



# Too big to fail? An analysis of the Colombian banking system through compositional data

Juan David Vega Baquero<sup>a</sup>, Miguel Santolino<sup>b,\*</sup>

<sup>a</sup> University of Barcelona Diagonal, Barcelona 690 08034, Spain

<sup>b</sup> Department of Econometrics Riskcenter-IREA, University of Barcelona Diagonal, Barcelona 690 08034, Spain



## ARTICLE INFO

### JEL classification:

G17  
G21  
G28

### Keywords:

Simplex  
Aitchison geometry  
Systemically important banks  
Vector autoregression

## ABSTRACT

Although still incipient in economics and finance, compositional data analysis (in which relative information is more important than absolute values are) has become more relevant in statistical analysis in recent years. This article constructs a concentration index for financial/banking systems via compositional analysis to establish the potential existence of “too big to fail” financial entities. The intention is to provide an early warning tool for regulators about this type of institution, so they can define thresholds and measures depending on their risk appetite and the systems’ specificities. The index has been applied to the Colombian banking system and assessed over time with a forecast to determine whether the system is becoming more concentrated. Results found that the concentration index has been decreasing in recent years and the model predicts this trend will continue. Regarding the methodology used, compositional models were shown to be more stable and to lead to better prediction of the index compared to the classical multivariate methodologies.

## 1. Introduction

The term “too big to fail” (TBTF) has been in use since the 1980s for institutions that can pose “significant risks to other financial institutions, to the financial system as a whole, and possibly to the economic and social order” (Stern and Feldman, 2004, p. 1). Nevertheless, some authors did not consider that the TBTF concept should be central to banking regulation (Mishkin et al., 2006) until the financial crisis in 2008, when evidence emerged of how the effect of shocks in these institutions can expand without control. An anecdotal explanation on the development of the crisis can be found in Sorkin (2010). Some of the initial literature on TBTF institutions after the crisis include Shull (2010) and Zhou (2010). More recently, authors such as Barth and Wihlborg (2017), Ioannou et al. (2019), Omarova (2019) and Cetorelli and Traina (2021) have continued the discussion on TBTF institutions, focusing on their development after the crisis.

As a response to the crisis from the regulatory side, the Basel Committee on Banking Supervision (BCBS) started working on a new framework to reduce the risk of such a disruption in the future. This led to new definitions, especially those of global systemically important banks (G-SIBs) and domestic systemically important banks (D-SIBs).<sup>1</sup> In the case of the D-SIBs, the definition includes not only the size, but also other characteristics such as interconnectedness, substitutability and complexity of the financial institutions, while for G-SIBs, the cross-national activity should also be considered. Both definitions are complementary, depending on the nature of each entity. The BCBS also mentions the importance of considering the potential impact of a D-SIB at an international level, either

\* Corresponding author.

E-mail address: [msantolino@ub.edu](mailto:msantolino@ub.edu) (M. Santolino).

<sup>1</sup> Basel Committee on Banking Supervision (2022)

regional or bilateral, even if the effect at a global level might be insignificant. Moch (2018) carried out an interesting review of recent literature about SIBs. Among the most important conclusions, the author noted that there was consensus that the relationship between the size of banks and systemic risk was nonlinear. In other words, the risk increases more than proportionally to increases in the size of banks. Furthermore, the author found that concentration and interconnectedness in systems could lead to an amplified effect of size on systemic risk. Li et al. (2020) developed a recent attempt to measure the effect of SIBs on systemic risk. They tried to determine the importance of an institution in the system via changes in the systemic risk measured with and without it (a leave-one-out method), using z-scores as the risk measure. Another approach can be found in Bezrodna et al. (2019), who proposed a risk indicator that considered the importance of banks in terms of the size of their assets with respect to the total assets the banking system held and connected this with the riskiness of the entity to determine whether an SIB needed greater supervision than others did. This approach relies on the assumption that highly concentrated, interconnected systems are more vulnerable to distress in one SIB.

More specific to the United States context, the concentration of the financial system and the risk it imposes on competition and stability has been largely studied. For instance, before the crisis Cetorelli et al. (2007) studied the effects of concentrated markets to stability. More recently, Vives (2016) focused on competition and its effect on stability in light of the events of the financial crisis and Bikker and Spierdijk (2019) published a handbook discussing the methodologies available for policymakers to assess competition in the financial system. Indeed, regulators in the US have a wide range of indicators that allow them to estimate the status of competition in the financial sector such as the Lorenz Index and the Herfindahl-Hirschman Index. Notably, the latter is used in assessing of mergers and acquisitions.<sup>2</sup>

As can be seen, the size of banks and financial institutions continues to play a key role in determining whether an entity is considered TBTF. This article proposes a novel approach to evaluating concentration in financial systems through constructing a concentration index based on the relative size of each entity in terms of assets. Then, the main assumption of this article is that the importance of a financial entity on a system depends on its relative size, that is, its size in relation to the size of the other entities within the system. The composition of the relative participation of each financial entity on the whole system will determine the system's degree of concentration. The methodology presented uses the compositional data framework to estimate an indicator of concentration that can be tracked over time, to gain insight into a defined system's current and expected concentration. In the compositional data framework, relative information is more relevant than absolute values are. Individuals (in this case, financial institutions and more specifically banks) are not considered independently but as part of a whole (here the financial/banking system), measured in relative terms instead of absolute terms (van den Boogaart and Tolosana-Delgado, 2013).

For this article, the variable of interest is a bank's relative weight with respect to others in terms of assets. Therefore, the value of assets each bank holds is expressed in compositional terms to compare the actual composition of the banking system with the benchmark, defined as an ideal composition in which all entities have the same participation. This comparison is made through the Aitchison distance (Aitchison, 1986), which measures the distance between two compositions, to create an indicator for the concentration of the system. However, this indicator per se does not explain the evolution of the system's concentration level. Hence, the indicator is assessed over time to define the system's concentration trend. Furthermore, a time series model could forecast the future behavior of the concentration of the system, which can be used as an early warning for regulators. For instance, regulators could set thresholds (following their risk appetite) for the deviations from the forecasted behavior or rules to react to facing changes in the trend toward a more concentrated system.

This methodology will be applied to the Colombian banking system, to estimate the concentration trend over the last decade and try to forecast its behavior in the following years. Furthermore, the estimated model is compared with other potential estimation methodologies and assessed for its estimation hypotheses.

Compositional methods have scarcely been used in finance and economics. Previous applications of this methodology in economics include Belles-Sampera et al. (2016), who made an initial approach to understanding capital allocation problems as compositional problems, and Boonen et al. (2019), who went beyond this to forecast risk allocations using the compositional data framework. Furthermore, approaches to forecasting compositional data have been taken by Mills (2010), Kynčlová et al. (2015) and Zheng and Chen (2017). All of them showed the benefits of using compositional data framework in different fields.

The rest of the article is divided into four sections. The next section describes the methodology, emphasizing the compositional data framework and its application. Section three explores the Colombian banking system and the data set used in the analysis, which is included in section four, together with the comparison and assessment of the proposed model. Finally, section five summarizes the findings and concludes.

## 2. Methodology

This section explains the methodology used for the analysis, starting with a brief explanation of the compositional data framework and the centered log-ratio (clr) and isometric log-ratio (ilr) transformations required to use regular statistical methods in compositional data. The definition of the econometric models to be used follows this. Finally, the criterion for the model selection and assessment of results are explained.

<sup>2</sup> More information on the application of this index in <https://www.federalreserve.gov/bankinforeg/competitive-effects-mergers-acquisitions-faqs.htm>.

2.1. Compositional data and Aitchison geometry

A composition is defined as a row vector  $\bar{x} = [x_1, \dots, x_n]$  whose components only carry relative information (Pawlowsky-Glahn et al., 2011), which usually adds up to a constant  $\kappa$ . For simplicity, it is usually assumed that  $\kappa = 1$  without loss of generality. Therefore, the sample space in which compositional data is defined as:

$$S^n = \{\bar{x} \in \mathbb{R}^n | x_i \geq 0, i = 1, \dots, n, \sum_{i=1}^n x_i = 1\}$$

For simplicity, throughout the article it will be assumed that  $x_i > 0$ , because the log transformations to be applied later do not allow for zeros. For further discussion of the case in which the composition has zeros, please refer to Aitchison (1986).

Therefore, the methods for analyzing the data are those defined for the algebraic-geometric structure of the simplex, which are commonly referred to as Aitchison geometry. Hence, following Aitchison (1986), the perturbation operation  $\oplus$  within the simplex is defined as  $\bar{x} \oplus \bar{y} = \left( \frac{x_1 \cdot y_1}{\sum_{i=1}^n x_i \cdot y_i}, \dots, \frac{x_n \cdot y_n}{\sum_{i=1}^n x_i \cdot y_i} \right)$  for  $\bar{x}, \bar{y} \in S^n$ . Moreover, the powering operation  $\odot$  is defined as  $\lambda \odot \bar{x} = \left( \frac{x_1^\lambda}{\sum_{i=1}^n x_i^\lambda}, \dots, \frac{x_n^\lambda}{\sum_{i=1}^n x_i^\lambda} \right)$  for  $\lambda \in \mathbb{R}$ .

With this in mind, Aitchison (1986) defines the compositional mean for  $M$  compositions as  $AM_\Delta(\bar{x}_1, \dots, \bar{x}_M) = \frac{1}{M} \odot \bigoplus_{m=1}^M \bar{x}_m$ . Additionally, according to Aitchison geometry the distance between two compositions  $\bar{x}$  and  $\bar{y}$  is defined by Eq. (1):

$$AD_\Delta(\bar{x}, \bar{y}) = \|\bar{x} \ominus \bar{y}\|_\Delta = \sqrt{\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left( \ln \frac{x_i}{x_j} - \ln \frac{y_i}{y_j} \right)^2} \tag{1}$$

where  $\|\cdot\|_\Delta$  is the norm and  $\bar{x} \ominus \bar{y} = \bar{x} \oplus [(-1) \odot \bar{y}]$ . This definition brings to light an important drawback of the Aitchison geometry: logarithmic transformations do not allow for zeros in the compositions. The usual approach when dealing with zeros in compositions is to replace them with sufficiently small values. In the next section, when describing the data, an explanation on how this might affect the results will be provided.

2.2. The clr and ilr transformations

Despite the advantages of using the compositional data framework to analyze multivariate series, the sum of the vector  $\bar{x}$  being equal to a constant is a constraint. This limitation causes issues when the usual multivariate econometric models are applied. To overcome this issue, the elements of  $\bar{x} \in S^n$  need to be translated into elements of  $\mathbb{R}$ . Aitchison (1986) suggested a first attempt by defining the centered log-ratio (clr) transformation:

$$clr(\bar{x}) = \left[ \ln \frac{x_1}{g(\bar{x})}, \dots, \ln \frac{x_n}{g(\bar{x})} \right] \text{ with } g(\bar{x}) = \left( \prod_{i=1}^n x_i \right)^{1/n}$$

where  $g(\bar{x})$  is the component-wise geometric mean of the composition. As can be seen, because the clr is a component-wise transformation, it follows that  $clr(\bar{x}) \in \mathbb{R}^n$ . However, by definition, the elements of  $clr(\bar{x})$  add up to zero, which imposes a new restriction.<sup>3</sup>

To apply the usual econometric models,  $\bar{x}$  must be translated into the Euclidean space without restrictions. This can be achieved through the isometric log-ratio (ilr) transformation (Pawlowsky-Glahn et al., 2011). Starting from an orthonormal basis  $e = \{e_1, \dots, e_{n-1}\}$ , it is possible to create the matrix  $V$ , whose rows are defined as  $clr(e_j)$ . Therefore,  $V$  is  $(n-1) \times n$ . Given its construction,  $V \cdot V' = \mathbf{I}_{n-1}$  and  $V' \cdot V = \mathbf{I}_n - (1/n) \mathbf{1}'_n \mathbf{1}_n$  where  $\mathbf{I}_n$  is the identity matrix of size  $n$  and  $\mathbf{1}_n$  is the  $n$ -row vector of ones. With  $V$ , the ilr transformation can be defined as:

$$ilr(\bar{x}) = clr(\bar{x}) \cdot V$$

Another advantage of the ilr transformation is that it fulfills  $\Delta(\bar{x}, \bar{y}) = d(ilr(\bar{x}), ilr(\bar{y}))$ , where  $d(\cdot, \cdot)$  is the Euclidean distance. More details can be found in Aitchison (1986) and Egozcue et al. (2003).

Now, the issue relies on defining the orthonormal basis  $e$ . The most common method used in compositional data literature is what Egozcue et al. (2003) proposed through binary partitions. This method has the advantage of interpretability because the partitions can be such that the groups can be translated into a principal component analysis. In this process, the parts of the composition are divided into two groups. The elements of one of the groups will be assigned +1 while those of the other will be given -1. This way, the elements of one group balance with those of the other. In the subsequent steps, one of the groups is taken and divided again into two subgroups, coded +1 and -1, while the elements of the other group(s) are assigned 0. Therefore, the balances between two groups are created at each step. This process has to be completed  $n-1$  times, until each subgroup contains only one element. The result is a matrix  $(n-1) \times n$  filled with +1, -1 and 0, which can be used as the orthonormal basis for the  $V$  matrix and the ilr transformation. After the transformation, the components of  $\bar{x}$  are now represented by coordinates in  $\mathbb{R}^{n-1}$ , which allows using conventional statistical methods for multivariate analysis.

<sup>3</sup> This implies that the covariance matrix of  $clr(\bar{x})$  is singular, that is, the determinant is zero (Pawlowsky-Glahn et al., 2011).

### 2.3. Compositional VAR model

Now that the compositional data has been transformed to use regular multivariate statistical methods, the model must be established for analyzing the Colombian banking system. In this case, the modeling strategy follows the approach in Boonen et al. (2019), who analyzed the market risk in a stock portfolio.

The vector  $\bar{a}_t = [a_{1,t}, \dots, a_{n,t}]$  contains the assets held by bank  $i \in \{1, \dots, n\}$  at time  $t$ . The first approach would be to analyze this series from the multivariate statistics framework. However, the analysis can be expanded through compositional data. To achieve this, the total assets the banking system holds at time  $t$  as  $A_t = \sum_{i=1}^n a_{i,t}$  must be defined. Now,  $\bar{a}_t$  can be transformed into compositional data by defining the vector  $\bar{x}_t = [x_{1,t}, \dots, x_{n,t}]$ , whose elements will be  $x_{i,t} = a_{i,t}/A_t$ , meaning  $\bar{x}_t$  is the composition across all entities of the assets the banking system holds at time  $t$ .

Therefore, the model to be estimated corresponds to a vector autoregressive (VAR) model. In this case, the general form of the model is given by:

$$ilr(\bar{x}_t) = b + B_1 \cdot ilr(\bar{x}_{t-1}) + \dots + B_p \cdot ilr(\bar{x}_{t-p}) + \epsilon_t \quad (2)$$

where  $b$  is a vector of parameters in  $\mathbb{R}^{n-1}$ ,  $B_1, \dots, B_p$  are each  $(n-1) \times (n-1)$  matrices of parameters in  $\mathbb{R}$ ,  $\epsilon$  is the error term that follows a multivariate normal distribution with a vector of means equal to zero and covariance matrix  $\Sigma_\epsilon$ ,  $N_{n-1}(0, \Sigma_\epsilon)$ , and  $p$  is the number of lags to be considered in the model. The model in (2) can be extended by adding the first lag of the total value of the assets the banks hold in logarithms  $\log(A_{t-1})$  as a control variable. Thus, the extended model is:

$$ilr(\bar{x}_t) = b + B_1 \cdot ilr(\bar{x}_{t-1}) + \dots + B_p \cdot ilr(\bar{x}_{t-p}) + \gamma \cdot \log(A_{t-1}) + \epsilon_t \quad (3)$$

where  $\gamma$  will be the  $(n-1) \times 1$  vector of coefficients associated with the control variable.

These models are compared to the classical approach for multivariate series. As mentioned previously, the series of assets held by banks  $\bar{a}_t$  can be analyzed directly through the multivariate VAR:

$$\bar{a}_t = c + C_1 \cdot \bar{a}_{t-1} + \dots + C_p \cdot \bar{a}_{t-p} + \epsilon_t \quad (4)$$

where  $c$  will be a vector of parameters in  $\mathbb{R}^n$ ,  $C_1, \dots, C_p$  will be  $n \times n$  matrices of parameters defined in the real space and  $\epsilon_t \sim N_n(0, \Sigma_\epsilon)$  is the error term. This model can also be extended by including  $\log(A_{t-1})$ :

$$\bar{a}_t = c + C_1 \cdot \bar{a}_{t-1} + \dots + C_p \cdot \bar{a}_{t-p} + \varphi \cdot \log(A_{t-1}) + \epsilon_t \quad (5)$$

where  $\varphi$  will be the  $n \times 1$  vector of coefficients associated with the control variable.

To define the number of lags to be used in every model, the Akaike's information criterion (AIC) is used. This is a likelihood criterion for measuring the fitness of a statistical model to a sample. It is defined as  $AIC = -2\ell + 2K$  where  $\ell$  is the log-likelihood of the estimation and  $K$  corresponds to the number of parameters in the model. As can be seen, the higher the goodness-of-fit with respect to the number of parameters used, the lower the values of the information criteria. Therefore, a lower value of the criteria will indicate a better model.

### 2.4. Stationarity

Even though the VAR model can be estimated in the absence of stationarity because the estimators exist and are consistent (Sims et al., 1990), this is one of the main assumptions to determine the estimators' distribution and be able to make an inference. Therefore, the stationarity of the data set is assessed to determine whether a VAR model is suitable and under which conditions.

In its strictest sense, stationarity refers to the property of a time series of having a time invariant distribution. This means that each realization of the variable follows the same distribution, independently of the moment  $t$  (Fuller, 1996). However, in limited samples, this characteristic is hard to prove because the distribution of each realization is unknown. Therefore, weak stationarity is usually a sufficient condition. In this case, only the first two moments are required to be constant over time (i.e. independent of  $t$ ).

In the absence of stationarity, classical theory usually proposes two ways of modeling the time series: a VAR model in differences and a vector error correction (VEC) model (Lütkepohl, 2007). In the first case, the non-stationary series are differenced to obtain stationarity. Then, a VAR model is estimated with the new variables. This approach has a limitation regarding interpretability because the new coefficients will not refer to the effect of one variable on the other, but to the effect of changes in one variable on changes in the other one. For the case of compositional data, the generic model would be given by:

$$\Delta ilr(\bar{x}_t) = d + D_1 \cdot \Delta ilr(\bar{x}_{t-1}) + \dots + D_p \cdot \Delta ilr(\bar{x}_{t-p}) + \epsilon_t \quad (6)$$

where  $\Delta$  is the difference operator defined as  $\Delta ilr(\bar{x}_t) = ilr(\bar{x}_t) - ilr(\bar{x}_{t-1})$ ,  $d$  is a vector of parameters in  $\mathbb{R}^{n-1}$  and  $D_1, \dots, D_p$  are  $(n-1) \times (n-1)$  matrices of parameters in  $\mathbb{R}$ . Similarly the model can be estimated for the assets banks hold:

$$\Delta \bar{a}_t = f + F_1 \cdot \Delta \bar{a}_{t-1} + \dots + F_p \cdot \Delta \bar{a}_{t-p} + \epsilon_t \quad (7)$$

where  $f$  is a vector of parameters in  $\mathbb{R}^n$  and  $F_1, \dots, F_p$  will be  $n \times n$  matrices of parameters in the real space.

A second approach relies on the fact that even if the series are not stationary, they can maintain a long-term relationship that may be stationary. Two non-stationary time series are cointegrated if it is possible to find a linear combination of them that forms a stationary series. Therefore, relying on this assumption, the VEC estimation intends to model this long-term relationship and add it to the VAR model. For the compositional data, the model would be:

$$\Delta ilr(\bar{x}_t) = g + G \cdot ilr(\bar{x}_{t-1}) + B_1^* \cdot \Delta ilr(\bar{x}_{t-1}) + \dots + B_{p-1}^* \cdot \Delta ilr(\bar{x}_{t-(p-1)}) + \epsilon_t \quad (8)$$

where  $g$  is a vector of parameters in  $\mathbb{R}^{n-1}$ ,  $G = \sum_{s=1}^p B_s - I$  and  $B_r^* = -\sum_{s=r+1}^p B_s$  ( $B_s$  corresponds to the matrix of the VAR model's coefficients in Eq. (2), meaning both  $G$  and  $B_r^*$  are of size  $n - 1 \times n - 1$ ). Furthermore,  $G = \alpha \cdot \beta'$  can be defined, where  $\alpha$  will denote the size of the cointegration effects and  $\beta$  the cointegration matrix leading to the stationarity of the series. The dimensions of both  $\alpha$  and  $\beta$  will be determined by the rank of  $\beta$ , which is explained below. Consequently, the VEC model can be interpreted as composed via the long-term dynamics (contained in  $G$ ) and the short-term effects (modeled through  $B_r^*$ ). Similarly, for the series of assets banking institutions hold, the VEC model is defined by:

$$\Delta \bar{a}_t = h + H \cdot \bar{a}_{t-1} + C_1^* \cdot \Delta \bar{a}_{t-1} + \dots + C_{p-1}^* \cdot \Delta \bar{a}_{t-(p-1)} + \varepsilon_t \tag{9}$$

where  $h$  is a vector of parameters in  $\mathbb{R}^n$ ,  $H = \sum_{s=1}^p C_s - I$  and  $C_r^* = -\sum_{s=r+1}^p C_s$ .

Nevertheless, to estimate a VEC model, the variables' order of integration must be known, meaning how many times it is necessary to difference each variable to obtain a stationary series, and whether the variables are cointegrated, that is if a long-term relationship exists between the variables and how many variables are needed to obtain it, which would correspond to the rank of the cointegration matrix  $\beta$ .

Therefore, the data set will be diagnosed to test for the assumptions. With the information obtained, the models that suit the data's characteristics will be estimated.

### 2.5. Model comparison

When the estimated models have been obtained, the next step is to compare them. In this case, the AIC cannot be used because the models to be compared now have different variables. Therefore, this comparison is achieved through the accuracy of the forecasts the models make (considering that the intended use for the estimated models is forecasting).

In this case, for a sample with  $T$  periods, an  $h > 1$  number of periods can be selected. Then, for each period  $k \in \{T - h, \dots, T - 1\}$ , the model can be estimated with the information up to  $k$ . Next, with the estimated model, forecast periods  $\{k + 1, \dots, T\}$  and the forecasted values, the deviation from the observed values can be calculated through the mean Aitchison distance of prediction errors (MADPE):

$$MADPE(k, m) = \frac{1}{T - k} \sum_{t=k+1}^T AD_{\Delta}(\bar{x}_t, \hat{x}_t^{k,m}) \tag{10}$$

where  $AD_{\Delta}(\cdot, \cdot)$  is the Aitchison distance defined in Eq. (1) and  $\hat{x}_t^{k,m}$  is the forecast for the period  $t$  with the model  $m$  estimated with information up to  $k$ . Again, as the MADPE measures the forecasts' accuracy through the distance between the observed and predicted values, the model with the lowest MADPE will be best for forecasting.

### 3. Colombian financial system data

The Financial Superintendence of Colombia<sup>4</sup> (financial regulator) defines credit establishments as all the entities that channel resources from the public (capturing them as deposits) to individuals and companies in need of liquidity (by providing credit). These are divided into four groups, depending on the means used to channel the resources: banking establishments that obtain resources from the market via current accounts or term deposits to provide credit; financial corporations, which channel resources to companies to promote their growth; commercial financing companies, which raise funds through fixed-term deposits to provide finance for the commercialization of goods and services, and leasing operations; and financial cooperatives that are authorized to provide credit to non-associated parties. Additionally, the Colombian financial system has special official institutions, which are government-financed entities providing development financing for specific purposes or to specific clients that the legal act creating each entity defines.

The study aims to analyze the financial system's risk of concentration with a low number of institutions managing most of the assets. Notably, the government funds special official institutions (exclusively or in a high proportion). In most cases, they are created with the purpose stabilizing the system through liquidity in situations of disruption or to provide liquidity to a specific public. Therefore, considering that their activity responds to rules different from those of the free competitive activity that govern the banking establishments, these entities will be excluded from the analysis.

Following financial regulation, all credit establishments ought to provide key financial indicators monthly. Considering that the aim is to analyze the financial system through the relative importance of each entity within the system, assets would be a good size indicator. The total assets in credit establishments have been growing in recent decades with banking establishments' considerable participation, which on average was 95.6% during the last decade. The next type of entity is financial corporations with 2.6%, followed by commercial financing companies with 1.4%, and finally financial cooperatives with 0.4%. Therefore, the composition of credit establishments shows very high relevance of banking establishments, while the rest are irrelevant regarding total assets. Furthermore, financial corporations, commercial financial companies and financial cooperatives use specific means to fund their credit operations. Therefore, the risk they are exposed to differs in some sense from that of banking establishments. Consequently, the data set to be used for the analysis will only contain banking establishments.

<sup>4</sup> <https://www.superfinanciera.gov.co>

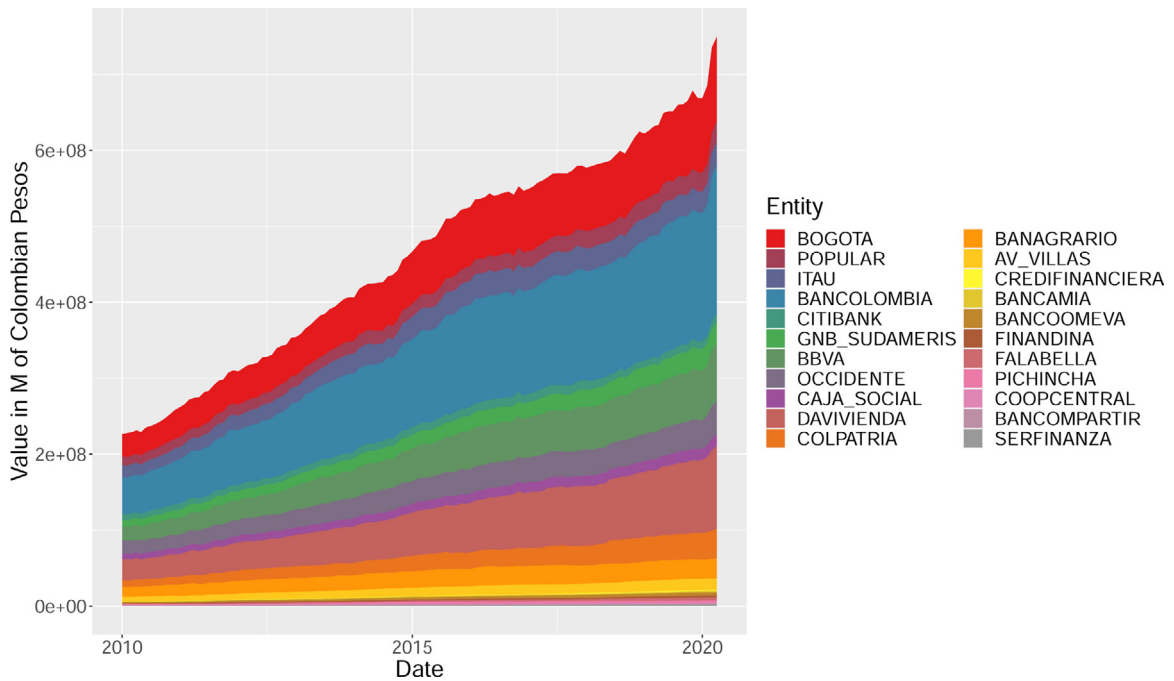


Fig. 1. Total assets per banking establishment in Colombia from January 2010 to April 2020.

### 3.1. Colombian banking system

The data set considered is from January 2010 to April 2020, which provides a base availability of 124 observations for each entity. The series consists of 26 banking establishments. However, the financial system is dynamic. Therefore, four establishments did not have observations for the full period, either because they started operating after January 2010 or they ceased operations before April 2020. These institutions participated marginally in the system. Therefore, it is assumed that their effect on the outcome is negligible, and they have been excluded to avoid issues with zeros in the data set.<sup>5</sup>

Hence, the data set to be used in the analysis consists of 22 banks and is summarized in Fig. 1, which shows the total assets each banking establishment held. Fig. 2 shows the relative proportion of the total assets (compositions) per entity. Three entities hold over 10.0% of the assets individually and can be considered as systemically important: Bancolombia, Banco de Bogota and Davivienda with 24.6%, 14.5% and 12.5%, respectively. Together, they account for more than 50.0% of the total assets in the banking system.

### 3.2. Concentration level

To analyze if there is a concentration of the assets within one or a small group of institutions, a benchmark is required for comparison. In this case, this is the uniform distribution of assets across all entities (or the neutral element of the perturbation operation for compositional data):  $\bar{z} = [z_1, \dots, z_n]$  with  $z_i = 1/n$ . Therefore, the distance between this hypothetical distribution and the actual distribution of the assets among the entities should be measured. Thus, the Aitchison distance between the neutral composition and the observed compositions of the Colombian banking system's relative assets is computed.<sup>6</sup> Note that the aim is to evaluate the trend rather than the absolute values obtained because it is not possible to define what a high concentration would be regarding the value of the distance from an objective perspective. For instance, the indicator can virtually go to infinite as the composition approaches extreme values. Therefore, the explicit threshold to define excessive market concentration would depend on risk appetite of decision makers. The calculation of the monthly distances is shown in Fig. 3. As observed, the distance has been decreasing throughout the period analyzed, which would signal a more equal distribution of assets among the banks, getting closer to the ideal uniform distribution.

<sup>5</sup> The concentration index was also calculated for the timeframes in which it was possible to have observations of each of these entities and the results are similar regarding the trend and the values, confirming the assumption that they can be neglected.

<sup>6</sup> The methodology proposed allows assessing the value of the distance between the observed composition and a composition used as benchmark. The neutral composition is chosen as benchmark, but other compositions may be chosen as benchmark based on the risk appetite of decision makers and the specificities of the system under analysis.



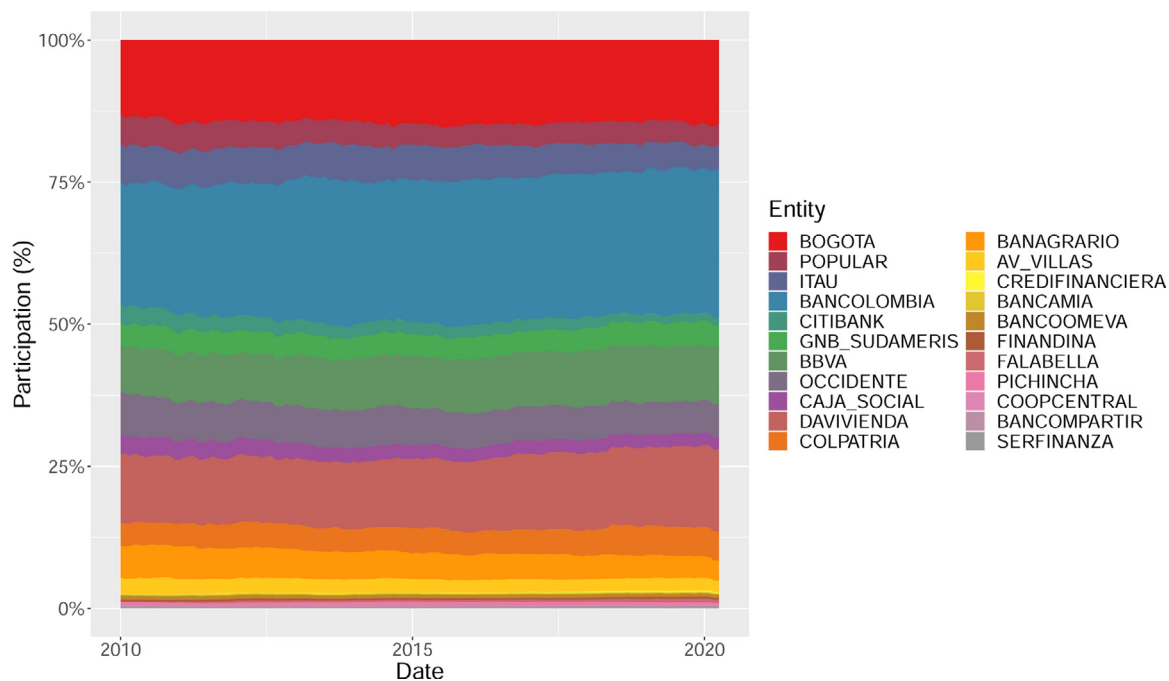


Fig. 2. Assets composition per banking establishment in Colombia from January 2010 to April 2020.

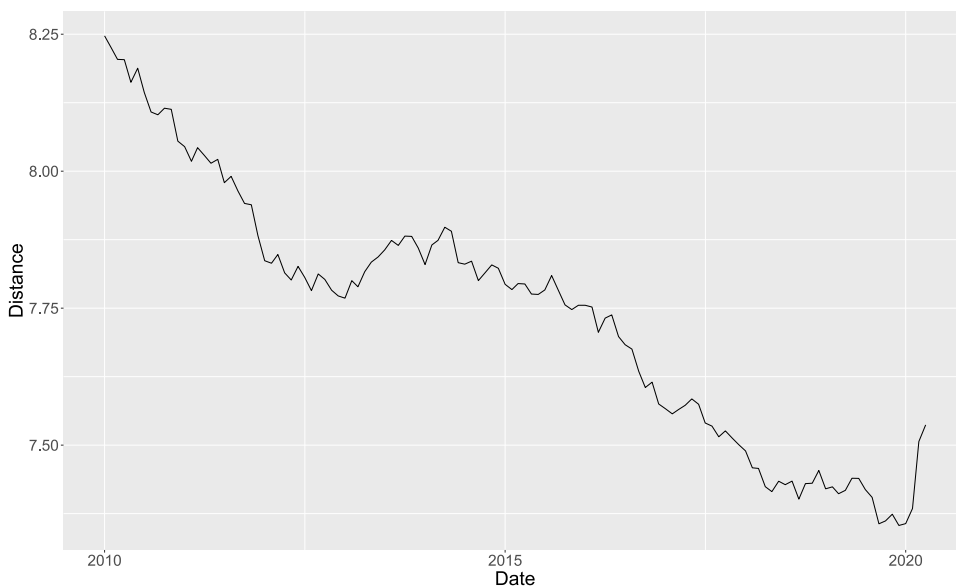


Fig. 3. Observed concentration level index of the Colombian banking system: Aitchison distance between the actual composition of assets and the hypothetical composition in which assets are equally distributed among all entities.

Now, the next step is to obtain a model that can be used to predict the asset composition’s behavior in the banking system, to assess the potential concentration of the market and use the outcome as an input for decision making on financial regulation, by creating an alert on potential TBTF institutions.

### 3.3. Data diagnosis

After defining the data set to be used, the time series needs to be diagnosed to determine which of the proposed models is more suitable to estimate. The first assumption in time series analysis refers to the stationarity of the series. To test for stationarity, unit root tests are commonly used. These tests rely on that, for a stationary series, all the roots of the characteristic equation lie inside

**Table 1**  
AIC results for different models and number of lags .

Model	N. of lags	AIC	
		Compositional model	Model with value of assets
VAR	1	-12,993.35	72,096.36
	2	-12,993.35	71,487.74
	3	-12,836.35	70,434.82
Extended VAR	1	-13,008.18	72,051.00
	2	-13,008.18	71,450.68
	3	-12,885.30	70,360.87
VAR in differences	1	-12,461.58	71,923.12
	2	-12,461.58	71,923.12
	3	-12,461.58	70,442.79

the unit circle. If the process has at least one root equal to one, the process is not stationary. A common test for stationarity is the augmented Dickey–Fuller test (Said and Dickey, 1984). In this case, the null hypothesis of unit roots in the characteristic equation is compared to the alternative of the stationarity of the process.

Hence, the augmented Dickey–Fuller test can be applied to each individual series in  $\bar{a}$  and  $\bar{x}$  (i.e., the values and compositions of assets in the Colombian banking system, respectively). For the series of the value of assets banking institutions hold, the results show that there is no stationarity for 21 out of 22 entities at 5.0% of significance. This non-stationarity was expected because most time series showed an increasing trend in Fig. 1. Similarly, for the series of compositions there is no stationarity in 19 out of 21 variables. Nevertheless, both sets are integrated of degree one, meaning that the first differences of each individual series of both  $\bar{a}$  and  $\bar{x}$  are stationary.

Because the series to be analyzed are not stationary and are integrated of degree one, it is worth testing for cointegration. In this case, considering the problem has more than two variables, the test to be used is the Johansen cointegration test (Johansen, 1991). This test uses the cointegration matrix of the data set (which contains the linear relations of the variables) and assesses the null hypothesis of no cointegration (the rank  $r$  of the matrix is zero) against the alternative of rank  $r > 0$ , meaning there is a cointegrating relationship between at least two of the variables. Subsequently, it evaluates  $r \leq 1$  against  $r > 1$  and continues recurrently until  $r \leq n - 1$  (where  $n$  is the number of variables tested). At this point, if the null hypothesis is rejected, meaning  $r > n - 1$ , it can be assumed that the matrix is of full rank ( $r = n$ ), that is, all the variables are cointegrated. Due to the data set’s high dimension (22 variables in the case of the values of assets and 21 in the case of the ilr-transformed compositions), confidence intervals for the tests cannot be computed. Nevertheless, the values of the statistics are high. Therefore, cointegration in the series can be assumed, although it is not possible to know the rank of the cointegration matrix.

The diagnosis showed that the data set is non-stationary and cointegrated, which allows estimating any of the models presented previously, each with its advantages and disadvantages. In the absence of stationarity, VAR in differences and VEC models are the usual approaches to obtain the assumed distribution of the residuals. However, the VAR model (and its extended version) can still be estimated with consistent results. Therefore, the models are compared using the MADPE described previously to determine which of the models performs better regarding forecasting.

#### 4. Results

Now that the methodology and the data set have been defined, the methodology must be applied and the results assessed. First, the number of lags  $p$  to be used for each model needs to be defined via the AIC. Then, the four proposed models will be assessed through the MADPE. Afterwards, the models’ specification assumptions will be assessed, to confirm their correct specification. Finally, the model with the best performance will be used to forecast the composition of the Colombian banking system and evaluate its concentration trend.

##### 4.1. Model selection

The AIC selection criteria’s results for the models are shown in Table 1. As seen, for the proposed specifications using compositional data, one lag should be used for the VAR models.<sup>7</sup> For the models using the value of the assets the banks held, three lags should be used in the VAR models.

Therefore, the initial VAR models from Eqs. (2) and (4) will correspond to:

$$ilr(\bar{x}_t) = b + B \cdot ilr(\bar{x}_{t-1}) + \epsilon_t \tag{11}$$

$$\bar{a}_t = c + C_1 \cdot \bar{a}_{t-1} + C_2 \cdot \bar{a}_{t-2} + C_3 \cdot \bar{a}_{t-3} + \epsilon_t \tag{12}$$

<sup>7</sup> In the case of equality, the model with the lowest number of lags is to be chosen for parsimony.



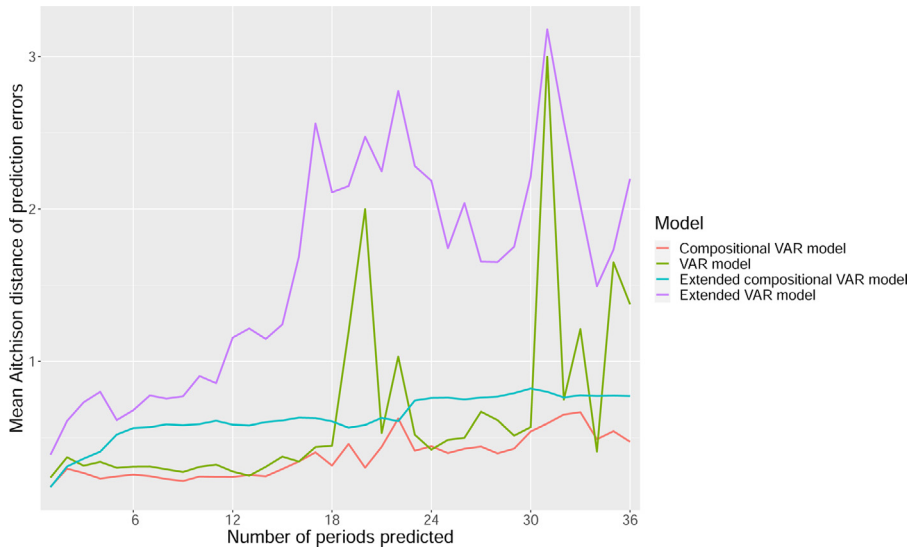


Fig. 4. Mean Aitchison distance of prediction errors (MADPE): Compositional VAR model, VAR model, extended compositional VAR model and extended VAR model.

Similarly, the extended models to be estimated from general Eqs. (3) and (5) are:

$$ilr(\vec{x}_t) = b + B_1 \cdot ilr(\vec{x}_{t-1}) + \gamma \cdot \log(A_{t-1}) + \epsilon_t \tag{13}$$

$$\bar{a}_t = c + C_1 \cdot \bar{a}_{t-1} + C_2 \cdot \bar{a}_{t-2} + C_3 \cdot \bar{a}_{t-3} + \varphi \cdot \log(A_{t-1}) + \epsilon_t \tag{14}$$

Likewise, for the models in differences, the equations are:

$$\Delta ilr(\vec{x}_t) = d + D \cdot \Delta ilr(\vec{x}_{t-1}) + \epsilon_t \tag{15}$$

$$\Delta \bar{a}_t = f + F_1 \cdot \Delta \bar{a}_{t-1} + F_2 \cdot \Delta \bar{a}_{t-2} + F_3 \cdot \Delta \bar{a}_{t-3} + \epsilon_t \tag{16}$$

For the VEC models, the rprogram automatically selects the number of lags during the estimation<sup>8</sup> and is set to two in both cases. Therefore, the models to be estimated will be:

$$\Delta ilr(\vec{x}_t) = g + G \cdot ilr(\vec{x}_{t-1}) + B_1^* \cdot \Delta ilr(\vec{x}_{t-1}) + B_2^* \cdot \Delta ilr(\vec{x}_{t-2}) + \epsilon_t \tag{17}$$

$$\Delta \bar{a}_t = h + H_1 \cdot \bar{a}_{t-1} + C_1^* \cdot \Delta \bar{a}_{t-1} + C_2^* \cdot \Delta \bar{a}_{t-2} + \epsilon_t \tag{18}$$

Now, starting from the base availability of data (124 observations), the MADPE is calculated for all models, taking  $h = 36$ , meaning the model will be used to predict up to 36 months ahead (i.e., 3 years).

Fig. 4 compares the basic VAR models with the extended models. The first notable aspect is that including the first lag of the total value of the assets (in log scale) as a control variable does not improve their forecasting power regarding the MADPE. Furthermore, although similar when predicting up to 18 periods ahead, the compositional models show a more consistent performance, with the lowest value of the MADPE almost all of the time.

Similarly, Fig. 5 includes models in differences. As seen, the compositional model continues to outperform the model with the values of the assets. In addition, the model in differences seems to perform better compared to the basic compositional VAR. However, the difference is very small, and it is not possible to conclude whether this difference is significant because the confidence intervals are not available.

The VEC models compared in Fig. 6 show interesting results. VEC models for compositional data and for the value of the assets perform very similarly to the basic compositional VAR model. This could be due to that the modeled data is cointegrated. A model that considers this stylized fact would perform better. Nevertheless, the basic VAR model for compositional data shows very similar results to those of the other models for compositional data and to the VEC for the values of assets. An initial hypothesis would suggest that even if cointegration is not taken into account, the data expressed in compositional terms already considers the interaction between variables through explaining them in terms of relative importance with respect to the others. This, combined with the fact that VAR coefficients are consistent even in the absence of stationarity, could determine why compositional models perform similarly

<sup>8</sup> The package *vars* in R estimated the VEC models.

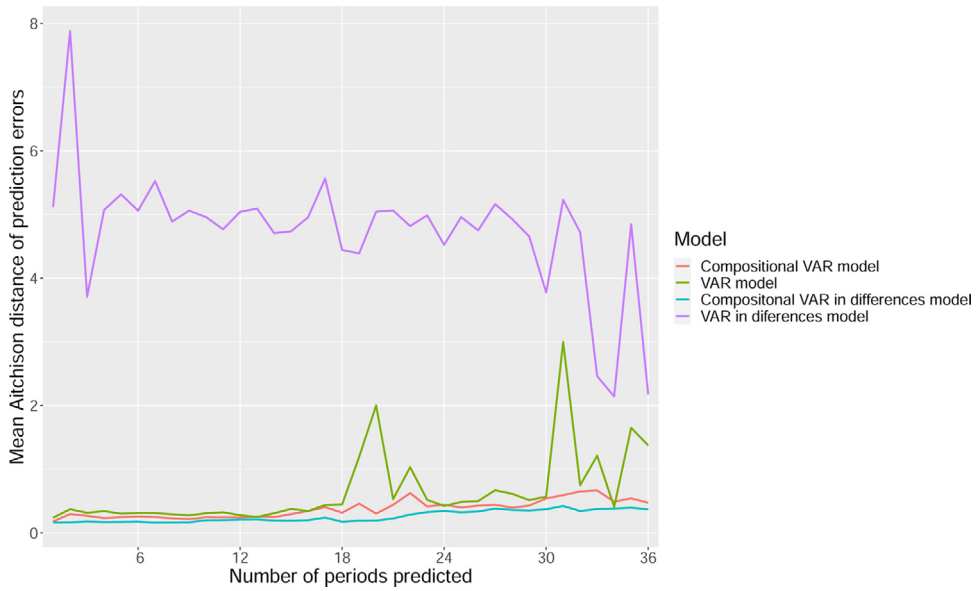


Fig. 5. Mean Aitchison distance of prediction errors (MADPE): Compositional VAR model, VAR model, compositional VAR in differences model and VAR in differences model.

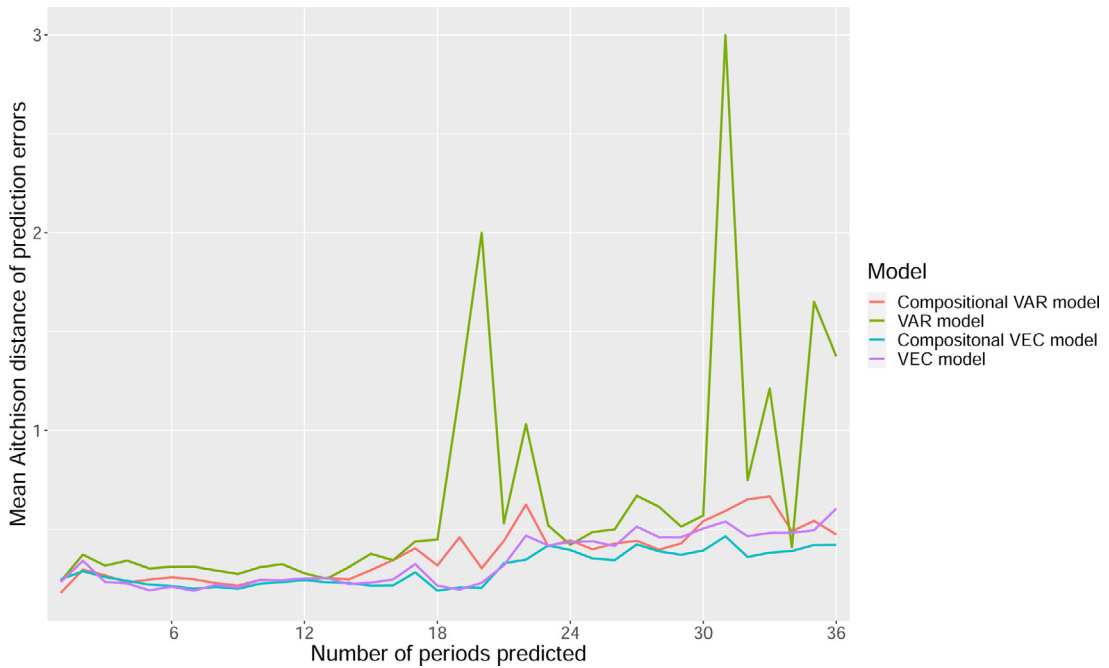


Fig. 6. Mean Aitchison distance of prediction errors (MADPE): Compositional VAR model, VAR model, compositional VEC model and VEC model.

regardless of the specification. In addition, the basic compositional VAR model has other advantages, such as reduced manipulation of the data set and the lower number of parameters to be estimated. Thus, the compositional VAR model would require the estimation of  $[(n - 1) \times (n - 1) \times p] + n - 1$  coefficients and the VAR model for the value of assets  $(n \times n \times p) + n$  coefficients. In the case of the VAR in differences models, the disadvantage comes from the fact that there are not  $T$  observations but  $T - 1$  observations. This reduces the estimation's degrees of freedom, while VEC estimations are even more complex, and the parameters cannot be easily interpreted.

Furthermore, this leads to another finding: The classical multivariate time series models for the value of asset banking institutions held are very sensitive to the specification. Indeed, misspecification can lead to a considerable decrease in the model's performance. Moreover, these models seem to be more sensitive to shocks in the series. This can be seen in the peaks in the MADPE when forecasting 19/20, 31 and 35/36 periods ahead, which are especially notorious in the basic VAR models. For example, for the model forecasting

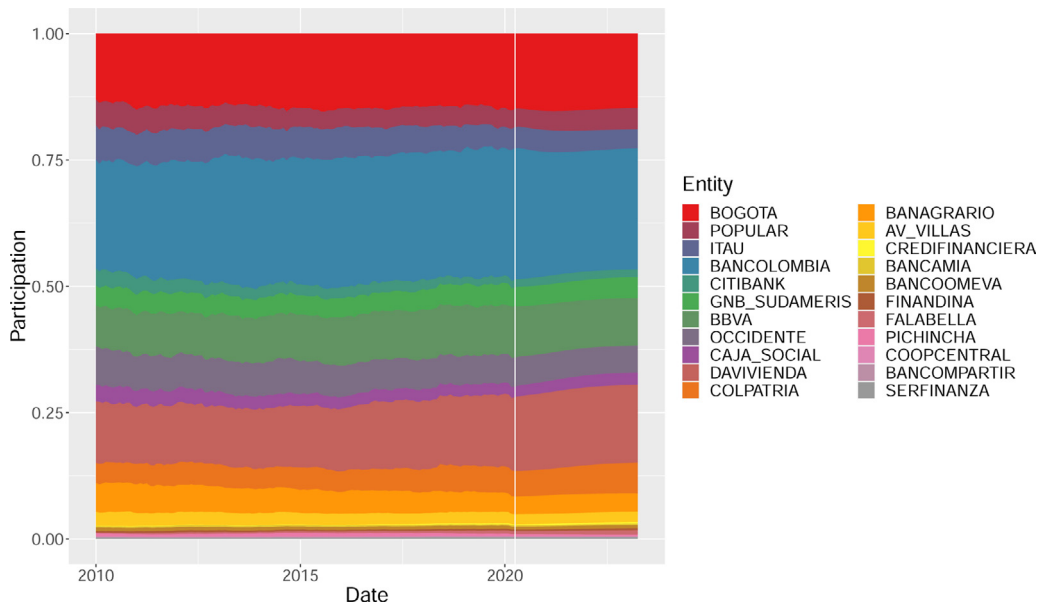


Fig. 7. Assets composition per banking establishment from January 2010 to April 2020 and forecast for May 2020 to April 2023.

20 periods ahead, the data set used for the estimation runs from January 2010 to August 2018 (and September 2018 for the forecast of 19 periods ahead). However, there was a sharp increase in the assets the two main banking institutions held in October 2018, which might explain why there is a bigger error when trying to forecast with a model estimated without information on this specific period. Relevant news for this period informed about good results for the financial system (notably for these entities), but there was no mention of acquisitions or any other important change in the assets' compositions. Similar relationships can be found in the other peaks in the MADPE for the VAR models for the value of assets.

Considering all these findings, the basic VAR model with compositional data (basic compositional VAR model) has more advantages compared to the other specifications. Therefore, the assumptions are assessed for this model. In addition, to maintain equivalence to the models using the value of assets, the same tests will be performed on the model in Eq. (12) (basic VAR model).

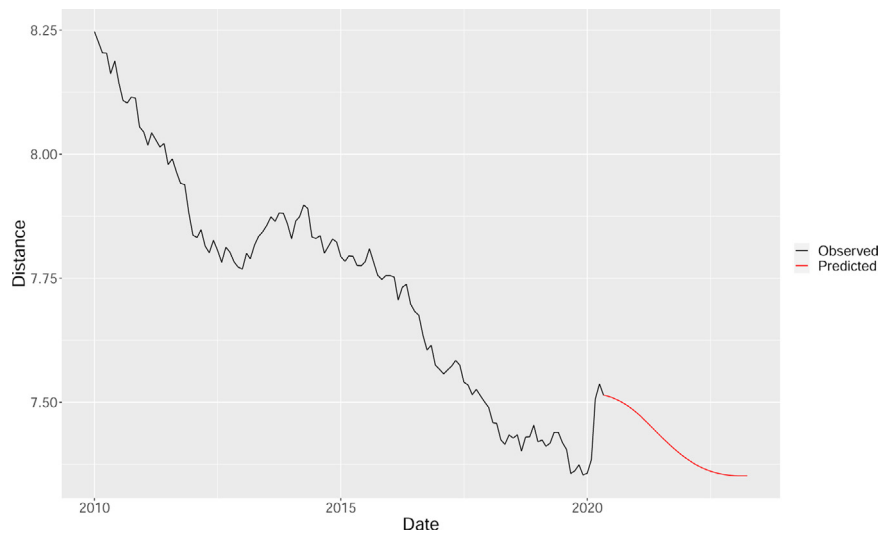
#### 4.2. Model diagnosis

Once the models have been defined, the assumptions of the models must be tested to determine whether they are correctly specified. For VAR models, the usual tests include Granger causality, autocorrelation, heteroscedasticity and normality of the residuals.

VAR models describe the joint generation process of numerous variables over time. Although in this article VAR models are used for forecasting, they can also be used for investigating relationships between variables, which the Granger causality test verifies (Granger, 1969). This test can be interpreted as a significance test for VAR models because it assesses whether adding lags of one variable improves the forecast of the other(s), that is, it contains valuable information for explaining the other variable(s) in the model. For the compositional VAR, the test shows that 5 out of the 21 series do not Granger cause the others at 5.0% significance (and only 2 when the significance is 10.0%). In the case of the VAR for the series of assets, only one of the 22 entities does not Granger cause the others. The autocorrelation test determines whether the residuals are independently distributed, which is one of the main assumptions for the estimation. Following the results for the Portmanteau test, the residuals from the compositional VAR do not show signs of serial correlation, while those from the VAR for the value of assets are autocorrelated.

Regarding the homoscedasticity of the residuals, a multivariate test cannot be performed because of the data's high dimensionality. Therefore, individual tests for the residuals of each equations are performed. For each residual series, conditional heteroscedasticity models with a different number of lags (up to 20 in this case) were estimated. Meaning, that for the compositional model, there were a total of 420 models (21 series of residuals by 20 number of lags), while for the model for the value of assets, the figure is 440 (22 series of residuals by 20 number of lags). For both the compositional and the value of assets models, at 5.0% significance, there is evidence of heteroscedasticity for at least one of the tested number of lags in four series of residuals. To obtain overall insight into the results, the proportion of models in which there is evidence of heteroscedasticity with respect to the total can be estimated to construct a pseudo-statistic for heteroscedasticity for the VAR models. For the compositional VAR, this ratio corresponds to 8.1% (34 out of 420), while for the VAR for the value of assets, it is 7.3% (32 out of 440). In both cases, at 5.0% of significance, the hypothesis of heteroscedasticity cannot be rejected, but this is possible at 10.0%.

The residuals from the compositional model do not appear to be normally distributed, while those from the model for the value of assets are normally distributed. Non-normal errors may violate some of the assumptions for the variance properties; therefore,



**Fig. 8.** Predicted concentration level index of the Colombian banking system: Aitchison distance between the predicted composition of assets and the hypothetical composition in which assets are equally distributed among all entities.

some caution should be taken in coefficient inference. However, in this case, the model is used for forecasting and the normality of residuals is not required (Lütkepohl, 2007).

#### 4.3. Forecast

After assessing the selected model (compositional model from Eq. (11)), the next step is to generate the forecast for 36 periods ahead (3 years from May 2020 to April 2023) to establish the expected composition of the Colombian banking system in the coming months and the concentration trend of the assets. Fig. 7 shows the expected composition of the assets including the forecasted period to the right of the white line. As seen, the participation of each bank in the system is not expected to undergo major changes in the coming years. There is only a slight increase in the participation of Davivienda, the third entity, regarding participation, which might lead it to surpass Banco de Bogota (the second one). Apart from that, the financial system's composition seems to remain similar to that observed in previous years.

However, as can be seen in Fig. 8, the decreasing trend of the distance between the benchmark and the actual composition of assets in the Colombian banking system is expected to continue according to the model, despite the increase seen in the beginning of 2020. It tends to stabilize by the end of the forecasted period. This result is important regarding financial stability policy because, at least for the moment, the model does not predict that the overall system is becoming more concentrated, which would threaten the stability. Furthermore, the composition of the assets across banking institutions is expected to remain stable, without major changes in the coming months.

## 5. Conclusions

Compositional data methodologies have gained popularity in recent years as a new perspective to study and model phenomena in which relative information is more relevant than absolute values. In this article, this framework is used to propose an indicator for concentration in financial systems, as applied to the Colombian banking system. The main goal was to analyze whether the difference in the participation between large and small financial institutions has been decreasing or increasing in recent years and to predict their expected evolution in the future. The concentration of the banking system was estimated as the distance between the actual composition of the financial system and the hypothetical composition in which assets are equally distributed among all entities, which was monitored over time. Therefore, the larger the distance, the higher the degree of concentration of the system in a few entities.

This analysis is relevant regarding policy making, considering the lessons learnt during the 2008 financial crisis when the global financial system was at high risk because of some “too big to fail” institutions. Thus, the proposed indicator can be used as an additional early warning for regulators about this kind of institution, to reinforce monitoring them and to maintain the overall system's stability. Furthermore, the indicator can be extended to assess the effect of mergers and acquisitions between entities on the concentration of the system and, consequently, on its stability. Conversely, it is worth mentioning that the definition of the financial system and the banking system may vary depending on the context. There are multiple other actors that have an influence in the financial system and its overall risk. For instance, shadow banking and more recently, fintech companies offer services that should also be considered when assessing financial risk (Adrian and Ashcraft, 2016). Thus, the model can be extended to account for the specificities of these entities and their effect in the system's stability.

Compositional multivariate time series methods were used to predict the banking system's future composition. These methods have shown increased performance in the prediction, considering multiple stylized facts the data showed. One of the most remarkable results is that using compositional methods provides a more robust model with lower sensibility to outliers in the data set. Furthermore, the compositional framework appears to catch the intrinsic connections between all entities in the system, which would need to be modeled through cointegrated models (such as vector error correction models). Additionally, the model has shown that it fulfills the estimation assumptions, particularly regarding autocorrelation and homoscedasticity of errors. In terms of the expected future behavior of the Colombian banking system's composition and its concentration index trend, the forecast for the next 3 years shows little variation in the participation of each entity. Furthermore, the deconcentration trend that the banking system has shown in the last decade is expected to continue over the coming months.

To conclude, the methodology applied in this article opens opportunities for multiple applications in the context of financial risk and stability. This methodology is flexible enough to be adapted to other contexts, where a different number of entities (either much higher or much lower) or a different concentration pattern of assets among the entities could lead to interesting results. For instance, special official institutions were excluded from the study because their activity does not always follow the free market rules. Given that the literature has identified a potentially significant policy trade-off between competition and stability, the concentration level index was predicted including these entities to have an idea of the effect of these firms omission on the results, and a more stable trend of the predicted concentration index was observed. Additionally, total assets were used in this study as a measure of size. In general, most definitions of SIBs include assets as the metric for the size of the entities because total assets reflect loans and other components of bank's portfolios, which multiple other types of assets can form and are also considered in the risk metrics regulators require. Nevertheless, using alternative size metrics such as total loans can also be an interesting extension of the analysis and can also be considered by regulators, depending on the specific context of each country. Likewise, the methodology can be used to measure risk within entities. Some potential applications are analyzing portfolios as compositional data and assessing risk exposure to specific assets. Nevertheless, the proposed methodology does not consider the entrance and exit of actors in the system because this would require dealing with zeros in the compositions. Indeed, this study has excluded those banks entering and leaving the market, considering they had low participation. However, this limitation can affect the results, especially in longer time series when relatively important participants can enter or leave the market. This opens an opportunity for further developing the model, considering the existing literature on how to transform compositional data in the presence of zeros (Aitchison, 1986).

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

The Spanish Ministry of Science and Innovation supported this study under grant PID2019-105986GB-C21, as did the Secretaria d'Universitats i Recerca of the Catalan Government under grant 2020-PANDE-00074.

## References

- Adrian, T., Ashcraft, A.B., 2016. *Shadow banking: a review of the literature*. Palgrave Macmillan UK, London, pp. 282–315.
- Aitchison, J., 1986. *The Statistical Analysis of Compositional Data*. Chapman & Hall, London.
- Barth, J.R., Wihlborg, C., 2017. Too big to fail: measures, remedies, and consequences for efficiency and stability. *Financ. Mark. Inst. Instrum.* 26 (4), 175–245. doi:10.1111/fmii.12083.
- Basel Committee on Banking Supervision, 2022. The basel framework. Bank Int. Settlements. <https://www.bis.org/publ/bcbs255.htm>.
- Belles-Sampera, J., Guillén, M., Santolino, M., 2016. Compositional methods applied to capital allocation problems. *J. Risk* 19 (2), 1–15.
- Bezrodna, O., Ivanova, Z., Onyshchenko, Y., Lypchanskyi, V., Rymar, S., 2019. Systemic risk in the banking system: measuring and interpreting the results. *Banks Bank Syst.* 14 (3), 34–47. doi:10.21511/bbs.14(3).2019.04. [www.scopus.com/inward/record.uri?eid=2-s2.0-85072066788](http://www.scopus.com/inward/record.uri?eid=2-s2.0-85072066788).
- , 2019. In: Bikker, J., Spierdijk, L. (Eds.), *Handbook of Competition in Banking and Finance*. Edward Elgar Publishing.
- van den Boogaart, K., Tolosana-Delgado, R., 2013. *Analyzing Compositional Data with R*. Springer Berlin Heidelberg.
- Boonen, T.J., Guillen, M., Santolino, M., 2019. Forecasting compositional risk allocations. *Insurance* 84, 79–86.
- Cetorelli, N., Hirtle, B., Morgan, D., Peristiani, S., Santos, J., 2007. Trends in financial market concentration and their implications for market stability. *Econ. Policy Rev.* 13 (1), 33–51.
- Cetorelli, N., Traina, J., 2021. Resolving “Too big to Fail”. *J. Financ. Serv. Res.* 60, 1–23.
- Egozcue, J.J., Pawłowsky-Glahn, V., Mateu-Figuera, G., Barcelo-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. *Math. Geol.* 35 (3), 279–300. doi:10.1023/A:1023818214614.
- Fuller, W.A., 1996. *Introduction to statistical time series*, second ed. John Wiley and Sons, New York. [http://www.ru.ac.bd/stat/wp-content/uploads/sites/25/2019/03/504\\_08\\_Fuller\\_Introduction-to-Statistical-Time-Series-Second-Edition.pdf](http://www.ru.ac.bd/stat/wp-content/uploads/sites/25/2019/03/504_08_Fuller_Introduction-to-Statistical-Time-Series-Second-Edition.pdf).
- Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438. doi:10.2307/1912791.
- Ioannou, S., Wójcik, D., Dymski, G., 2019. Too-Big-To-Fail: why megabanks have not become smaller since the global financial crisis? *Rev. Political Econ.* 31 (3), 356–381. doi:10.1080/09538259.2019.1674001.
- Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica* 59 (6), 1551–1580. <http://www.jstor.org/stable/2938278>.
- Kynčlová, P., Filzmoser, P., Hron, K., 2015. Modeling compositional time series with vector autoregressive models. *J. Forecast.* 34, 303–314. doi:10.1002/for.2336.
- Li, X., Tripe, D., Malone, C., Smith, D., 2020. Measuring systemic risk contribution: the leave-one-out z-score method. *Finance Res. Lett.* 36 (C). doi:10.1016/j.frl.2019.101316. [www.scopus.com/inward/record.uri?eid=2-s2.0-85075481176](http://www.scopus.com/inward/record.uri?eid=2-s2.0-85075481176).
- Lütkepohl, H., 2007. *New Introduction to Multiple Time Series Analysis*. Springer-Verlag Berlin Heidelberg.
- Mills, T., 2010. Forecasting compositional time series. *Qual. Quant.* 44 (4), 673–690. doi:10.1007/s11135-009-9229-8.

- Mishkin, F.S., Stern, G., Feldman, R., 2006. How big a problem is too big to fail? a review of gary stern and ron Feldman's "too big to fail: the hazards of bank bailouts". *J. Econ. Lit.* 44 (4), 988–1004. <http://www.jstor.org/stable/30032392>.
- Moch, N., 2018. The contribution of large banking institutions to systemic risk: what do we know? a literature review. *Rev. Econ.* 69 (3), 231–257. doi:10.1515/roe-2018-0011. [www.scopus.com/inward/record.uri?eid=2-s2.0-85058407923](http://www.scopus.com/inward/record.uri?eid=2-s2.0-85058407923).
- Omarova, S.T., 2019. The "Too big to Fail" problem. *Minn. Law Rev.* 103, 2495–2541.
- Pawlowsky-Glahn, V., Egozcue, J., Tolosana-Delgado, R., 2011. *Lecture notes on compositional data analysis*. University of Girona.
- Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71 (3), 599–607. doi:10.1093/biomet/71.3.599.
- Shull, B., 2010. Too big to fail in financial crisis: motives, countermeasures, and prospects. *Levy Econ. Inst Work. Pap. No. 601* doi:10.2139/ssrn.1621909.
- Sims, C.A., Stock, J.H., Watson, M.W., 1990. Inference in linear time series models with some unit roots. *Econometrica* 58 (1), 113–144. [https://www.princeton.edu/~mwatson/papers/Sims\\_Stock\\_Watson\\_Ecta\\_1990.pdf](https://www.princeton.edu/~mwatson/papers/Sims_Stock_Watson_Ecta_1990.pdf).
- Sorkin, A.R., 2010. *Too Big to Fail: the Inside Story of How Wall Street and Washington Fought to Save the Financial System—and Themselves*. Penguin Books.
- Stern, G., Feldman, R., 2004. *Too big to fail: The hazards of bank bailouts*. Washington, D.C.: Brookings Institution Press. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84900956521&partnerID=40&md5=10f00dd5267bf7a710fa24a39527ae79>.
- Vives, X., 2016. *Competition and Stability in Banking: the Role of Regulation and Competition Policy*. Princeton University Press.
- Zheng, T., Chen, R., 2017. Dirichlet arma models for compositional time series. *J. Multivar. Anal.* 158, 31–46. doi:10.1016/j.jmva.2017.03.006.
- Zhou, C., 2010. Are banks too big to fail? Measuring systemic importance of financial institutions. *Int. J. Central Banking* 6 (4), 205–250. <https://www.ijcb.org/journal/ijcb10q4a10.htm>.