IV SPECIAL FEATURES

A IDENTIFYING LARGE AND COMPLEX BANKING GROUPS FOR FINANCIAL SYSTEM STABILITY ASSESSMENT

For the purposes of financial system stability assessment, it is important to identify and monitor the activities of banking groups whose size and nature of business is such that their failure and inability to operate would most likely have adverse implications for financial intermediation, the smooth functioning of financial markets or other financial institutions operating within the system. A simple and common approach for identifying such institutions – often grouped under the heading large and complex banking groups (LCBGs) – is to rank them by the size of their balance sheets. However, asset size alone may fail to shed much light on the importance and complexities of the interconnections that a banking group may have within a financial system, especially given the growing importance of banks' off-balance sheet activities. Knowledge about such interconnections is important because it can help in mapping how, or if, strains in a large banking group could spread to other institutions or markets. Based on a multiple indicator approach, this Special Feature takes a first step towards statistically identifying banking institutions that meet certain "largeness" characteristics that go beyond balance sheet size.

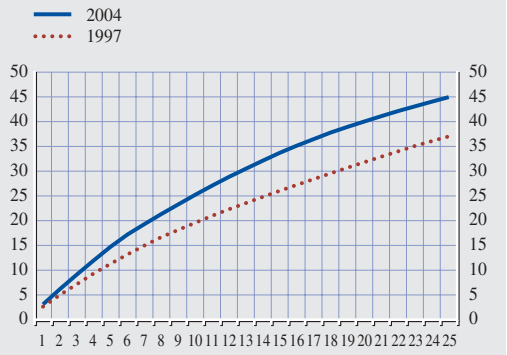
INTRODUCTION

Thanks to mergers, acquisitions and organic growth, a relatively small number of banking groups in the euro area now control a significant and growing share of total banking sector assets (see Chart A.1). Because of their importance from a financial system stability perspective, this Special Feature considers a framework for identifying LCBGs. The "largeness" of a banking group clearly depends on the size of its balance sheet. Indeed, the most rudimentary method for identifying large banking groups is to rank institutions by their total assets. In practical terms, however, this approach has at least two shortcomings for financial system stability assessment. First, there is no commonly agreed threshold for the percentage of banking sector assets, or the number of large banking groups, that should be monitored. Second, given the growing importance of off-balance sheet activity, the size of a financial institution's balance sheet may not necessarily accurately reflect its complexity or the importance of the role it plays in the various forms of financial intermediation, risk transformation and management processes that take place within the financial system.[1] For instance, should a banking group fail that is relatively large, but which has few linkages with other parts of the financial system, it may have little impact on the functioning of other financial institutions. By contrast, a smaller bank with few but important linkages could have a disproportionately larger adverse impact on the functioning of financial markets or other financial institutions. This could occur, for instance, in the case of a bank offering brokerage services in derivatives markets, or acting as a custodian in security settlements. Neither of these activities is necessarily associated with a bank's portfolio size.

To assess how important a banking group is for the smooth functioning of the various

Chart A.1 Cumulative banking sector asset shares of the 25 largest euro area banking groups

(% of total euro area banking sector assets)



Sources: Bureau van Dijk (Bankscope) and ECB calculations.

1 See, for example, I. W. Marsh and I. Stevens (2003), "Large Complex Financial Institutions: Common Influences on Asset Price Behaviour", *Bank of England Financial Stability Review*, December.

intermediation, risk transformation and management processes that take place within the system, a wide set of key business activity characteristics is needed. Clearly, the wider the set of activities that are considered to be important for the stable functioning of the financial system, the more complex the conceptual and technical challenges to ranking – or even selecting – large banking groups become. In addition to technical difficulties related to making league table comparisons (i.e. lists of banking groups ordered by the relative size of the different indicators), or the fact that indicators of different banking activities will inevitably be measured using different scales, two essential problems arise. First, without a weighting of the importance of different activities that take place within a financial system, there is no natural prescription for aggregating what might be conflicting rankings for the same institutions across league tables of different indicators. Second, league table rankings offer no information on magnitude, as the ordering obscures the measure of the "largeness" of a banking group in a particular banking activity.

This Special Feature explores a methodology for data analysis that aims at addressing these issues. While the procedure is simple, transparent and lends itself to dealing with expanding information sets, it is by no means unique or even necessarily the best among all available procedures. What it does represent, however, is a clear improvement on simply selecting banks on the basis of asset size and arbitrarily choosing a threshold asset value or number of institutions, and in this regard should be seen as a first step.

The remainder of this Special Feature is organised as follows: the following section discusses banks' presence and linkages. This is followed by a section describing the data and data processing and a section describing the methodology and the main findings. Conclusions and implications for financial stability monitoring are provided in the final section.

BANKS' PRESENCE AND LINKAGES

Large and complex banking groups can be considered as institutions whose size and nature of business is such that their failure and inability to operate would most likely spread and have adverse implications for the smooth functioning of financial markets or other financial institutions operating within the system. If the disturbance were large enough to threaten financial system stability it could be transmitted through various channels – including payment systems and markets – but would most likely originate from an institution being unable to meet its payment and settlement obligations.[2]

With a view to selecting suitable business activity variables for identifying LCBGs, conceptual work on systemic risk can be helpful when it comes to pinpointing potential contagion channels though which adverse disturbances could be transmitted throughout the financial system. The literature distinguishes between contagion channels that are "pure" (resulting from either idiosyncratic or systemic shocks), those that are information-based (stemming from information asymmetries among investors and/or depositors), and those that are a combination of both.[3] Of these two potential channels, case studies of systemic banking crises have not found information-based channels to be important, and in any case they pose significant challenges in terms of monitoring.[4] By contrast, pure contagion channels are more amenable to surveillance as they are based on measurable quantities. Two types of pure shocks to a banking system can be distinguished: systemic and idiosyncratic. At the core of financial stability monitoring are systemic (common, and often macroeconomic) shocks that affect all banks in the system simultaneously. A common finding in the

2 For an overview of systemic banking crises since the late 1970s, see G. Caprio and D. Klingebiel (2003), "Episodes of Systemic and Borderline Financial Crises", World Bank.

3 For a review of this literature, see O. De Bandt and P. Hartmann (2000), "Systemic Risk: A Survey", *ECB Working Paper,* No 35.

4 See C. W. Calomiris and J. R. Mason (1997), "Contagion and Bank Failures during the Great Depression: The June 1932 Chicago Banking Panic", *American Economic Review,* Vol. 87.

empirical literature is that the level of banks' exposure to systemic shocks tends to determine the extent and severity of a systemic crisis. However, an individual bank can – through failure or inability to operate – also be a source of systemic risk. The transmission channel of the idiosyncratic shock can be direct – for example if the bank was to default on its interbank liabilities – or indirect, whereby a bank's default leads to serious liquidity problems in one or more financial markets where it was involved.

The degree to which individual banking groups are "large" in the sense that this could be a source of systemic risk would therefore seem to depend on the extent to which they can be a conduit for diffusing systemic and idiosyncratic shocks through a banking system. Attempts at estimating the degree of interconnectedness of banking groups can be divided into two main strands: one which measures the degree of co-movement of indicators based on security prices, and the other which is based on simulation exercises using interbank lending data.[5] Both approaches have some shortcomings. Clearly the first approach can only be followed for banking groups that are listed on stock exchanges. Moreover, using co-movement measures to make inferences about probable behaviour in times of distress often offers only limited insight into the nature of the relationship.[6] Simulation exercises, on the other hand, are less than ideal as they ignore remaining shock transmission channels and only provide a "lower bound" of the potential degree of spillover from one banking group to another following the crystallisation of a shock.[7] A shortcoming common to both types of studies is that they lack criteria for selecting which indicators of banks' interconnectedness are useful for identifying relevant banking groups from a financial system stability assessment viewpoint. Typically, total assets – or a combination of balance sheet items – serve as *a priori* criteria for sample pre-selection.[8]

The methodologies explored in this study aim at both expanding the set of possible indicators as well as quantifying the degree to which a bank is interconnected with the rest of the banking system, so as to determine endogenously a bank's "size" in the financial system.

DATA DESCRIPTION AND PROCESSING

As the purpose of this study is to identify those banking groups active in the euro area that play important roles in various forms of financial intermediation, risk transformation and management processes, the key business activity characteristics of around 260 banks, both domiciled within and outside the euro area, were examined. The analysis was restricted by the availability and comparability of publicly available information.[9] Hence, it does not include off-balance sheet positions, even though these are often important sources of interconnectedness. The business characteristics of banks were pre-screened and they were included in the analysis if they met one or more of the following three criteria in 2005:

– domiciled in Europe and with total assets in excess of one billion euro; or

– included in the top 30 bookrunners in the European equity, bond and syndicated lending markets; or

5 An approach using both methodologies is taken by H. Elsinger, A. Lehar and M. Summer (2006), "Using Market Information for Banking System Risk Assessment", *International Journal of Central Banking,* March. For an application based on payment systems data, see E. Amundsen and H. Arnt (2005), "Contagion Risk in the Danish Interbank Market", *Danmarks Nationalbank Working Papers,* No 29.

6 See for example G. De Nicolo and M. L. Kwast (2002), "Systemic Risk and Financial Consolidation: Are They Related?", *Journal of Banking and Finance,* Vol. 26, No 5, May.

7 Furfine provides a seminal study of interbank positions determining banks' systemic relevance (C. H. Furfine (2003), "Interbank Exposures: Quantifying the Risk of Contagion", *Journal of Money, Credit and Banking,* Vol. 35, No 1).

8 See for example P. Hartmann, S. Straetmans and C. G. de Vries (2005), "Banking System Stability: A Cross-Atlantic Perspective", *NBER Working Papers,* No. 11698. Furfine (2003) uses interbank federal funds exposures, while Elsinger et al. (2004) focus on total interbank positions (assets and liabilities) on banks' balance sheets. A notable exception is Marsh and Stevens (2003), who also recognise the importance of proxies of off-balance sheet items such as foreign exchange trading revenues or assets held in custody.

9 The data sources included Bureau van Dijk's Bankscope, Thomson Financial's Thomson ONE Banker – Deals and GlobalCustody.net.

– among the top 48 worldwide custodian banks according to Global Custody.

All banking groups domiciled in Europe – and not just euro area banking groups – that met at least one of the criteria above were included in the analysis, in order to permit the identification of large banking groups domiciled outside the euro area which could be seen as being important for euro area financial stability assessment. Similarly, banking groups domiciled outside Europe were included in the analysis if they were among the top 30 bookrunners in the European equity, bond or syndicated lending markets or among the top 48 worldwide custodian banks.

For the purpose of financial system stability assessment, the banking activity indicators selected for identifying LCBGs should ideally encompass relevant dimensions of importance with regard to various aspects of financial intermediation, as well as the degree of interconnectedness of the institution within the system. In this respect, the scale of a given banking group's activities in different banking market segments, in interbank markets and of its total assets are essential dimensions of its size. For instance, if a banking group has a particularly large share of the residential mortgage market, then the smooth functioning of that market segment may depend on the financial condition of the intermediary concerned, which would call for it to be included in the set of LCBGs. Other variables for selecting LCBGs might include a bank's activity level (gauged, for instance, by revenues – which might not be correlated with entries on the balance sheet, but which may be an important measure of size), as well as the institution's role and importance as a bookrunner in the issuance of equities, bonds and syndicated loans and its role as a custodian bank. The indicators used in the analysis conducted for this Special Feature can be grouped as follows:

– Traditional banking balance sheet items: loans, mortgages, other earning assets, deposits and contingent liabilities;

Table A.1 Correlation of indicators with total assets

(2005, cross-sectional correlations)

Assets under custody	0.24
Contingent liabilities	0.36
Interbank assets	0.45
Interbank liabilities	0.49
Net interest revenue	0.50
Proceed amount from equity issuance	0.51
Deposits	0.54
Customer loans	0.57
Net non-interest revenue	0.68
Proceed amount from syndicated loan issuance	0.68
Other assets	0.70
Proceed amount from bond issuance	0.79
Mortgages	0.84

Sources: Bureau van Dijk (Bankscope), Thomson ONE Banker-Deals, GlobalCustody.net and ECB calculations.

– Traditional indicators of banking activity: net interest revenue and net non-interest revenue;

– Interbank assets and liabilities;

– Bookrunner role: proceed amount in European equity, bond and syndicated loans markets; and

– Custodian role: worldwide assets under custody.

As previously mentioned, these indicators are used because a banking group's total assets may not necessarily provide an indication of the institution's complexity or of the importance of the role it plays in various forms of financial intermediation, risk transformation and management processes. Indeed, many of the indicators used in the analysis display rather low correlations with total assets (see Table A.1).

For the non-euro area banking groups, the indicators analysed were re-scaled to approximate the share of their business that was carried out in the euro area. The variables were scaled down to 50% if the banking group is domiciled in a non-euro area EU country, 40% if in a non-EU European country, and 10% if located outside Europe. While this scaling is to some extent rather arbitrary, some form of scaling is nevertheless needed to approximate

the share of the business conducted by non-euro area banking groups in the euro area, since the purpose is to identify large banking groups which are important from a euro area perspective. It should, however, be borne in mind that the scaling does not affect the identification of LCBGs that are domiciled in the euro area. Proceed amounts from bookrunner activity and worldwide assets under custody were not scaled down since these markets are generally internationally integrated.

A common approach to dealing with the aggregation of quantitative information measured in different units is to construct so-called league tables. The advantage of this approach is that rankings are measure-independent and therefore comparable.[10] However, league table rankings do not take magnitude into account. For instance, a bank that is ranked second in a league table may indeed have an absolute indicator value that is almost as large as the one ranked above it; however, this value could also be considerably smaller.

In order to utilise information on the magnitude of indicators, when both a variable metric and a cross-variable comparison are needed, it is necessary to standardise the variables in such a way that the relative sizes are preserved and are, at the same time, independent of the measurement unit. In this Special Feature, indicators are divided by the indicator's largest value (i.e. the value of the indicator for the banking group with the highest value of the indicator). Accordingly, all standardised indicators range between zero and one, and the relative distances between banking groups' indicator values are retained.

METHODOLOGY AND FINDINGS

THRESHOLD SELECTION BY MEANS OF CLUSTER ANALYSIS

One relatively simple way to address the threshold selection issue for either the percentage of banking sector assets or the number of large banking groups that need to be monitored when assessing the stability of the financial system is to conduct cluster analysis. This is a statistical method that separates a sample population into natural groups according to measures that define the characteristics of the population (see Box A.1). The business activities of banking groups are natural dimensions for grouping banks into clusters. When assessing a bank's size and importance, three broad types of banks are distinguished:

1. Banks with low values across all characteristic indicators – generally small banks.

2. Banks with one or a few medium-sized indicator values – generally medium-sized banks.

3. Banks with one or several high indicator values – large banks.

10 Marsh and Stevens (2003), op. cit., for instance, select their sample on the basis of those banks that appear in more than a threshold number of league table rankings.

Box A.1

CLUSTERING METHODS

The term "cluster analysis" encompasses a number of different algorithms and methods for grouping similar objects into respective categories. In other words, it is an exploratory data analysis tool which aims at sorting different objects into groups in such a way that the degree of association between two objects is maximal if they belong to the same group and minimal otherwise. The approach simply discovers structures in data without explaining why they exist.

Because it lacks the formal distribution models required for statistical analysis, it is typically used as a complement to other data analysis methods.[1]

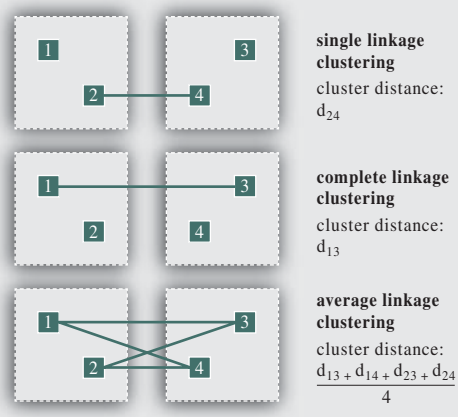
There are two basic clustering algorithm techniques: partitioning and hierarchical. With partitioning techniques a certain number of final clusters have to be assumed in advance. This is not the case with hierarchical techniques, where a series of successive merges or divisions are produced. Because the number of clusters is a priori unknown in this study, a hierarchical technique is used.

The hierarchical techniques are further divided into two main methods: agglomerative and divisive. With divisive methods an initial single group of objects is divided into two subgroups that are as dissimilar as possible. These subgroups are then further divided into dissimilar subgroups. Agglomerative methods, on the other hand, start with individual objects, so that there are initially as many clusters as objects. Objects are then grouped and merged according to their similarities until all objects have been grouped. This Special Feature uses an agglomerative method, which is also more commonly used and widely implemented in software solutions.

There are three main agglomerative hierarchical clustering methods: single linkage, complete linkage and average linkage. They differ in the way that they measure the Euclidean distance – that is, the geometric distance in a multidimensional space – between the clusters. In the single linkage method, the distance between two clusters is determined by the distance of the two closest objects (nearest neighbours) in the different clusters (see Figure B.A.1). This approach effectively strings objects together to form clusters that resemble long chains. This "chaining" can however be misleading if items at opposite ends of the chain are, in fact, quite dissimilar. The advantage of the complete linkage method is that the distances between clusters are determined by the greatest distance between any two objects in the different clusters (i.e. by the so-called furthest neighbours). This method usually performs quite well in cases when the objects actually form naturally distinct clumps, but is inappropriate if the clusters tend to be somewhat elongated or of a chain-type nature.

This Special Feature uses the so-called average linkage method, which combines the single and complete clustering methods by measuring the average distance between clusters as the average distance between all objects in the different clusters. This method reduces some of the problems encountered when using the single and complete linkage methods on their own, and can therefore be seen as a compromise solution.

Figure B.A.1 Intercluster distance



1 See, for example, R. A. Johnson and D. W. Wichern (1998), *Applied Multivariate Statistical Analysis,* Upper Saddle River: Prentice-Hall; B. S. Everitt (1993), *Cluster Analysis*, 3rd edition, London: Arnold; B. S. Everitt and T. Hothorn (2006), *A Handbook of Statistical Analyses Using R,* Boca Raton: Chapman & Hall/CRC; and W. N. Venables and B. D. Ripley (2002), *Modern Applied Statistics with S,* New York: Springer-Verlag.

A typical hierarchical cluster analysis procedure starts out by considering each of n banks as a separate group in a p-dimensional space, where p is the number of relevant characteristics. For example, if the only measure desired is the size of total assets, then p=1. The natural distance between banks in this p-dimensional space is the Euclidian distance. Therefore, the Euclidean distance gives a measure of the banks' (dis)similarity – as more similar banks are characterised by shorter distances. An iterative procedure is then run to group the banks hierarchically in terms of the distance between them. For example, banks with the lowest Euclidean distance are combined into one group, resulting in n-1 groups after the first step, and so forth until only one group is left. Looking at the ranking of the distances between groups then allows distinct jumps in the grouping process to be selected using a pre-specified criterion, creating a natural separation between groups. These distinct jumps can then be analysed using so-called stopping-rules to determine the number of groups which, statistically, represents the most significant division of a population's sample.

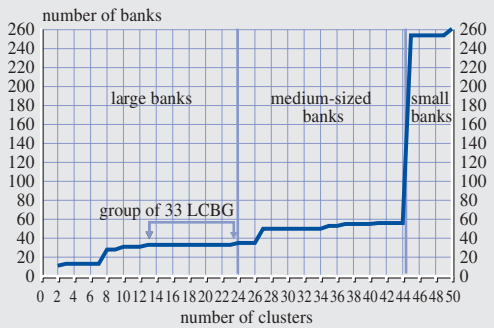
The cluster analysis applied to data for 2005 categorises the 260 banks into 50 different hierarchical clusters (see Chart A.2). The cluster groups permit a relatively clear demarcation of the line distinguishing large banks from the rest by looking at 13 to 23 clusters which contain a stable set of 33 banking groups. This is the most stable set identified during the clustering procedure, and all banks in this set share the characteristics of being both large and complex (i.e. they are important players in a range of banking activities).

By looking at 24 hierarchical clusters, the set of banking groups is extended to 35 by adding two banks that have a fairly high indicator value in only one of the three bookrunner markets considered; these are not considered to be LCBGs in this Special Feature. If the number of clusters is extended to 27, the set of banks increases to 50, adding 15 banks with medium-size balance sheets but that have no role as bookrunners or custodian banks, and are therefore also not deemed to be LCBGs.

Among the 33 banking groups identified as large in this analysis, 20 are located in euro area countries, seven in the US, four in the UK and two in Switzerland.

Chart A.2 Number of banks per number of clusters

(2005)



Sources: Bureau van Dijk (Bankscope), Thomson ONE Banker – Deals, GlobalCustody.net and ECB calculations.

ROBUSTNESS OF THE CLUSTER ANALYSIS

The robustness of the analysis to the indicators used to identify LCBGs needs to be analysed further in light of the need for possibly more expanded coverage of indicators over time. Periodic financial system stability assessment requires a relatively stable set of institutions to be monitored so as to ensure continuity. However, as the euro area banking landscape is likely to be transformed over time, primarily – but not only – through the consolidation of the banking sectors in EMU participating states, some changes in the way that large banking groups are identified are to be expected. As the sources of risk and vulnerability for financial system stability can change over time thanks to financial innovation, along with the development of new markets and the changing strategic emphasis given by banks to different activities,

Table A.2 Large and complex banking groups excluded when excluding indicators

	variable excluded from the cluster analysis						
	balance sheet items					activity	
	customer loans	mortgages	other earning assets	deposits	contingent liabilities	net interest income	net non-interest income
Number of banks excluded	1	3	0	0	0	3	0

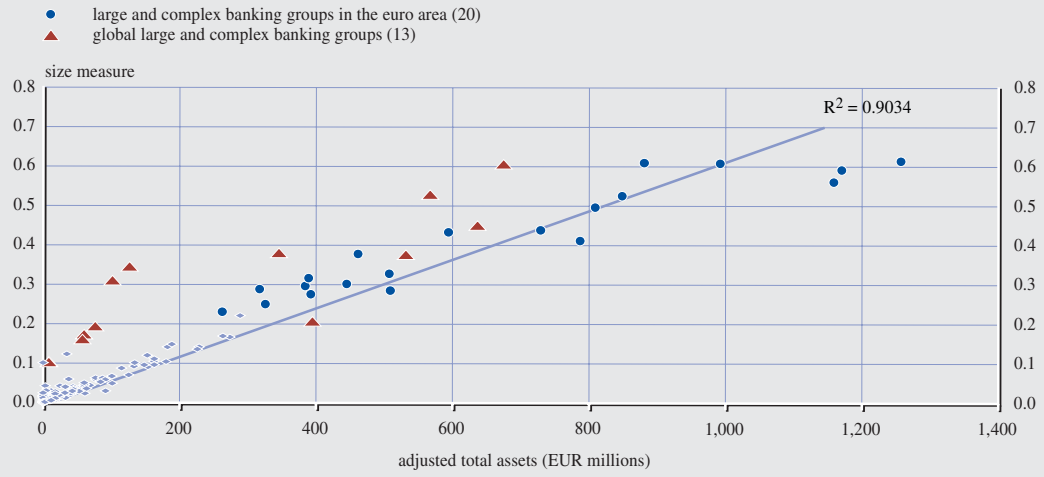
	variable excluded from the cluster analysis					
	interbank positions		bookrunner activity			custodian role
	interbank assets	interbank liabilities	bond issuance	equity issuance	loan issuance	custody assets
Number of banks excluded	0	9	0	3	0	2

it is also to be expected that the set of indicators will change too. Nevertheless, adding or removing indicators judged relevant would ideally only marginally affect the composition of the group, which would make this a good measure of robustness.

Adding or removing indicators from the cluster analysis can also shed light on the dependence of size on the available indicators. Applying the clustering analysis to the sample of banks after dropping a given indicator also suggests that the procedure is fairly robust to variation in the availability of variables. It turns out that seven of the 13 variables do not affect the "largeness" demarcation (see Table A.2). Often only one to three banking groups are excluded from the selection when excluding one variable, with the exception of the omission of the "interbank liabilities" variable, which excludes eight euro area banks and one non-euro area bank.

Chart A.3 Size measure vs. adjusted total assets

(2005)



Sources: Bureau van Dijk (Bankscope), Thomson ONE Banker – Deals, GlobalCustody.net and ECB calculations.
Note: The size measure ranges between 0 and 1 and is the average of the 13 balance sheet, activity, interbank positions, bookrunner and custodian indicators used in the cluster analysis. The adjusted total assets is total assets scaled down to 50% if the bank is domiciled in a non-euro area EU country, 40% if in a non-EU European country, and 10% if located outside Europe.

A COMPARISON WITH SOME MEASURES OF SIZE

To shed some light on the characteristics of the 33 large banks identified above, it is useful to correlate total assets – the traditional variable employed for selecting large banks – with some other measures of importance. The relationship between an adjusted total assets measure and a composite size measure, based on the 13 indicators used in the cluster analysis, is tight but imperfect (see Chart A.3). The fact that the correlation is less than one suggests that this methodology adds value over and above a selection based simply on total assets. The reason why the correlation is imperfect is because some banking groups with relatively low levels of total assets have other characteristics that make them important for the financial system.

CONCLUDING REMARKS

Despite the fact that cluster analysis is best characterised as an explanatory data analysis technique, it can provide a robust identification of LCBGs for periodic analysis of financial system stability. It should be emphasised that such banking groups are not necessarily those that are often called "systemically relevant institutions". Rather, they are banking groups whose size and nature of business is such that their failure and inability to operate would most likely have adverse – albeit not necessarily severe – implications for various forms of financial intermediation, the smooth functioning of financial markets or other financial institutions operating within the system. Judging their systemic relevance would however require far more information, especially on off-balance sheet positions. In that respect, supervisory knowledge and information can further enhance the assessment of an institution's importance for financial system stability.

As shown, the importance of a banking group in the financial system can go beyond traditional measures of size: the role it plays in specific banking activities, or the interconnections it has with other parts of the financial system, are also important considerations that need to be taken into account. Some of this information is publicly available, but there are important gaps in information – for example on the off-balance sheet positions of banks, the degree of their participation in relatively new financial markets (e.g. structured finance, traditional credit issuance, etc.), or on cross-border activities – that leave room for further refinement of the filtering procedure. The variables used in this study represent natural choices given the paucity of publicly available data for a large number of banks. The methodology would benefit greatly from expanding and refining the set of variables used. Fortunately, a very positive feature of the methodology used is that it can easily accommodate a growing number of indicators and banking groups, thus allowing for future enhancements in the availability of quantitative information. The set of banking groups identified is therefore likely to change over time.