0,076	0,065	0,070	0,013	0,107	0,067	0,014
JP Morgan	0,001	0,000	0,003	0,008	0,008	0,000	0,000	0,000	0,108	0,104	0,097	0,107	0,077	0,097	0,218	0,000
BNA	0,000	0,000	0,000	0,035	0,035	0,001	0,000	0,064	0,073	0,000	0,075	0,169	0,068	0,206	0,133	0,213
Tokyo	0,000	0,000	0,001	0,041	0,041	0,000	0,000	0,118	0,128	0,048	0,088	0,146	0,051	0,078	0,167	0,004
Deutsche	0,002	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,060	0,070	0,073	0,054	0,019	0,085	0,011	0,004
Rabobank	0,001	0,001	0,001	0,082	0,082	0,000	0,000	0,024	0,085	0,016	0,128	0,052	0,020	0,084	0,056	0,132
Consortio	0,003	0,001	0,005	0,073	0,084	0,046	0,024	0,014	0,122	0,022	0,053	0,043	0,037	0,188	0,057	0,007
Penta	0,002	0,001	0,002	0,125	0,125	0,000	0,000	0,034	0,111	0,043	0,032	0,038	0,076	0,158	0,095	0,004
Volatilidad	Baja	Baja	Baja	Alta	Alta	Baja	Baja	Media	Alta	Media	Alta	Alta	Alta	Alta	Alta	Alta
Incluido	✓							✓	✓				✓			

Fuente: Cálculo de los autores a partir de información entre enero 2008 y mayo 2013.

a. Corresponde a la desviación estándar. Nótese que no se consideran algunas de las variables con volatilidad media o baja, por estar correlacionadas con otras (ver cuadro 1).

III. RESULTADOS

Esta sección presenta los resultados de la aplicación del análisis de *cluster* sobre la banca en Chile a partir del análisis de dendrogramas de disimilitud de la metodología de Ward (1963) sobre la distancia euclidiana. Un dendrograma es una representación gráfica en forma de árbol que resume el proceso de agrupación en un análisis de *cluster*. Los objetos similares se conectan mediante enlaces cuya posición en el diagrama está determinada por la distancia entre los objetos¹¹.

El análisis se realiza en dos etapas. La primera consiste en un análisis general, cuyo objetivo es distinguir dos grandes grupos de bancos (típicamente distinguidos como banca universal o comercial y banca de tesorería). La segunda es un análisis por separado para la banca comercial y la banca de tesorería, haciendo uso de un set de variables distintas para cada caso. La razón de aplicar la metodología en dos etapas radica en el peso que tiene cada variable en el cálculo de la distancia. Las variables que permiten identificar grupos de bancos en la banca comercial no aportan información o introducen ruido en la identificación de la banca de tesorería, y viceversa.

A continuación se muestran los resultados para el período entre enero del 2008 y mayo del 2013. Cabe señalar que se muestran, a modo ilustrativo, gráficos y cuadros a marzo del 2013.

1. Análisis general

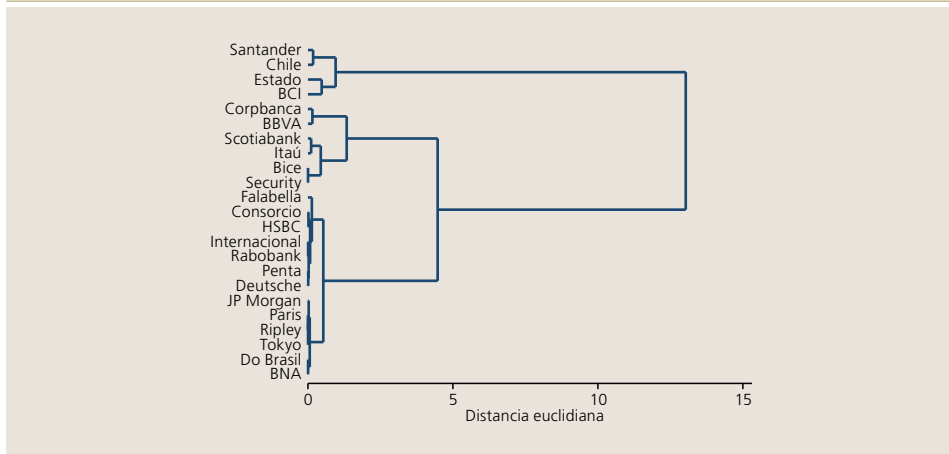
Agrupación jerárquica

El gráfico 1 muestra el dendrograma de agrupación utilizando exclusivamente variables de tamaño. Del gráfico 1 se desprende que, para una distancia igual o superior a 5, es posible distinguir dos grupos de bancos. El primer grupo está compuesto por los grandes bancos (Santander, Chile, Estado y BCI), mientras el segundo grupo incluye el resto de las instituciones bancarias. En este sentido, esta agrupación es poco informativa dado que el segundo grupo mantiene un alto nivel de heterogeneidad. En particular, bancos como Banco Ripley y Banco de Tokio o JP Morgan, aparecen juntos a pesar de que sus líneas de negocio o estructura de balance difieren significativamente.

¹¹ Es importante mencionar, que en este caso, la distancia no se mide en unidades físicas, sino en términos de semejanza entre las características intrínsecas de las entidades consideradas. Aquí, la distancia euclidiana mide la diferencia entre grupos. Es decir, un mayor valor representa una diferencia mayor.

Gráfico 1**Dendrograma de agrupación: modelo general**

(solo tamaño)

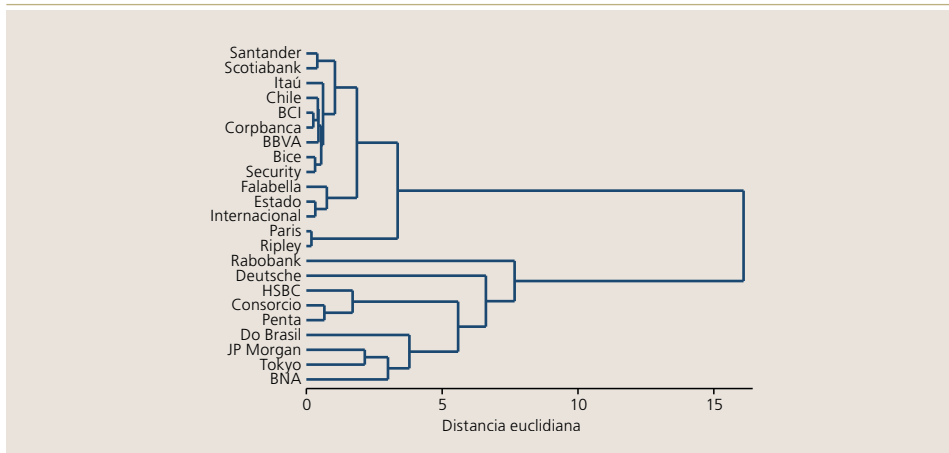


Fuente: Cálculo de los autores a partir de información a marzo del 2013.

Nota: Aplicación de la metodología de Ward (1963) haciendo uso de la distancia euclidiana.

Gráfico 2**Dendrograma de agrupación: modelo general**

(todas las variables)



Fuente: Cálculo de los autores a partir de información a marzo del 2013.

Nota: Aplicación de la metodología de Ward (1963) haciendo uso de la distancia euclidiana.

Para abordar la heterogeneidad existente en las líneas de negocio o estructura de balance, el gráfico 2 muestra el dendrograma de agrupación que surge al obviar la variable tamaño y considerar en cambio las siguientes variables: (1) colocaciones totales como porcentaje de los activos totales,

(2) activos en moneda extranjera como porcentaje de los activos totales, (3) derivados (valor razonable de activos) como porcentaje de los activos totales, (4) patrimonio como porcentaje de los activos totales y (5) financiamiento externo como porcentaje de los pasivos totales.

Del gráfico 2 se desprende que para una distancia superior a 8 sea posible identificar dos grandes grupos de bancos: la banca comercial (ej.: Santander, Banco Estado, Banco Paris) y la banca de tesorería (ej.: Rabobank, Penta, Tokio).

Estabilidad grupal en el tiempo

El cuadro 3 muestra el porcentaje de veces que una institución bancaria pertenece en el tiempo al grupo de bancos identificados en el gráfico 2 (banca comercial y banca de tesorería), es decir, cuando solo se utilizan indicadores de la estructura de balance. Del cuadro se desprende que las agrupaciones sugeridas son bastante estables en el tiempo, a pesar de que no consideran la variable de tamaño en su construcción. Casi la totalidad de los bancos permanecen en el mismo grupo el 100% de las veces. De esta forma, podemos distinguir y dividir los bancos (sin considerar el tamaño relativo) en dos grupos: (1) *banca comercial o universal*, es decir, aquellos bancos enfocados en el otorgamiento de crédito comercial, de consumo e hipotecario y que se financian principalmente con depósitos y (2) *banca de tesorería*, es decir, aquellos bancos especializados en créditos comerciales sectoriales, comercio exterior, inversión financiera y mercado cambiario y que se financian principalmente de créditos en el exterior y capital propio.

Cuadro 3

Ubicación en la agrupación primaria de bancos^a (porcentaje del período total)

Banco	Comercial	Tesorería	Banco	Comercial	Tesorería
Bice	100		Corpbanca	100	
BBVA	100		Scotiabank	100	
Chile	100		Rabobank	15	85
BCI	100		Consorcio	5	95
Estado	100		Penta	2	98
Falabella	100		BNA		100
Internacional	100		DoBrasil		100
Itaú	100		Deutsche		100
Paris	100		HSBC		100
Ripley	100		JP Morgan		100
Santander	100		Tokyo		100
Security	100				

Fuente: Cálculo de los autores a partir de información entre enero del 2008 y mayo del 2013.

a. Porcentaje de meses del total de la muestra en los que la institución pertenece al grupo específico.

En las líneas siguientes, es razonable considerar que las variables que caracterizan a la banca comercial no necesariamente son aquellas que caracterizan a la banca de tesorería. Por ejemplo, la composición de la cartera de colocaciones no es un factor relevante para la banca de tesorería, debido a que su exposición a créditos de consumo o vivienda es prácticamente inexistente; mientras que, al utilizar estas variables en la agrupación de la banca comercial permite distinguir entre bancos enfocados a un tipo específico de crédito. En particular, las variables consideradas son aquellas que presentan una volatilidad media o baja en cada subgrupo de la banca comercial y de tesorería por separado, y que no están fuertemente correlacionadas con otras variables (ver cuadro 2 para más detalles).

2. Banca comercial o tradicional y banca de tesorería

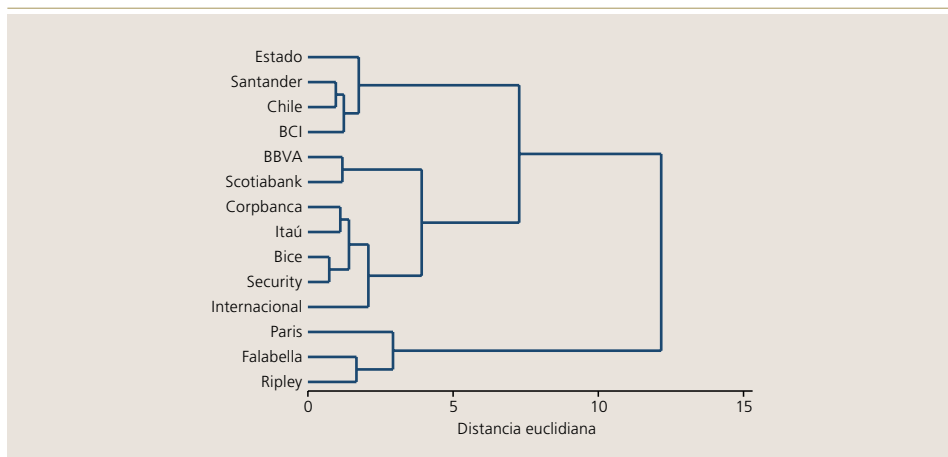
Agrupación jerárquica

En el gráfico 3 se puede observar el árbol de agrupación de los bancos comerciales utilizando indicadores sobre la composición general del portafolio del banco y de su participación de mercado: (1) participación de mercado medida a partir de los activos totales, (2) participación de las colocaciones por tipo en el total de las colocaciones, (3) participación de los instrumentos financieros no derivados en los activos totales, (4) participación de las colocaciones de comercio exterior sobre el total de colocaciones a empresas, (5) participación del valor nominal de los instrumentos derivados sobre los activos totales, (6) participación de los depósitos en los pasivos, (7) participación del financiamiento externo en los pasivos y (8) participación del capital básico en los activos totales.

Para una distancia igual o superior a 5, es posible identificar tres grupos de bancos comerciales homogéneos: (1) multibancos grandes, (2) bancos medianos y (3) la denominada banca de personas (Bancos Paris, Falabella y Ripley). Para un nivel de tolerancia algo menor, la banca mediana es posible separarla en dos subgrupos: (1) multibancos medianos y (2) banca especializada mediana.

Gráfico 3

Dendrograma de agrupación: banca comercial



Fuente: Cálculo de los autores a partir de información a marzo del 2013.

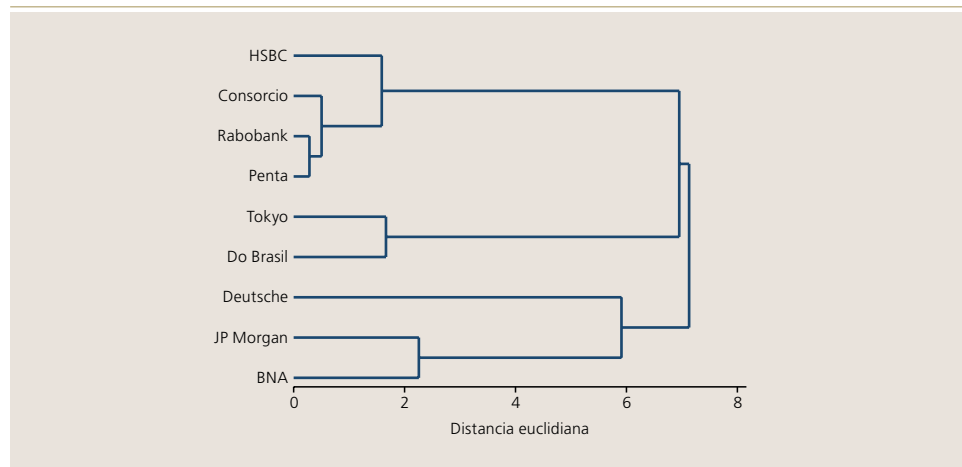
Nota: Aplicación de la metodología de Ward (1963) haciendo uso de la distancia euclidiana.

En el gráfico 4 se puede observar la agrupación de los bancos de tesorería al utilizar los siguientes indicadores: (1) participación de mercado medida a partir de los activos totales, (2) participación de las colocaciones de comercio exterior sobre el total de colocaciones a empresas, (3) participación del valor nominal de los instrumentos derivados sobre los activos totales y (4) participación del capital básico en los activos totales.

Para una distancia igual o superior a 6 es posible identificar a tres grupos de bancos de tesorería relativamente homogéneos: (1) bancos de tesorería (Deutsche, JP Morgan y BNA), (2) bancos de comercio exterior (Tokyo y Do Brasil) y (3) bancos basados en la administración de instrumentos financieros (HSBC, Consorcio, Rabobank y Penta).

Gráfico 4

Dendrograma de agrupación: banca de tesorería



Fuente: Cálculo de los autores a partir de información a marzo del 2013.

Nota: Aplicación de la metodología de Ward (1963) haciendo uso de la distancia euclidiana.



¿Existe un número óptimo de *clusters*? Notar que la definición del número de grupos analizados en la sección anterior no está predeterminada y requiere, en la práctica, del criterio y la experiencia de quien decide. Existe, sin embargo, extensa literatura que analiza distintas estrategias para identificar el número de grupos óptimos (por ejemplo, Milligan y Cooper, 1985). De estas estrategias, los criterios de Calinski y Harabasz (1974) y de Duda-Hart (2001) son considerados los más efectivos para la identificación del número óptimo de grupos en el análisis de *cluster* jerárquico.

El criterio de Calinski y Harabasz (*CH*) equivale a la razón de la varianza entre grupos SS_b y la varianza dentro del grupo (SS_w), ajustado por el número de grupos (k) y el número de observaciones (N):

$$CH = \frac{SS_b}{SS_w} \frac{(N-k)}{k-1} \quad (3)$$

En este caso, el número de *cluster* óptimo es aquel que genera la mayor varianza entre grupos y la menor varianza dentro del grupo¹².

El criterio de Duda-Hart (*DH*) también representa una razón de varianzas. En este caso, para cada partición del proceso jerárquico (por ejemplo, al pasar de dos a tres grupos) se calcula la razón entre la suma de los errores al cuadrado del grupo antes de la partición ($Je(1)$) y la suma de los errores al cuadrado de los dos grupos resultantes ($Je(2)$):

$$DH = \frac{Je(2)}{Je(1)} \quad (4)$$

En consecuencia, mayores valores de esta razón de varianzas indican grupos claramente distintos¹³.

El cuadro 4 muestra el cálculo de estos criterios para el período 2008-2013, separando el análisis para la banca comercial y la banca de tesorería. Los valores destacados representan el número óptimo de grupos, de acuerdo con cada criterio. Como se puede observar, los criterios no entregan resultados concluyentes. En el caso de la banca comercial, el número óptimo de grupos puede ir de dos a cinco en el 2013 y de tres a seis en el 2012, según el criterio utilizado.

En el caso de la banca de tesorería, el criterio *CH* no es informativo, ya que su valor crece monotónicamente; mientras que el criterio *DH* entrega valores óptimos entre tres y siete, dependiendo de la fecha de su evaluación.

12 Cabe notar que, al igual que la mayoría de los criterios de aglomeración, la aplicación de *CH* representa un método heurístico. En consecuencia, no existe un valor crítico o límite aceptable que determine el número óptimo de grupos. En la práctica, lo que se sugiere es comparar los valores de *CH* y escoger el número de grupos que maximice el valor de *CH*, siempre que exista un máximo. Si en cambio los valores de *CH* aumentan o disminuyen en forma monotónica, no es posible elegir un número de grupos óptimo a partir de este criterio.

13 Cabe notar que *DH* es cero cuando los dos subgrupos resultantes de la partición tienen varianza cero. Un ejemplo de esto es cuando el grupo que está siendo dividido tiene dos valores distintos que están siendo divididos en subgrupos únicos. En este caso, $Je(1)$ nunca será cero ya que no se dividen grupos que no tienen variabilidad.

Cuadro 4

Pruebas de validación^{a,b}

Marzo 2013

Comercial		
N	Calinski/ Harabast	Duda/ Hart
2	13,13	0,49
3	14,29	0,44
4	14,09	0,25
5	15,27	0,46

Tesorería		
N	Calinski/ Harabast	Duda/ Hart
2	3,24	0,15
3	4,28	0,13
4	15,41	0,00
5	19,33	0,00

Marzo 2012

Comercial		
N	Calinski/ Harabast	Duda/ Hart
2	13,21	0,45
3	16,12	0,49
4	14,59	0,18
5	16,06	0,49

Tesorería		
N	Calinski/ Harabast	Duda/ Hart
2	4,38	0,59
3	5,09	0,25
4	10,58	0,00
5	16,36	0,12

Marzo 2010

Comercial		
N	Calinski/ Harabast	Duda/ Hart
2	13,43	0,46
3	15,73	0,12
4	14,38	0,51
5	15,38	0,48

Tesorería		
N	Calinski/ Harabast	Duda/ Hart
2	6,15	0,49
3	7,88	0,42
4	10,63	0,00
5	13,17	0,00

Marzo 2008

Comercial		
N	Calinski/ Harabast	Duda/ Hart
2	12,45	0,53
3	13,78	0,28
4	11,90	0,57
5	12,05	0,23

Tesorería		
N	Calinski/ Harabast	Duda/ Hart
2	8,80	0,27
3	9,88	0,09
4	16,39	0,29
5	19,63	0,00

Fuente: Cálculo de los autores.

a. Mayor valor indica *clusters* más distintos.

b. Las celdas sombreadas corresponden al grupo óptimo, según cada criterio.

Agrupación sugerida para la industria bancaria de Chile

Dado que no existe un criterio estadístico único para identificar el número óptimo de grupos, se sugiere la siguiente estrategia: (1) seleccionar un número de grupos a partir del análisis de los dendrogramas y los criterios de optimalidad descritos anteriormente, (2) evaluar el grado de diferenciación entre grupos, (3) evaluar su estabilidad en el tiempo, (4) evaluar el grado de similitud al interior del grupo y (5) evaluar la relevancia de los grupos en términos de reducir el sesgo que se genera al analizar las variables/indicadores en términos agregados.



Respecto al grupo de bancos sugeridos, se opta por seleccionar cuatro grupos en el caso de la banca comercial (multibancos grandes, multibancos medianos, banca especializada y banca de personas) y tres grupos en la banca de tesorería (banca de tesorería propiamente tal, banca dedicada al comercio exterior y la banca de servicios financieros).

¿Son estos grupos suficientemente distintos entre ellos? Los cuadros 5 y 6 intentan responder esta pregunta. En el cuadro 5 se muestra el valor promedio de cada indicador utilizado en la aplicación del análisis de *cluster*. En ella se puede observar que la variable tamaño (activos como porcentaje de los activos del sistema) es una de las principales variables que identifica el grupo de multibancos grandes (16%), lo mismo que el porcentaje que representa los préstamos de vivienda sobre las colocaciones totales y la importancia relativa de los depósitos a la vista sobre los depósitos totales. Los multibancos medios se diferencian de la banca especializada en términos de su exposición a los préstamos hipotecarios y su menor exposición relativa a la banca comercial.

El cuadro 6, por otra parte, muestra la variable/indicador que mejor explica el cálculo de la distancia promedio de cada grupo. De ella se desprende, por ejemplo, que la banca de tesorería se caracteriza por una mayor exposición a los instrumentos derivados (cerca del 50%), mientras que en la banca de comercio exterior, la exposición a este tipo de préstamos explica en mayor medida el cálculo de la distancia de este grupo. La distancia del grupo de bancos de servicios financieros, por otro lado, se explica principalmente por su exposición a instrumentos financieros como porcentaje de los activos totales.

Cuadro 5

Indicador promedio por grupo de bancos^{a,b} (porcentaje)

	Multibanco grande	Multibanco medio	Especial media	Personas	Tesorería	Comercio exterior	Servicios financieros	Total
Activos/Activos del sistema	16,33	5,67	3,27	0,46	0,77	0,05	0,58	4,28
	(2,89)	(1,53)	(1,97)	(0,37)	(0,37)	(0,04)	(0,39)	(5,99)
Comercial/Activos	42,14	41,74	57,97	2,26	1,75	32,86	27,30	33,04
	(7,52)	(4,57)	(5,71)	(2,40)	(1,85)	(27,03)	(20,98)	(23,70)
Consumo/Activos	9,65	7,50	4,88	64,53	0,00	0,12	1,23	12,04
	(2,22)	(1,98)	(3,79)	(16,32)	(0,00)	(0,16)	(3,10)	(21,58)
Vivienda/Activos	19,22	22,26	9,15	16,77	0,00	0,00	0,68	9,57
	(5,15)	(2,80)	(5,24)	(7,60)	(0,00)	(0,00)	(1,68)	(9,48)
Comex/Comercial	11,17	15,82	11,43	0,00	0,00	45,86	10,97	13,69
	(2,78)	(5,87)	(3,35)	(0,00)	(0,00)	(21,33)	(10,22)	(16,24)
Ins. Financieros/Activos	12,12	9,56	13,23	7,92	20,16	8,41	37,55	16,23
	(5,77)	(2,13)	(4,53)	(4,43)	(8,74)	(10,89)	(17,12)	(13,86)
Derivados/Activos	3,66	4,42	1,93	0,55	39,09	1,91	6,08	6,22
	(2,22)	(3,19)	(1,00)	(1,40)	(12,48)	(3,47)	(8,46)	(11,65)
Activos ME/Activos	12,69	12,55	15,30	1,90	22,44	59,08	23,45	20,61
	(3,85)	(1,59)	(2,92)	(2,98)	(8,56)	(24,99)	(14,12)	(19,77)
Disponible/Activos	5,03	3,20	3,97	3,94	17,50	27,17	7,48	8,90
	(1,83)	(1,14)	(1,84)	(2,41)	(12,08)	(19,79)	(5,70)	(11,64)
Patrimonio/Activos	7,20	10,71	7,64	12,66	23,10	48,93	11,87	15,95
	(1,20)	(6,76)	(1,54)	(2,81)	(16,26)	(22,31)	(5,96)	(16,76)
Depósitos/Pasivos	67,17	59,66	70,55	72,24	23,94	68,72	59,67	63,05
	(4,59)	(6,91)	(5,19)	(13,92)	(9,59)	(24,06)	(23,28)	(19,56)
Dep. Vista/Depósitos	31,16	20,25	15,27	4,86	14,52	39,95	10,60	19,45
	(3,87)	(4,52)	(5,03)	(4,23)	(20,34)	(23,47)	(7,40)	(15,97)
Fin. Externo/Pasivos	8,45	10,59	6,82	0,00	0,03	20,36	13,25	8,84
	(3,77)	(3,31)	(2,45)	(0,01)	(0,31)	(28,60)	(22,96)	(15,53)

Fuente: Cálculo de los autores a partir de información entre enero del 2008 y mayo del 2013.

a. Desviación estándar entre paréntesis.

b. Indicador promedio ponderado por grupo de bancos, expresado en porcentaje. La intensidad del color verde indica grupos con un mayor indicador; la intensidad del color naranja, grupos con indicadores más bajos.

Cuadro 6
Importancia relativa de indicadores en cálculo de la distancia^{a,b}
 (porcentaje)

	Multibanco grande	Multibanco medio	Especial media	Personas	Tesorería	Comercio exterior	Servicios financieros
Activos/Activos del sistema	63,36	1,96	1,46	3,41	1,97	3,22	8,29
Comercial/Activos	2,21	5,01	52,23	13,45	9,40	0,18	2,12
Consumo/Activos	0,19	1,71	5,66	49,88	1,79	1,98	5,54
Vivienda/Activos	15,56	60,29	0,19	4,74	5,58	6,31	18,32
Comex/Comercial	0,37	2,27	1,24	5,65	3,85	23,75	0,72
Ins. Financieros/Activos	1,50	8,37	3,18	2,98	1,25	2,81	47,85
Derivados/Activos	0,77	1,00	6,93	2,07	48,07	1,02	0,57
Activos ME/Activos	1,84	4,17	2,71	5,44	0,37	17,40	0,82
Disponible/Activos	1,31	7,20	7,40	1,22	4,68	11,84	1,03
Patrimonio/Activos	3,09	3,10	8,69	0,26	1,11	17,26	1,10
Depósitos/Pasivos	0,69	2,72	6,00	1,85	17,88	1,03	4,70
Dep. Vista/Depósitos	9,01	0,97	3,68	7,02	2,70	10,31	7,07
Fin. Externo/Pasivos	0,11	1,24	0,64	2,03	1,36	2,88	1,87
Total	100,00	100,00	100,00	100,00	100,00	100,00	100,00

Fuente: Cálculo de los autores a partir de información entre enero del 2008 y mayo del 2013.

a. Porcentaje en que cada indicador explica distancia absoluta de cada grupo de bancos.

b. La intensidad del color verde muestra la importancia del indicador en explicar la porción de la distancia en cada grupo por separado.

¿Son estos grupos estables en el tiempo? El cuadro 7 permite analizar la estabilidad de cada grupo de bancos a partir del porcentaje de veces que una institución bancaria pertenece al mismo grupo en el período 2008-2013. Como se puede observar, si bien algunos bancos migran de un grupo a otro, estos se ubican en la frontera de los mismos. En términos generales, los grupos definidos son estables en el tiempo.

¿Qué tan similares son los bancos de cada grupo? El cuadro 8 muestra las diferencias al interior de cada grupo, a partir de la distancia euclidiana del banco en relación con la media de los otros bancos del mismo grupo. Este indicador tiene un mínimo de cero cuando todos los integrantes del grupo son idénticos. Como se aprecia a partir del cuadro 8, los bancos comerciales son más similares entre sí que el grupo de tesorería. Asimismo, se puede observar qué bancos son más distintos entre sí, dentro de cada grupo.

La evolución en el tiempo de estas distancias nos da información acerca de cuáles son los períodos en los que los bancos son más similares en estructura. En este sentido, el gráfico 5 muestra una ligera tendencia a tener grupos de bancos más compactos hacia el 2013, en particular para la banca comercial. En tanto, la banca de tesorería ha tendido a diferenciarse en los últimos años.

Cuadro 7**Ubicación en la agrupación secundaria de bancos^a**
(porcentaje del período total)

	Multibanco grande	Multibanco medio	Especial media	Personas	Tesorería	Comercio exterior	Servicios financieros
Chile	100						
BCI	100						
Estado	100						
Santander	100						
Scotiabank		100					
BBVA		92	8				
Itaú		32	68				
Corpbanca		29	71				
Bice			100				
Internacional			100				
Security			100				
Falabella				100			
Paris				100			
Ripley				100			
Deutsche					100		
JP Morgan					92		8
BNA					15	68	17
DoBrasil						100	
Tokyo					2	58	40
HSBC					18	23	58
Consortio							100
Penta							100
Rabobank							100

Fuente: Cálculo de los autores a partir de información de marzo del 2013.

a. Porcentaje de meses en que la institución permanece en el grupo específico.

**Cuadro 8****Distancia promedio entre bancos de un mismo *cluster*^a**

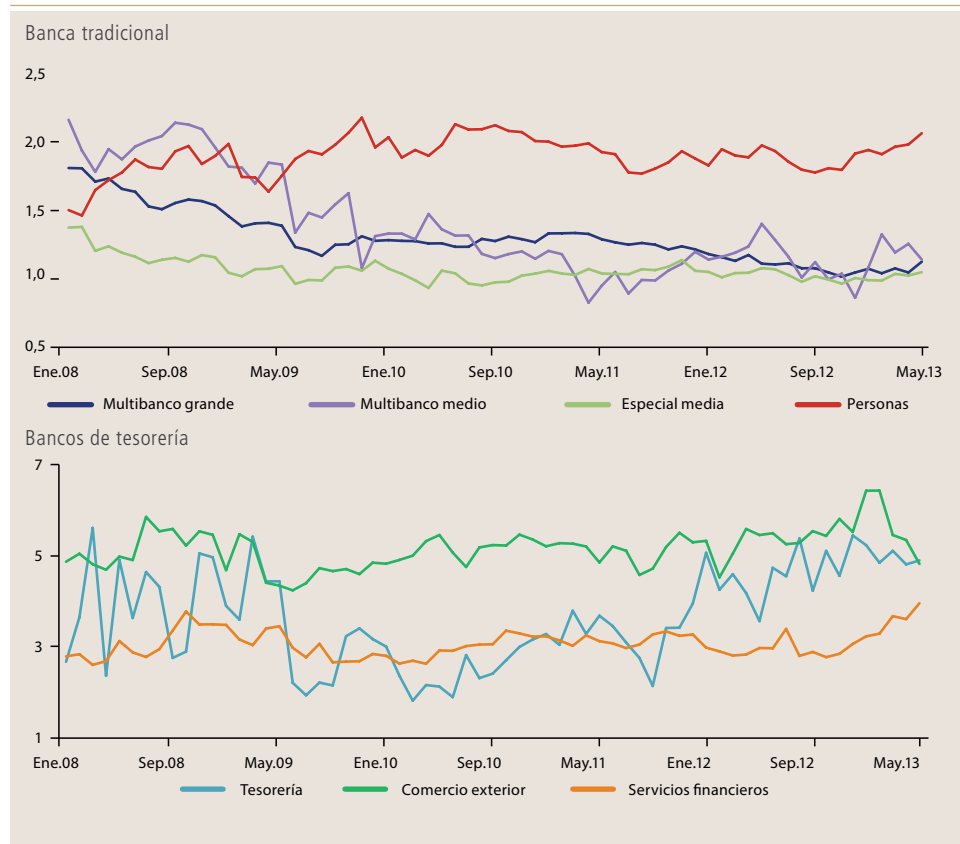
<i>Cluster</i>	Banco	Distancia	<i>Cluster</i>	Banco	Distancia
Multibanco grande	Chile	0,778	Servicios financieros	HSBC	2,661
	Santander	0,996		Consortio	2,663
	BCI	1,116		Penta	3,229
	Estado	1,397		Rabobank	6,035
Multibanco medio	BBVA	1,187	Tesorería	JP Morgan	5,068
	Scotiabank	1,187		Deutsche	5,068
	Ripley	1,020		BNA	4,864
Personas	Falabella	2,308	Comercio exterior	Tokyo	5,374
	Paris	2,538		DoBrasil	5,997
	Security	0,482			
	Bice	0,631			
Especializada mediana	Corpbanca	1,179			
	Itaú	1,233			
	Internacional	1,633			

Fuente: Cálculo de los autores a partir de información de marzo del 2013.

a. Distancia absoluta de cada banco respecto de la distancia promedio de cada *cluster*.

Gráfico 5

Distancia promedio por tipo de cluster^a



Fuente: Cálculo de los autores.

a. Corresponde a la distancia euclidiana promedio por cluster.

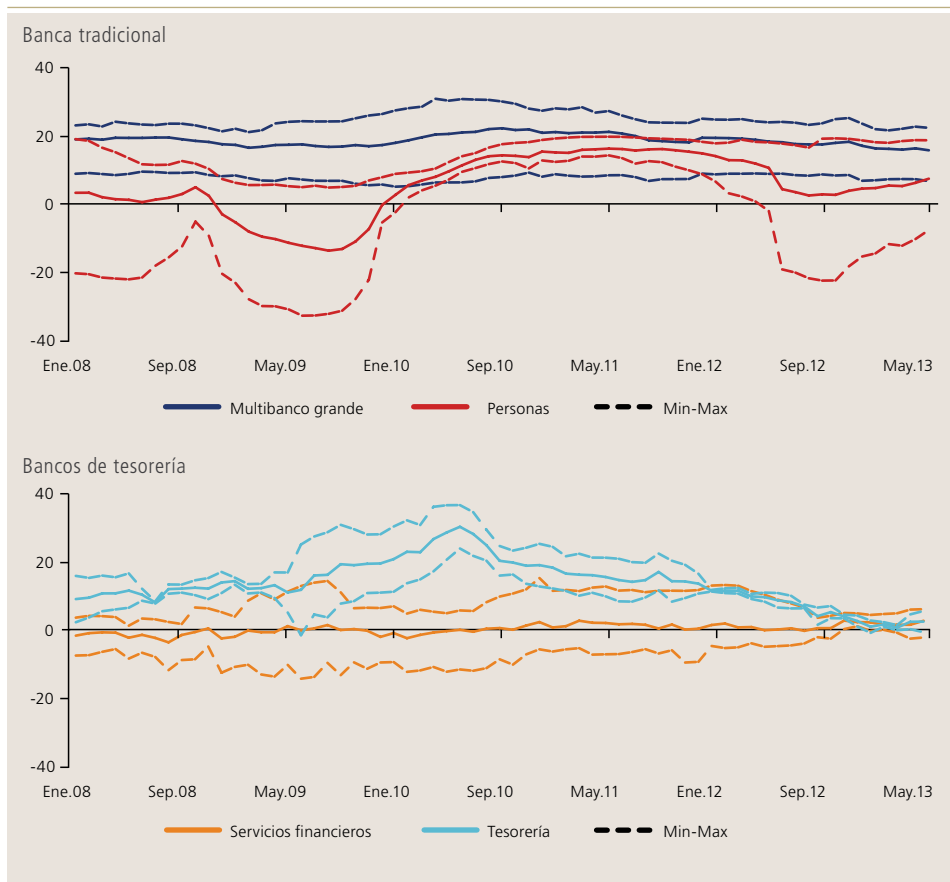
¿Qué tan relevante es la agrupación sugerida para eliminar el sesgo que se genera al analizar las variables y los indicadores en términos agregados? Para responder a esta pregunta se analiza la evolución de la rentabilidad sobre el capital (ROE) para cada grupo en forma separada. Tal como muestra el gráfico 6, el ROE de los multibancos grandes es más estable en el tiempo que el ROE de la banca de personas. Por otra parte, los bancos de tesorería, cuya exposición es mayor a los riesgos de mercado (ej.: tipo de cambio y tasas de interés) muestra una diferencia sustancial, tanto en nivel como en variabilidad, a la rentabilidad de la banca de servicios financieros, la que al estar mayormente dedicada a los créditos comerciales, presenta un ROE más estable¹⁴.

14 Nótese que el rango de ROE (mínimo y máximo) de los grupos no se cruza en la mayor parte del tiempo y se mueve en la forma similar a su media.



Gráfico 6

Rentabilidad del capital por tipo de *cluster*^{a,b}



Fuente: Cálculo de los autores.

a. Equivale al promedio de rentabilidad sobre patrimonio por *cluster* seleccionado.

b. Líneas segmentadas representan el mínimo y el máximo de cada grupo.

En consecuencia, los grupos sugeridos caracterizan la heterogeneidad del sistema bancario y reducen el sesgo que se genera al considerar las variables en términos agregados. Si bien este análisis no es concluyente, este tipo de agrupación aporta información al análisis y monitoreo de la industria bancaria. Asimismo, es una herramienta práctica para establecer bases de comparación para los bancos respecto de sus pares.

IV. CONCLUSIONES

Esta nota presenta una herramienta que permite distinguir aglomeraciones de bancos a partir de un análisis multidimensional de indicadores de la banca en Chile. El set de indicadores considerados incorpora básicamente información del tamaño y de la estructura de activos y de pasivos bancarios. De esta forma, se busca agrupar los bancos de acuerdo con el grado de exposición a riesgos comunes. Haciendo uso de información de balance para el período 2008-2013, se aplica el método jerárquico del análisis de *cluster* a 23 bancos en Chile, encontrando que la industria bancaria puede ser agrupada en siete grupos de instituciones bancarias homogéneas: (1) los multibancos grandes, (2) los multibancos medianos, (3) la banca especializada mediana; (4) la banca de personas, (5) los bancos de tesorería, (6) los bancos de comercio exterior y (7) la banca dedicada a los servicios financieros.

Los grupos formados, a partir del análisis de conglomerados, presentan alta estabilidad durante el período de estudio. Sin embargo, la evolución del sector, eventuales cambios normativos y los cambios en las estrategias de negocio que suelen implementar los bancos, podrían alterar la composición de cada grupo.

Finalmente, esta primera aproximación metodológica a la segmentación de la banca chilena sienta las bases para futuras clasificaciones *ad hoc*, utilizando un conjunto de indicadores distintos, dependiendo del objetivo que persiga la agrupación de los bancos



REFERENCIAS

Ahumada, L.A. y J. Marshall (2001). "The Banking Industry in Chile: Competition, Consolidation and Systemic Stability". BIS background paper.

Amel, D.F. y S.A. Rhoades (1988). "Strategic Groups in Banking". *Review of Economics and Statistics* 70(4): 685–9.

Banco Central de Chile. *Informe de Estabilidad Financiera*, varios números.

Calinski, T. y J. Harabasz (1974). "A dendrite method for cluster analysis". *Communications in Statistics* 3: 1–27.

Dardac N. y Boitan A. I., 2009. "A Cluster Analysis Approach for Bank's Risk Profile: The Romanian Evidence". *European Research Studies Journal, European Research Studies Journal, vol. 0(1)*, pages 109-118.

Duda, R. O.; P. E. Hart y D. G. Stork (2001). "Pattern Classification". 2nd ed. New York: Wiley.

Jara, A. y D. Oda (2007). "Análisis de *Cluster* de la Banca en Chile". Mimeo, Banco Central de Chile.

Kaufman, L. y P.J. Rousseeuw (1990). *Finding Groups in Data*. Nueva York, EE.UU.: John Wiley & Sons.

Koller, W. (2001). "Strategic Groups in Austrian Banking 1995-2000". Research Institute for European Affairs (IEF). Mimeo, Universidad de Economía y Negocios de Viena, Austria.

Mao, J. y A.K. Jain (1996). "A Self-organizing Network for Hyperellipsoidal Clustering (HEC)". *IEEE Transactions Neural Networks*. 7(1): 16–29.

Milligan, G.W. y M.C. Cooper (1985). "An Examination of Procedures for Determining the Number of Clusters in a Data Set". *Psychometrika* 50:159.

Terrones, C. y P. Vargas (2013). "Clasificación de la Banca Comercial Peruana: Un Análisis de *Cluster* Jerárquico". Documento de Trabajo, Superintendencia de Banca, Seguros y AFP del Perú.

Ward, J.H., Jr. (1963). "Hierarchical Grouping to Optimize an Objective Function". *Journal of the American Statistical Association* 58: 236–44.

Wigodski, T. y C. Torres (2004). "Análisis Estratégico de la Industria Bancaria Chilena". *Trend Management* 7: 10–7.