ISSN: 1576-0162
DOI: http://dx.doi.org/10.33776/rem.v0i61.5155

Monetary Integration in South America: Election of Candidates Through Unsupervised Machine Learning

Integración monetaria en Sudamérica: elección de candidatos a través del aprendizaje automático no supervisado

León Padilla
leon.padilla@udla.edu.ec
Universidad de las Américas

Eduardo Marín Nicolalde
em22en@gmail.com
handytec Academy

Recibido: enero 2021; aceptado: enero 2022

Abstract

Applying Unsupervised Machine Learning techniques to a set of nominal variables (based on the optimum currency area [OCA] theory and the Maastricht Treaty criteria) and industrial indicators (based on similar production patterns), this paper aims to identify potential candidates for a monetary integration in South America (SA). The main conclusion is that, according to the clustering of the nominal and industrial indicators, the countries in best position for a hypothetical monetary integration in SA are Chile, Colombia, and Perú (and Ecuador to a lesser extent); countries that are generally members of the same cluster. This group of economies, which belong to the Pacific Alliance, are in a better position to meet various criteria for regional monetary integration, such as nominal convergence and similar production patterns.

Keywords: Optimum currency areas, Monetary unions, South America, Unsupervised machine learning, Cluster analysis.

Resumen

Resumen Aplicando técnicas de Aprendizaje Automático no Supervisado para un conjunto de variables nominales (señaladas por la teoría de las áreas monetarias óptimas [OCA] y los criterios del Tratado de Maastricht) e indicadores industriales (basados en patrones de producción similares), este documento tiene como objetivo identificar candidatos potenciales para una integración en Sudamérica (SA). La principal conclusión es que, según el agrupamiento de los indicadores nominales e industriales, los países en mejor posición para una hipotética integración monetaria en SA son Chile, Colombia y Perú (y Ecuador en menor medida). Este grupo de economías, pertenecientes a la Alianza del Pacífico, se encuentran en mejores condiciones para cumplir con diversos criterios de integración monetaria regional, como la convergencia nominal y patrones productivos similares.

Palabras clave: áreas monetarias óptimas, uniones monetarias, América del sur, aprendizaje automático no supervisado, análisis clúster.

JEL Classification / Clasificación JEL: F45; F15; E52.

1. Introduction

Recent changes in the economic scenario have contributed to the movement of the International Monetary System towards consolidation of currencies by regions. After a long period of economic convergence in Europe, the euro finally came into circulation in January 2002. Indeed, the euro has acted as a catalyst for deepening and broadening European integration since early in the 2000s; however, the euro crisis called into question the integrity of the Eurozone, whose structural and institutional fault lines have been revealed by the financial crisis (Pegkas et al., 2020). Moreover, other economic blocks have initiated formal processes for the adoption of a common currency. The member countries of the West African Monetary Zone (WAMZ) —The Gambia, Ghana, Guinea-Conakry, Liberia, Nigeria, and Sierra Leone— agreed that they would adopt a common currency known as the eco in the coming years. Furthermore, despite having no formal process of monetary integration, Asian economies would meet certain requirements to adopt a regional currency (De Grauwe, 2016). In the case of Latin America (LA), the debate started with Bayoumi and Eichengreen (1994), and Larrain and Tavares (2003). However, these works found little support for the idea of a common currency area in LA.

An important obstacle to achieve the economic integration in LA nations is their reluctance to lose their sovereignty (Dutta et al., 2020). In fact, most research agrees that LA countries maintain a low level of integration. Dorrucci et al. (2004) found that LA was less economically integrated than the European Union (EU) was in the 1960s and that a stronger institutional integration has indeed led to deeper economic integration. Aminian et al. (2009) stressed that despite the relative lack of formal regional trade treaties, East Asia is more integrated among its nations than LA. Márquez-Ramos et al. (2017) prove that economic, geographic, institutional, and political factors influence economic integration in the LA region. They also point out that the integration of the labor market in LA seems very far from being complete. Nevertheless, Basnet and Sharma (2013) suggested that the group with the seven largest economies LA —Argentina, Brazil, Chile, Colombia, Mexico, Peru, and Venezuela— can lead the path of integration in the region more smoothly given its macroeconomic conditions.

However, recent work has once again revived interest in monetary unions in LA. Hafner and Kampe (2018) found that the countries belonging to the CAN present better homogeneity (in terms of openness and factor mobility) compared to MERCOSUR countries. In the same vein, Padilla and Rodriguez García-Brazales (2021) demonstrate that SA as a whole is not considered not an optimal monetary area. Nevertheless, the researchers identified a group of countries (comprised of Chile, Peru, Ecuador, Brazil, and Argentina) for which the costs of a hypothetical monetary union would be relatively lower. Moreover, although most research of monetary unions in LA analyze *business cycles synchronization* through VAR models, recent studies (Bénassy-Quéré and Coupet, 2005; Issiaka and Gnimassoun, 2013; Tsangarides and Qureshi, 2008) have also used different clustering methods to establish potential candidates who share similar economic characteristics and meet different criterions needed to adopt a common currency. In the same vein, we employ Unsupervised Machine Learning techniques to find potential candidates to form a monetary area in South America (SA). Even though previous research using the cluster methodology also incorporated *nominal* variables (based on the optimal currency areas [OCA] theory and the Maastricht Treaty criteria) to explore the suitability of adopting a common currency, we include *industrial* indicators to explore the degree of *industrial* similarities among SA nations. McKinnon (1963) and Kenen (1969) point out the importance of industrial similarities to form a monetary union. In addition, while other research uses a simple average for each variable and apply these averages to identify the clusters, we go further and extract other time series components from nominal and industrial series to build a cross-sectional dataset for each set of indicators. Specifically, for each country and series in each set of indicators, the mean, the standard deviation, and linear trend (slope of a linear regression) were calculated and used in each algorithm. The structure of the paper is as follows: Section 2 presents the literature review regarding the OCA theory; in Section, 3 the literature regarding monetary unions in LA is revised. Section 4 details the clustering methodology used, and Section 5 describes the variables and date used. Section 6 outlines the preliminary findings, and the clustering method's results. The final section presents the main conclusions.

2. Theoretical Framework

The standard criteria to determine the feasibility to form a monetary union has been established in the optimum currency area (OCA) theory. Mundell (1961) proposed *factor mobility* and *labor market flexibility* as adjustment mechanisms that replace the monetary policy and the variation of exchange rates when economies face *asymmetric shocks*. McKinnon (1963) established two types of worker mobility: geographic and industrial. The first refers exclusively to the flow of employees between regions, similar to Mundell's perspective, while the second refers to the transfer of workers between industries. Later, Kenen (1969) argued that in order to have intra-industrial mobility, a high degree of similarities in terms of qualification by workers is necessary. Consequently, labor mobility works as long as the productive structure of the countries and the qualification of their workers are similar.

McKinnon (1963) integrated *economic openness* as a determinant to adopt single currency, that is, the higher the level of economic opening, the more likely the optimal exchange rate would be a fixed one. Kenen (1969) pointed out that *product diversification* could avoid the need for frequent changes in terms of trade and national exchange rates. Another criterion used for the adoption of a common currency is the *symmetry of shocks* across countries (Mundell, 1961). If the business cycles of the members of a monetary area are correlated, the cost of losing monetary policy to face imbalances should be lower (Alesina et al., 2002). Frankel and Rose (1997) believe that the rise of trade can lead to a greater correlation of economic cycles, which occurs if the demand shocks prevail.

Furthermore, the OCA theory compares benefits and cost in order to determine the adequacy of adopting a single currency. The costs of giving up the local currency includes the loss of autonomy in monetary and exchange rate policies as a mechanism to adjust to external imbalances; the loss of the ability to manage the interest rate and money supply; the loss of sovereignty for giving up the national currency; and restrictions for financing fiscal budget deficits. On the other hand, the benefits of adopting a common currency are intra-regional trade improvement by reducing transaction costs and eliminating exchange rate risk; welfare gains from less uncertainty; reducing accounting costs; credibility enhancements made by adopting an international currency; price convergence; and better economic performance (Alesina et al., 2002; De Grauwe, 2016; Obstfeld and Rogoff, 1996; Visser, 2004).

Although traditional criteria are generally of a qualitative nature, recent research has focused on designing empirical methodologies to quantify the OCA theory (Regmi et al., 2015). Among these methodologies, the "convergence criteria" of the Maastricht Treaty have been widely used to evaluate the suitability to form a monetary area. Kopits (2002) argued that *nominal convergence* within the EU was expressed formally as the convergence of reference values for inflation, interest rates, fiscal balance, public debt, adherence to the independence of the central bank, and an exchange rate policy. In this context, several authors consider that the eurozone experience as a reference that offers lessons to potential monetary integration projects, including the LA integration project. Hochreiter and Siklos (2002) and Hochreiter et al. (2002) emphasized the importance of common policies to achieve economic convergence, such as the Maastricht Treaty and the Stability and Growth Pact in the case of Europe. The authors found that in the LA region, there was a low level of convergence between Brazil and the rest of the countries; and LA shows a high level of heterogeneity given the different economic sizes, structures, and politics.

3. Monetary Unions in the Literature Review

Most of the research regarding the suitability of LA economies to form a monetary area has focused on certain groups of countries that maintain trade agreements, such as MERCOSUR[1] member countries or the Andean Community (CAN). Bayoumi and Eichengreen (1994) found low correlations in supply and demand shocks in the case of MERCOSUR. The authors concluded that there is no evidence to support the formation of a monetary union, either between LA countries or with the United States of America (USA) or Canada. Eichengreen (1998) demonstrated that a single currency among MERCOSUR members was not a good choice to reduce the exchange rate volatility. Licandro (2000) demonstrated that supply shocks correlations among MERCOSUR countries were low compared to the EU and the North American Free Trade Agreement (NAFTA). Larrain and Tavares (2003) showed that dollarization may be an option in Central America, but neither the dollarization nor a common currency would be a right decision for SA nations.

Bresser-Pereira and Holland (2009) found out that a regional currency could improve the integration process in LA by reducing the nominal exchange rate volatility, particularly for MERCOSUR. Basnet and Pradhan (2017) demonstrated that MERCOSUR countries share common trends in their main macroeconomic indicators. Hafner and Kampe (2018) demonstrated that LA and its different Regional Trade Agreements are far from being considered an optimal monetary area because the countries of LA have marked heterogeneities in terms of income, growth, and economic structure. However, the most important conclusion of his research is that the countries belonging to the CAN present better homogeneity (in terms of openness and factor mobility) compared to MERCOSUR countries. In the same vein, Padilla and Rodriguez García-Brazales (2021) demonstrate that the output trajectory of SA countries is mainly explained by country specific shocks; therefore, SA as a whole is not considered not an optimal monetary area. Nevertheless, the researchers identified a group of countries (comprised of Chile, Peru, Ecuador, Brazil and Argentina) for which the costs of a hypothetical monetary union would be relatively lower.

4. Methodology

Cluster analysis integrates a numerical method to add objects sequentially according to some metric (Kaufman and Rousseeuw, 1990). This analysis was used by Artis and Zhang (1997) to sequentially aggregate European countries according to their "economic distance". Most recent studies (Bénassy-Quéré and Coupet, 2005; Issiaka and Gnimassoun, 2013; Tsangarides and Qureshi, 2008) have also used different clustering methods to establish potential candidates who share similar economic characteristics and meet different criterions needed to adopt a common currency. These studies have been carried out specifically to determine the feasibility of forming monetary unions in African economies using several variables related to the OCA criteria. In the same vein, this paper

[1] The Southern Common Market is conformed by Brazil, Argentina, Uruguay, and Paraguay.

employs Unsupervised Machine Learning techniques to find potential candidates to form a currency area in SA. Four clustering algorithms were trained and compared in R software: (i) *K-means* clustering, (ii) Partitioning Around Medoids algorithm (PAM), (iii) *Fuzzy C-Means* clustering and (iv) *Agglomerative Hierarchical Clustering*.

4.1. K-Means clustering

K-means clustering, developed by MacQueen (1967), consists in finding clusters where intra-cluster variance —defined as the sum of squared Euclidian distances between the centroid and the observations— is minimized by using the Hartigan and Wong (1979) algorithm. At the first stage, the algorithm randomly assigns each observation to one of the clusters and computes centroids (vectors of means of all variables for the observations in the cluster). Next, Euclidean distances are computed among all data points and the centroids. Each data point is assigned to a new cluster based on the minimal distance among the centroids. The process continues iteratively until a stable solution is obtained. As this clustering method requires random initialization, simulations with different random starts were performed until a solution with a minimal sum of intra-clusters variances was found. It should be noted that the k-means algorithm is the most widely used partitional clustering algorithm (Jain, 2010). Among the reasons for its popularity, Celebi et al. (2013) emphasize: a) its conceptual simplicity, b) it is conceptually simple and easy to implement, and c) its versatility (i.e. several aspects of the algorithm can be modified). This is evidenced by hundreds of publications over the last fifty years that extend k-means clustering method (Celebi et al., 2013).

4.2. K-Medoids clustering

K-medoids is related to K-*means* clustering. Instead of centroids, Partitioning Around Medoids algorithm (PAM), developed by Kaufman and Rousseeuw (1987), identifies a representative observation (medoid) whose average dissimilarity between it and the remaining observations in the cluster is minimal (Kaufman and Rousseeuw, 1990). In other words, each centroid is the most centrally located data point in the cluster. The medoid value is chosen by the distance between every two data points of all objects. Once the first medoids are randomly assigned, each observation is grouped with the closest medoid based on a dissimilarity matrix. The algorithm then checks whether swapping the medoids in a given cluster decreases the average dissimilarity; if it does, a new medoid is selected. The process continues iteratively until a stable solution is obtained. In addition, the k-medoids algorithm has been shown to be very robust to the existence of noise or outliers and generally produces clusters of high quality (Kaufman and Rousseeuw, 1990; Ng and Han, 2002).

4.3. Fuzzy c-means clustering

Fuzzy C-Means clustering (FCM) is closely related to K-Means clustering; nevertheless, it returns a probability of belonging to every cluster for all observations. Probabilities are calculated based on the distance to the centroid of the cluster. Consequently, data points which are close to a given cluster centroid have a higher degree of belonging than those in the edge (Bezdek, 1981; Dunn, 1973). This research considered the maximum belonging probability in order to assign a country to one of the obtained clusters. The fuzzy clustering algorithm has already been used for studies related to monetary integration (Boreiko, 2003; Tsangarides and Qureshi, 2008). The advantage of fuzzy clustering is that this method takes into account the possibility that a country is similar to a country or group of countries and also shares other characteristics with another country or group of countries; therefore, it provides more information about the data than conventional clustering methods (Tsangarides and Qureshi, 2008).

4.4. Agglomerative hierarchical clustering

Agglomerative Hierarchical Clustering is an alternative approach to partitioning algorithms. According to Kassambara (2017), this algorithm starts by treating each object as a cluster singleton. Next, pairs of clusters are successively merged based on similarity information (linkage function) until all clusters have been merged into one big cluster containing all data points. Aiming to select the linkage function that best fits the data, Cophenetic Coefficient (Sokal and Rohlf, 1962) was calculated for Single, Complete, Average, and Ward linkage functions. Finally, the result is a tree-based representation of the objects (dendrogram). An advantage of using this method in this research is that at each stage, the countries that share the greatest similarities join together to form a cluster until the last group formed by all the countries is reached (Tsangarides and Qureshi, 2008).

4.5. Optimal number of clusters per method

We use the *Gap Statistic* for estimating the optimal number of clusters in our data set. Following Tibshirani, Walther, and Hastie (2001), the data $\{x_{ij}\}$, $i=1,2,\ldots,n$, $j=1,2,\ldots,p$, depend on p characteristics measured on n independent observations. $d_{ii'}$ is the distance between observations i and i'. The common distance for $d_{ii'}$ used is the squared Euclidean distance $\sum_j \left(x_{ij}-x_{i'j}\right)^2$. It is assumed that the data has been clustered into k clusters $C_1,C_2,\ldots,C_k$, with C_r denoting the indices of observations in cluster r, and $n_r=|C_r|$. Where:

$$D_r=\sum_{i,i'\in C_r} d_{ii'} \tag{1}$$

The sum of pairwise distances for all points in cluster $\mathbf{r}$, and set

$$W_k = \sum_{r=1}^{k} \frac{1}{2n_r} D_r \tag{2}$$

Considering that de distance $\mathbf{d}$ is the squared Euclidean distance, then W_k is the pooled within-cluster sum of squares around the cluster means. Tibshirani, Walther, and Hastie (2001) propose standardizing the graph of $\log(W_k)$ by comparing it with its expectation given a null reference distribution of the data. Therefore, the optimal number of clusters is then the value of $\mathbf{k}$ for which $\log(W_k)$ falls the farthest below this reference curve. Tibshirani, Walther, and Hastie (2001) define the gap statistic such as:

$$Gap_n(k) = E_n^*\{\log(W_k)\} - \log(W_k) \tag{3}$$

Finally, E_n^* denotes expectation under a sample of size $\mathbf{n}$ from the reference distribution. Thus, $\hat{k}$ will be the value that maximizes $Gap_n(k)$ after taking the sample distribution into account. This estimate is applicable to any clustering method and distance measure $d_{ii'}$.

5. Data and variables

For each country considered in the research[2], Nominal and Industrial annual time-series indicators (2001–2017) were employed to choose potential candidates for a monetary area in SA. On the one hand, *nominal* indicators include seven variables related to the OCA theory and the Maastricht Treaty criteria: inflation, government balance, debt, interest rate, nominal exchange rate variation, regional trade intensity, and labor flexibility. The aim of forming clusters from these variables is to make it possible to establish the level of *nominal* convergence between SA countries. Other works have also incorporated several of these *nominal* variables to explore the suitability of adopting a common currency (see Bénassy-Quéré and Coupet 2005; Tsangarides and Qureshi 2008).

On the other hand, following the approach pointed out by McKinnon (1963) and Kenen (1969), which establishes the importance of industrial similarities to form a monetary union, we include *industrial* indicators to explore the degree of *industrial* similarities among SA nations. According to Blackman (1998) industrial convergence occurs when economies share a trend in the evolution of technological services and industrial structures. Bröring and Leker (2007) distinguished between two types of industrial convergence. The input-side convergence, the first type, is mainly driven by technological factors. The second type is output-side convergence, which is related to market-

[2] Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Paraguay, Peru, and Uruguay. Venezuela was excluded due to extreme values in all variables.

driven factors. Therefore, *industrial* indicators include innovation, market size, competitiveness, real effective exchange rate variation, productivity growth, energy consumption of the industry, and market concentration.

Table 1 shows a brief description of each variable by indicator group, and Table 13 specifies the data sources, periodicity, and imputation method for each variable. For an incomplete time series, three imputation methods were employed. The first available value of the series was used for imputation on missing values before 2006; the mean was used for missing values between 2007 and 2015. Finally, missing values for 2017 were forecasted using seasonal autoregressive integrated moving average (SARIMA) models. In addition to the SARIMA model accounting for seasonal effects, this model also has a stable layout and it is expressly designed for time series data (Divisekara, Jayasinghe, and Kumari, 2020).

Traditional clustering algorithms require cross-sectional data to be trained. In that regard, time series components were extracted from nominal and industrial series to build one cross-sectional dataset for each set of indicators. Hence, for each country and series in each set of indicators, the mean (*mean*) volatility measured as standard deviation (*sd*) and linear trend (*slp*) —slope of a linear regression on the indicator with time as independent variable— were calculated. Additionally, correlations (*corr*) among annual GDP cyclical components (1960-2017) of each country and the three main trading partners of the region[3] were calculated and merged to each cross-sectional dataset. The Annual GDP cyclical components was obtained by using the Baxter-King filter. Table 2 below shows the correlations among cyclical components. Finally, an automatic variable selection procedure was performed over each cross-sectional dataset to avoid skewed clustering results due to highly correlated variables (Sambandam, 2003). According to Cohen (1988), a correlation coefficient of |0.10| represents a weak association; a correlation coefficient of |0.30| is considered a moderate correlation; and a correlation coefficient of |0.50| or greater represents a strong correlation. In this sense, we decided to use a threshold of |0.75|; that is, we use a medium threshold within the range of strong correlations. Figures 1 and 2 show a heat plot based on Pearson correlations among variables in the *nominal* and *industrial* indicators' dataset. The variables are grouped according to a high correlation intensity motivating variable selection procedure, which keeps 13 for nominal variables and 15 for industrial variables as shown in Table 3

6. Results

6.1. Preliminary results

The nominal indicators are shown in Table 4. The countries with the best price stability are Bolivia, Chile, Colombia, Ecuador, and Peru. Furthermore, the low level of inflation in these countries also coincides with lower interest

[3] USA, EU, and China.

Table 1. Description of variables

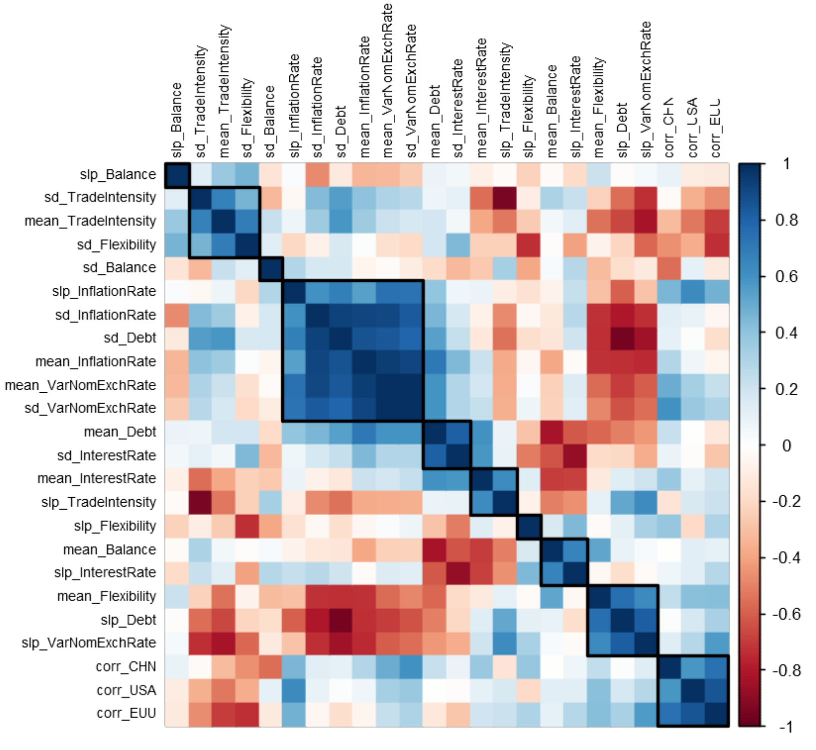
Nominal and/or traditional	Description
(i) Inflation	The inflation variable was constructed by variation of consumer price index at the *end of* the period for each country.
(ii) Government balance	This variable was calculated by general government primary net lending/borrowing sheet as a percentage of the GDP.
(iii) Debt	The debt variable was calculated by gross debt as percentage of the GDP for each country.
(iv) Lending interest rate	In the absence of continuous issuance of bonds by some SA governments, the *lending interest rate* was used as proxy variable at the long-term interest rate.
(v) Nominal exchange rate Variation	In the absence of a similar mechanism in SA, the variation of the nominal exchange rate of each country was calculated as proxy variable of the use of exchange rate policy.
(vi) Business cycle correlations	Correlations among the annual GDP cyclical component (1960-2017) of each country and the three main trading partners of the region (USA, EU, and China) were obtained and merged to the summarized dataset.
(vii) Regional trade intensity	This indicator was calculated as the ratio between intraregional trade and total trade. For each country (i) is the total of intraregional imports and exports divided by the total imports and exports. That is: (.
(viii) Work flexibility	This indicator was obtained by the "Flexibility" component of the Global Competitiveness Index (GCI).
Industrial	Description
(a) Concentration	We calculate the Herfindahl–Hirschman Index (HHI) with the data of the productive sectors of the SA.
(b) Innovation	According to Sancho (2007), the most competitive companies are those that have greater capacity for innovation. Therefore, innovation is a key factor in a country's industry. The "Innovation pillar" component of the Global Competitiveness Index (GCI) was used.
(c) Market Size	According to Segura (1993), the industrial structure is also directly related to the number of competitors. We used the "Market Size pillar" component of the GCI.
(d) Competition	Segura (1993) states that internal competition in a market determines important characteristics of an industry. The "Competition" component of the GCI was used.
(e) Variation of real effective exchange rate	Issiaka and Gnimassoun (2013) argue that the real exchange rate as well as being an indicator of competitiveness is also a useful indicator for determining the viability of a monetary union.
(f) Labor Productivity growth	Productivity change can increase during an expansionary economic cycle and decline during a recession. Hence, productivity growth is used as a proxy variable for the correlation of economic cycles in addition to being directly related to industrial patterns.
(g) Energy consumption of the industry	As a proxy variable for industrial activity, an index of energy demand was constructed by the industrial sector. The ratio was calculated with the energy consumption of the industry on the total energy consumption for each country.

rates. On the other hand, Chile, Colombia, Ecuador, Paraguay, and Peru show better macro-fiscal performance with low levels of primary fiscal deficits and lower debt levels. Brazil has the largest primary fiscal deficit in the region; and Argentina, Brazil, and Uruguay have the highest levels of debt. Argentina is the country with the highest exchange rate volatility. In addition, Argentina, Bolivia, Paraguay, and Uruguay show a tendency to monetary devaluations. The countries with the greatest intensity of regional trade are Bolivia, Argentina, Paraguay, and Uruguay. However, a negative aspect is that all SA countries show a decreasing trend in the intensity of regional trade. Chile, Colombia, Peru, and Uruguay are the economies with the greatest labor flexibility. Finally,

Table 2. Correlations among annual GDP cyclical components (1960–2017)

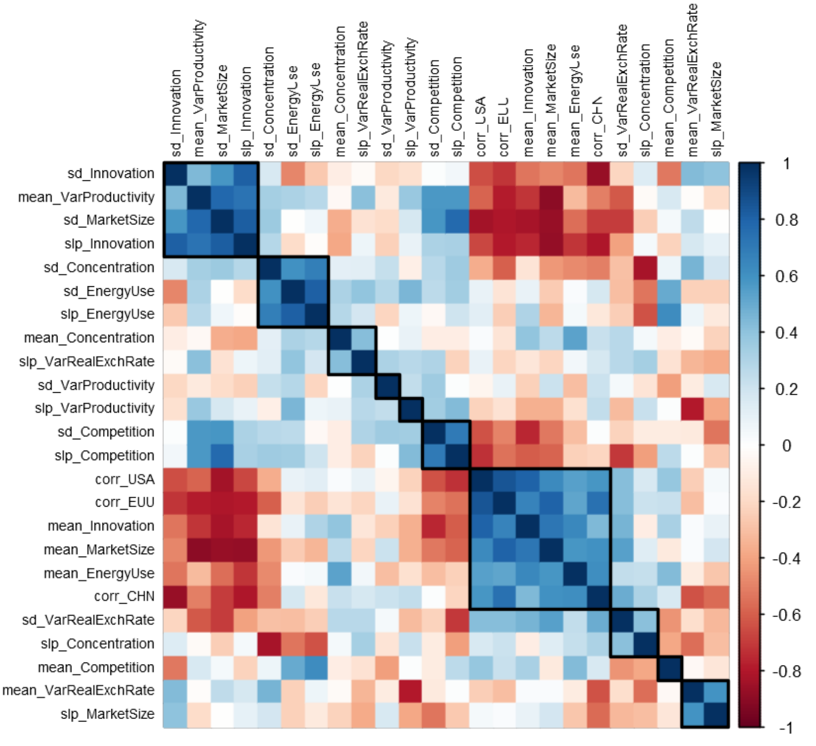
	USA	China	European Union
Argentina	0.387	0.440	0.421
Bolivia	0.117	0.069	0.061
Brazil	0.240	0.355	0.354
Chile	0.508	0.310	0.476
Colombia	0.326	0.441	0.518
Ecuador	0.039	0.001	0.087
Peru	-0.022	0.318	0.096
Paraguay	0.254	0.337	0.371
Uruguay	0.172	0.256	0.042

Figure 1. Correlations among nominal indicators' time-series components



Note: The procedure keeps only one variable from couples of variables with correlation coefficients higher than a threshold of |0.75|.
Source: Author's calculations.

Figure 2. Correlations among industrial indicators' time-series components



Note: The procedure keeps only one variable from couples of variables with correlation coefficients higher than a threshold of |0.75|.
Source: Author's calculations.

Table 3. Retained variables by time-series components

Time-series components	Nominal Indicator	Industrial Indicator
Mean	Balance Interest rate Trade intensity Flexibility	Competition Var Real Exchange Rate Var Productivity Energy use Concentration
Volatility (SD)	Inflation rate Balance Flexibility	Innovation Competition Var Real Exchange Rate Var Productivity
Linear Trend (slope)	Balance Interest rate Trade intensity Flexibility	Competition Market Size Var Real Exchange Rate Energy use Concentration
Correlation	USA China	USA

Argentina, Brazil, Chile, Colombia, and Peru have a high correlation of the economic cycle with the USA, China, and the EU. In summary, the observed patterns reveal that the countries with the best macroeconomic performance who would best adjust to the requirements to adopt a common currency are Chile, Colombia, and Peru (to a lesser extent Ecuador). However, a negative aspect is that these countries have a low level of intraregional trade.

Table 4. Statistics of Nominal Indicators

Indicator	stats	Argentina	Bolivia	Brazil	Chile	Colombia	Ecuador	Paraguay	Peru	Uruguay
Inflation	mean	13.927	5.058	6.759	3.269	4.923	5.377	6.494	2.749	8.721
	sd	12.053	3.255	2.388	2.214	1.944	5.067	3.654	1.733	4.906
	slp	0.940	0.110	-0.105	0.051	-0.189	-0.591	-0.507	0.154	-0.204
Gover-nment balance	mean	-1.695	-2.091	-4.095	0.900	-1.843	-1.196	0.128	0.017	-2.177
	sd	3.025	4.451	2.392	3.494	1.263	3.256	1.616	1.900	1.268
	slp	-0.364	0.204	-0.229	-0.253	-0.003	-0.546	-0.074	0.031	-0.057
Debt	mean	67.378	47.633	67.345	11.074	39.771	25.985	25.986	31.683	71.972
	sd	34.864	17.121	6.037	4.974	5.997	10.289	13.015	9.974	18.184
	slp	-4.844	-2.706	-0.353	0.379	0.186	-1.212	-1.913	-1.916	-2.400
Lending interest rate	mean	21.246	10.382	57.954	12.877	15.794	9.873	15.248	20.741	25.156
	sd	8.627	5.325	16.670	1.907	2.989	2.243	1.321	3.272	27.222
	slp	0.343	-0.959	-2.902	-0.268	-0.531	-0.375	-0.002	-0.573	-3.180
Nominal exchan-ge rate	mean	0.244	0.008	0.053	0.018	0.031	0.000	0.039	-0.001	0.072
	sd	0.517	0.041	0.163	0.089	0.128	0.000	0.140	0.049	0.184
	slp	-0.011	-0.005	0.005	0.003	0.007	0.000	-0.006	0.005	-0.009
Regional trade intensity	mean	0.406	0.569	0.159	0.202	0.187	0.261	0.528	0.218	0.415
	sd	0.047	0.037	0.012	0.029	0.039	0.038	0.055	0.035	0.056
	slp	-0.009	-0.005	-0.001	-0.006	-0.007	-0.006	-0.010	-0.007	-0.009
Work fle-xibility	mean	3.047	3.444	3.758	4.922	4.391	3.589	3.867	4.440	4.021
	sd	0.176	0.493	0.184	0.299	0.162	0.222	0.346	0.183	0.547
	slp	0.000	-0.088	0.014	-0.057	0.021	0.003	0.067	0.026	-0.107
Business cycle correla-tions	USA	0.387	0.117	0.240	0.508	0.326	0.039	-0.022	0.254	0.172
	China	0.440	0.069	0.355	0.310	0.441	0.001	0.318	0.337	0.256
	EU	0.421	0.061	0.354	0.476	0.518	0.087	0.096	0.371	0.042

The industrial variables—which represent the productive characteristics of the economies—are presented in Table 5. Brazil, Chile, and Colombia are the countries with the best potential in terms of innovation. The economies with the greatest economic weight within the region—Brazil, Argentina, Colombia, Chile, and Peru—are also those with the best market size index within SA. Chile, Peru, and Uruguay have the markets with the best competition. In relation to the level of international commercial competitiveness with respect to other countries, Colombia has gained the most competitiveness since 2002 because its real effective exchange rate (REER) has depreciated. The countries with the greatest increases in productivity are Bolivia, Paraguay, and Uruguay. The countries with the highest energy consumption are Brazil, Chile, and Paraguay. Brazil and Paraguay present the highest levels of productive concentration, measured on the Herfindahl and Hirschman Index (IHH). In summary, it is difficult to determine clear patterns in this set of indicators due to the industrial heterogeneity between the economies.

Table 5. Statistics of Industrial Indicators

Indicator	stats	Argentina	Bolivia	Brazil	Chile	Colombia	Ecuador	Paraguay	Peru	Uruguay
Innovation	mean	3.055	2.528	3.439	3.463	3.201	2.692	2.237	2.747	3.094
	sd	0.094	0.350	0.133	0.059	0.045	0.294	0.161	0.030	0.099
	slp	-0.001	0.052	-0.021	0.002	0.000	0.038	0.030	-0.001	0.017
Market Size	mean	4.924	3.214	5.608	4.401	4.672	3.771	3.233	4.309	3.263
	sd	0.065	0.142	0.078	0.098	0.092	0.156	0.170	0.117	0.124
	slp	-0.001	0.016	0.012	0.008	-0.002	0.022	-0.014	0.016	-0.006
Competition	mean	3.188	3.374	3.603	4.922	3.790	3.509	3.923	4.161	4.211
	sd	0.230	0.132	0.090	0.056	0.119	0.213	0.457	0.268	0.261
	slp	-0.043	0.008	0.009	-0.007	0.013	0.034	0.090	0.053	0.049
Variation of real effective exchange rate	mean	0.002	0.001	0.019	0.015	-0.014	0.042	0.002	0.003	0.004
	sd	0.169	0.060	0.105	0.076	0.117	0.075	0.064	0.032	0.085
	slp	0.011	0.004	-0.003	-0.002	-0.009	-0.009	0.001	0.000	0.007
Productivity growth	mean	0.009	0.023	0.006	0.014	0.004	0.013	0.024	0.015	0.020
	sd	0.050	0.030	0.034	0.014	0.028	0.033	0.030	0.059	0.031
	slp	0.000	0.004	0.001	-0.001	0.003	-0.002	0.003	0.004	0.003
Energy consumption	mean	0.299	0.244	0.376	0.366	0.282	0.217	0.306	0.270	0.285
	sd	0.024	0.016	0.019	0.032	0.018	0.009	0.024	0.067	0.087
	slp	-0.005	-0.003	-0.003	0.006	-0.003	0.000	-0.004	0.002	0.017
Concentration	mean	1543.5	1487.8	1802.7	1430.8	1365.3	1332.6	1466.2	1402.1	1716.8
	sd	24.863	24.181	21.683	19.935	10.332	61.571	20.915	44.915	72.214
	slp	3.022	4.329	-2.378	-3.289	-0.169	-11.472	0.642	-9.049	-10.063
Correlation	USA	0.387	0.117	0.240	0.508	0.326	0.039	-0.022	0.254	0.172
	China	0.440	0.069	0.355	0.310	0.441	0.001	0.318	0.337	0.256
	EU	0.421	0.061	0.354	0.476	0.518	0.087	0.096	0.371	0.042

6.2. Clustering methods results

Each clustering method used different criteria to establish the resultant groups of the study[4]. For hierarchical clustering, we calculated the Cophenetic Coefficient[5] (Sokal and Rohlf, 1962) for Single, Complete, Average, and Ward linkage functions (see Table 6). With a value of 0.772 and 0.769, Average linkage obtained the highest Cophenetic Coefficient in *nominal* and *industrial* sets of variables, respectively, and was retained for the analysis. The latter implies that distances between two clusters will be computed as the average distances between all the data points in cluster 1 and 2. In the case of the Fuzzy clustering, the probability of belonging coefficients obtained for each cluster are in Table 7. We assigned each country to the cluster whose probability is maximal. With the aim of selecting the number of clusters that best fits the data, the GAP statistic was calculated iteratively for each clustering algorithm and data set. Table 8 shows the GAP statistic value and the number of clusters obtained.

The results of clustering are shown in the Figures 1-8. In addition, the most important results of the *nominal* and *industrial* indicators are summarized in Tables 9 and 10, respectively. In the case of *nominal* indicators, the results of the different clustering methods identified three groups of countries. The first group is formed by Chile, Colombia, and Peru. These economies coincide in all clustering methods according to *nominal* variables. In other words, within the SA region, these three countries have better convergence in nominal terms compared to the rest of the countries. More importantly, this group of economies could represent a central nucleus to adopt a common currency in SA, and these countries meet various criteria for regional monetary integration. Argentina, Bolivia, Ecuador, Paraguay, and Uruguay, three of which are part of MERCOSUR, have an inferior macroeconomic performance compared with the first group. Lastly, Brazil does not show a clear cluster pattern.

The second group of indicators that could capture industrial characteristics has fewer clear patterns than *nominal* ones. This can be explained by the high productive heterogeneity among the countries of the region. However, we can identify certain country pairings that are repeated between different clustering methods. In the hierarchical and Fuzzy clustering methods, Chile is grouped with Colombia, Peru with Ecuador, and Argentina with Brazil. In the K-means and Pam cluster methods, Argentina and Bolivia are grouped. Paraguay and Uruguay do not present significant coincidences with other countries.

[4] In order to determine the optimal number of clusters for each algorithm the Gap Statistic, developed by Tibshirani et al. (2001), was computed.
[5] This coefficient measures the correlation between the original distances and the distance obtained according to the method used. The higher the co-behavior coefficient, the better the grouping adjustment. Further, for final validation it is necessary to determine the *optimum number of clusters* that represent the structure.

Table 6. Cophenetic coefficient (hierarchical clustering) for SA

Clustering method	Nominal	Industrial
Ward. D	0.5636071	0.7409198
Single	0.5203678	0.7492399
Complete	0.7417099	0.7300584
Average	0.7721513*	0.768598*

Note: *Best linkage function.
Source: Author's calculations.

Table 7. Fuzzy clustering—Probability of belonging for SA

Country	Probability of belonging				
	Nominal		Industrial		
	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3
Argentina	0.61	0.39	0.44	0.25	0.31
Bolivia	0.74	0.26	0.22	0.42	0.37
Brazil	0.41	0.59	0.55	0.23	0.23
Chile	0.35	0.65	0.46	0.26	0.29
Colombia	0.18	0.82	0.57	0.20	0.23
Ecuador	0.70	0.30	0.18	0.55	0.27
Paraguay	0.59	0.41	0.21	0.31	0.48
Peru	0.21	0.79	0.22	0.47	0.32
Uruguay	0.60	0.40	0.23	0.31	0.47

Note: The coefficients of belongingness of the country to each cluster.
Source: Author's calculations.

Table 8. GAP Statistic and optimal number of clusters for SA

Clustering method	Nominal		Industrial	
	GAP Statistic	Number of clusters	GAP Statistic	Number of clusters
K-means	0.332	2	-0.498	5
K-medoids	0.049	4	-0.681	5
Agglomerative Hierarchical Clustering	-1.265	2	-0.675	5
Fuzzy C-Means Clustering	0.212	2	0.157	3

Source: Author's calculations.

Table 9. Clustering methods summary for SA: nominal variables

Number of clusters	Clustering method			
	K-means cluster	PAM cluster	Hierarchical cluster (Average)	Fuzzy cluster
1	Chile Colombia Peru Paraguay Uruguay Argentina Bolivia Ecuador	Chile Colombia* Peru Paraguay Uruguay	Chile Colombia Peru Brazil	Chile Colombia Peru Brazil
2	Brazil	Brazil*	Argentina Bolivia Ecuador Paraguay Uruguay	Argentina Bolivia Ecuador Paraguay Uruguay
3	-	Bolivia Ecuador*	-	-
4	-	Argentina*	-	-

Note: *Country is the center of the cluster (medoid).
Source: Author’s calculations.

Table 10. Clustering methods summary for SA: industrial variables

Number of clusters	Clustering method			
	K-means cluster	Pam cluster	Hierarchical cluster (Average)	Fuzzy cluster
1	Chile Peru	Chile Brazil*	Chile Colombia	Chile Colombia Argentina Brazil
2	Colombia Ecuador	Colombia Paraguay*	Ecuador Peru	Ecuador Peru Bolivia
3	Argentina Bolivia Paraguay	Argentina Bolivia*	Argentina Brazil	Paraguay Uruguay
4	Brazil	Peru* Uruguay	Bolivia Paraguay	-
5	Uruguay	Ecuador*	Uruguay	-

Note: *Country is the center of the cluster (medoid).
Source: Author’s calculations.

Figure 3 K-means for SA: nominal variables

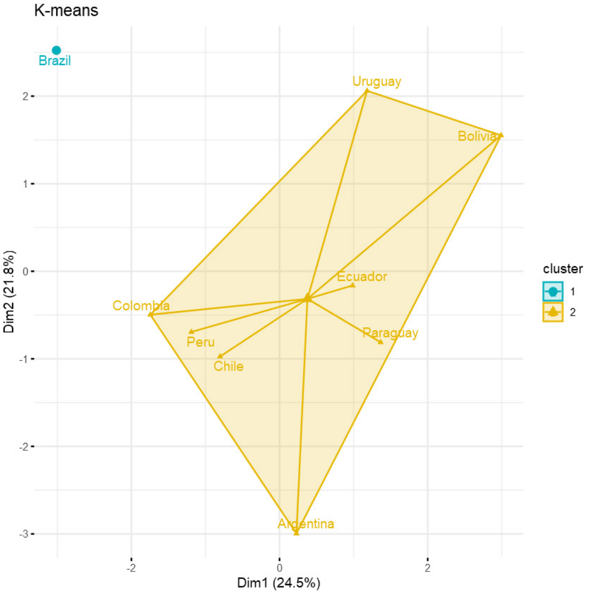


Figure 4 K-medoids for SA: nominal variables

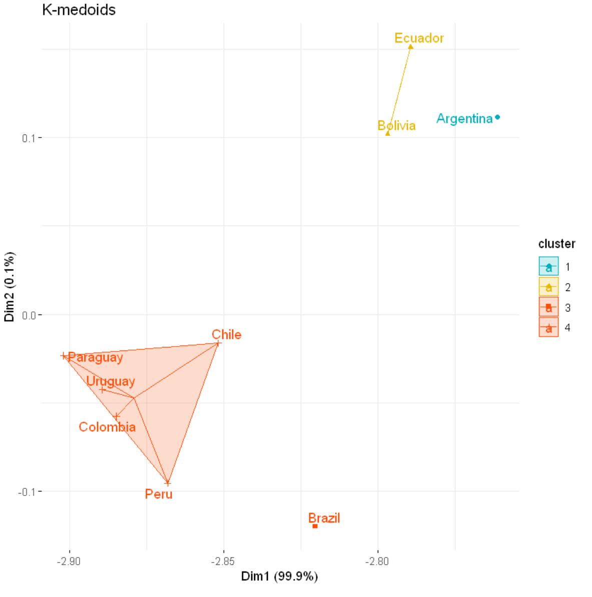


Figure 5 Hierarchical Average linkage for SA: nominal variables

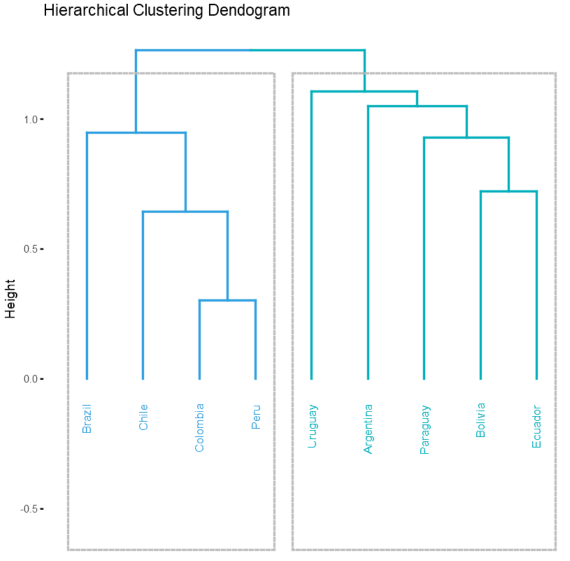


Figure 6 Fuzzy clustering for SA: nominal variables

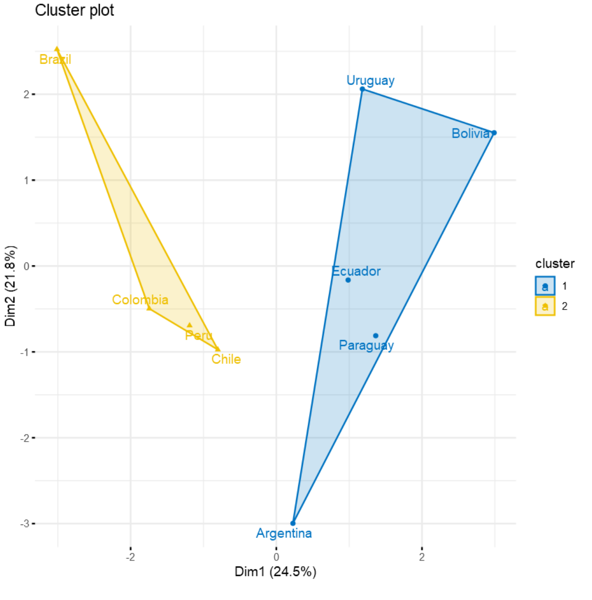


Figure 7 K-means for SA: industrial variables

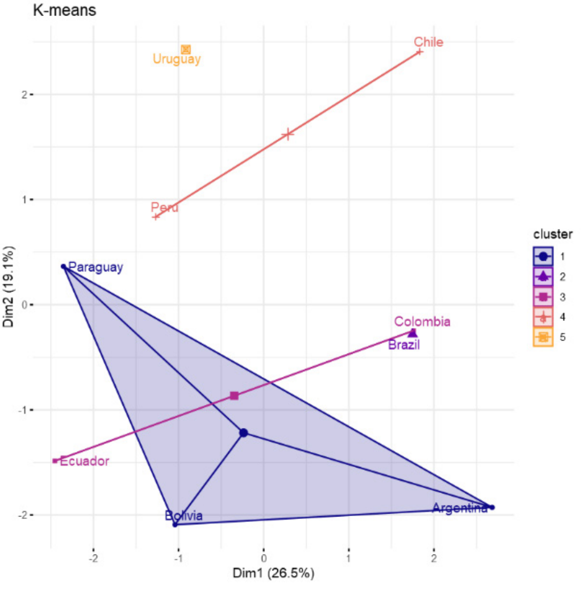


Figure 8 K-medoids for SA: industrial variables

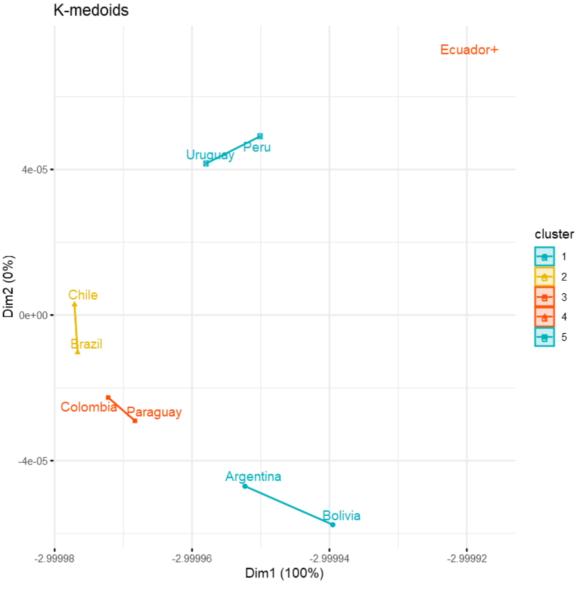


FIGURE 9 HIERARCHICAL AVERAGE LINKAGE FOR SA: INDUSTRIAL VARIABLES

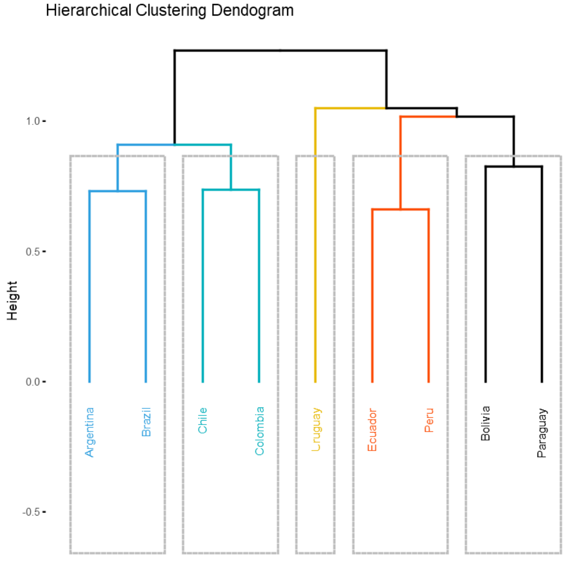
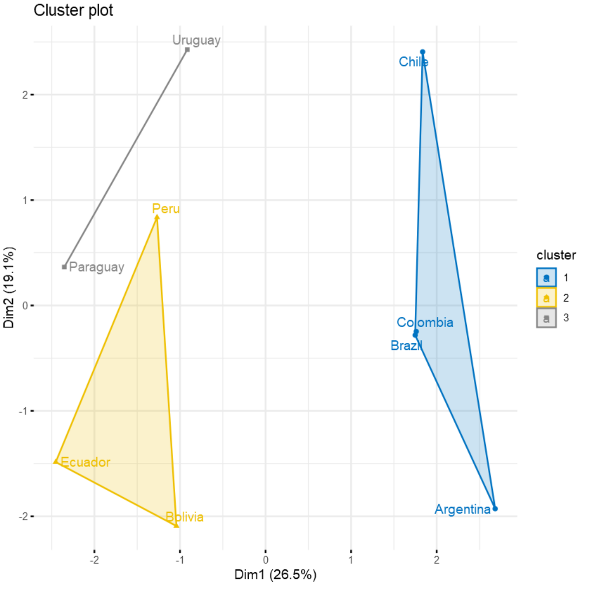


FIGURE 10 FUZZY CLUSTERING FOR SA: INDUSTRIAL VARIABLES



7. Concluding remarks

The results of the different clustering methods for both groups of indicators (*nominal* and *industrial*) shows that the degree of convergence of the SA countries coincides with the main regional integration processes. Despite that the clusters using *industrial* indicators show unclear patterns, the most relevant conclusion (according to the clustering of the *nominal* indicators) is that we were able to identify a group of SA countries, which generally share the same cluster, that are in the best position for a hypothetical monetary integration: Chile, Colombia, and Peru. These economies have greater macroeconomic stability—low inflation, debt-to-GDP levels below 50%, low fiscal deficits, and exchange rate stability—and they would best adjust to the requirements to adopt a common currency, that is, within SA region, these three countries have better convergence in *nominal* terms compared to the rest of the countries. Additionally, these economies belong to the Pacific Alliance, and Colombia and Peru are part of the Andean Community. In addition, according to *industrial* indicators, Chile and Colombia maintain similar productive patterns. However, a negative aspect is that this cluster exhibits a low level of intraregional trade; therefore, the benefits of adopting a common currency would be limited. To a lesser extent, Ecuador could be another candidate country to form a possible monetary union in SA. Ecuador has low levels of inflation, debt of less than 60% of GDP, and absence of volatility in the nominal exchange rate as it is a dollarized economy; and, according to *industrial* indicators, it has similarities with Peru. On the other hand, although Argentina and Brazil (MERCOSUR members) have some industrial convergence, the macroeconomic patterns of these countries are not suitable for a monetary integration process.

References

Alesina, A., Barro, R., Tenreyro, S., 2002. Optimal Currency Areas. Natl Bur Econ Res Work Pap Ser No. 9072. Https://Doi.Org/10.3386/W9072

Aminian, N., Fung, K.C., Ng, F., 2009. A Comparative Analysis of Trade and Economic Integration in East Asia and Latin America. Econ Chang Restruct 42, 105–137. Https://Doi.Org/10.1007/S10644-008-9059-Z

Artis, M., Zhang, W., 1997. On Identifying the Core Of EMU: An Exploration of Some Empirical Criteria, CEPR Discussion Papers. C.E.P.R. Discussion Papers.

Basnet, H.C., Pradhan, G., 2017. Regional Economic Integration in Mercosur: the Role of Real and Financial Sectors. Rev Dev Financ 7, 107–119. Https://Doi.Org/10.1016/J.Rdf.2017.05.001

Basnet, H.C., Sharma, S.C., 2013. Economic Integration in Latin America. J Econ Integr 28, 551–579. Https://Doi.Org/10.11130/Jei.2013.28.4.551

Bayoumi, T., Eichengreen, B., 1994. One Money or Many? Analysing the Prospects for Monetary Unification in Various Parts of The World, Princeton

Studies in International Economics. International Economics Section, Departement of Economics Princeton University,. Https://Doi.Org/DOI:

Bénassy-Quéré, A., Coupet, M., 2005. On the Adequacy of Monetary Arrangements in Sub-Saharan Africa. World Econ 28, 349–373. Https://Doi.Org/10.1111/J.1467-9701.2005.00649.X

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms, Pattern Recognition with Fuzzy Objective Function Algorithms. Https://Doi.Org/10.1007/978-1-4757-0450-1

Blackman, C.R., 1998. Convergence Between Telecommunications and Other Media: How Should Regulation Adapt? Telecomm Policy 22, 163–170.

Boreiko, D., 2003. EMU And Accession Countries: Fuzzy Cluster Analysis of Membership. Int J Financ Econ 8, 309–325. Https://Doi.Org/10.1002/Ijfe.216

Borsi, M.T., Metiu, N., 2015. The Evolution of Economic Convergence in the European Union. Empir Econ 48, 657–681. Https://Doi.Org/10.1007/S00181-014-0801-2

Bresser-Pereira, L.C., Holland, M., 2009. Common Currency and Economic Integration in Mercosul. J Post Keynes Econ 32, 213–234. Https://Doi.Org/10.2753/PKE0160-3477320206

Bröring, S., Leker, J., 2007. Industry Convergence and its Implications for the Front End of Innovation: A Problem of Absorptive Capacity. Creat Innov Manag 16, 165–175. Https://Doi.Org/Https://Doi.Org/10.1111/J.1467-8691.2007.00425.X

Cavallaro, E., Villani, I., 2021. Club Convergence in EU Countries. J Econ Integr 36, 125–161.

Celebi, M.E., Kingravi, H.A., Vela, P.A., 2013. A Comparative Study of Efficient Initialization Methods for the K-Means Clustering Algorithm. Expert Syst Appl 40, 200–210. Https://Doi.Org/Https://Doi.Org/10.1016/J.Eswa.2012.07.021

Cieślik, A., Wciślik, D.R., 2020. Convergence Among the CEE-8 Economies and their Catch-Up Towards the EU-15. Struct Chang Econ Dyn 55, 39–48. Https://Doi.Org/Https://Doi.Org/10.1016/J.Strueco.2020.07.006

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd Ed. Ed. Routledge. Https://Doi.Org/10.4324/9780203771587

De Grauwe, P., 2016. Economics of Monetary Union, 11th Ed. Oxford University Press.

Divisekara, R.W., Jayasinghe, G.J.M.S.R., Kumari, K.W.S.N., 2020. Forecasting the Red Lentils Commodity Market Price Using SARIMA Models. SN Bus Econ 1, 20. Https://Doi.Org/10.1007/S43546-020-00020-X

Dorrucci, E., Firpo, S., Fratzscher, M., Mongelli, F.P., 2004. The Link Between Institutional and Economic Integration: Insights for Latin America from the European Experience. Open Econ Rev 15, 239–260. Https://Doi.Org/10.1023/B:OPEN.0000037699.34047.Be

Dunn, J.C., 1973. A Fuzzy Relative of the ISODATA Process and its Use in Detecting Compact Well-Separated Clusters. J Cybern 3, 32–57. Https://Doi.Org/10.1080/01969727308546046

Dutta, T., Rawat, A., Mishra, A., 2020. Latin America: Problems and Opportunities of Integration. Pp. 241–255. Https://Doi.Org/10.4018/978-1-7998-1730-7.Ch013

Eichengreen, B., 1998. Does Mercosur Need a Single Currency, National Bureau of Economic Research Working Paper Series. Cambridge, MA. Https://Doi.Org/10.3386/W6821

Frankel, J.A., Rose, A.K., 1997. Is EMU More Justifiable Ex Post than Ex Ante? Eur Econ Rev 41, 753–760. Https://Doi.Org/10.1016/S0014-2921(97)00034-2

Hafner, K.A., Kampe, L., 2018. Monetary Union in Latin America: An Assessment in the Context Of Optimum Currency Area. Appl Econ 50, 5672–5697. Https://Doi.Org/10.1080/00036846.2018.1489116

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. J R Stat Soc Ser C (Applied Stat 28, 100–108. Https://Doi.Org/10.2307/2346830

Hochreiter, E., Schmidt-Hebbel, K., Winckler, G., 2002. Monetary Union: European Lessons, Latin American Prospects. North Am J Econ Financ 13, 297–321. Https://Doi.Org/10.1016/S1062-9408(02)00097-9

Hochreiter, E., Siklos, P.L., 2002. Alternative Exchange-Rate Regimes: The Options for Latin America. North Am J Econ Financ 13, 195–211. Https://Doi.Org/10.1016/S1062-9408(02)00099-2

Issiaka, C., Gnimassoun, B., 2013. Optimality of a Monetary Union: New Evidence from Exchange Rate Misalignments in West Africa. Econ Model 32, 463–482. Https://Doi.Org/10.1016/J.Econmod.2013.02.038

Jain, A.K., 2010. Data Clustering: 50 Years Beyond K-Means. Pattern Recognit Lett 31, 651–666. Https://Doi.Org/Https://Doi.Org/10.1016/J.Patrec.2009.09.011

Kassambara, A., 2017. Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning, Multivariate Analysis. STHDA.

Kaufman, L., Rousseeuw, P.J., 1990. Finding Groups in Data: An Introduction to Cluster Analysis (Wiley Series In Probability And Statistics), Eepe.Ethz.Ch. Https://Doi.Org/10.1007/S13398-014-0173-7.2

Kaufman, L., Rousseeuw, P.J., 1987. Clustering by Means of Medoids. Stat Data Anal Based L1-Norm Relat Methods.

Kenen, P., 1969. The Theory Of Optimum Currency Areas: An Eclectic View, in: Mundell, R. And Swoboda., A. (Ed.), Monetary Problems Of The International Economy. University Of Chicago Press, Chicago, Pp. 41–60.

Kočenda, E., Kutan, A.M., Yigit, T.M., 2008. Fiscal Convergence In The European Union. North Am J Econ Financ 19, 319–330. Https://Doi.Org/Https://Doi.Org/10.1016/J.Najef.2008.06.003

Kopits, G., 2002. Central European EU Accession and Latin American Integration: Mutual Lessons in Macroeconomic Policy Design. North Am J Econ Financ 13, 253–277. Https://Doi.Org/10.1016/S1062-9408(02)00092-X

Larrain, F., Tavares, J., 2003. Regional Currencies Versus Dollarization: Options for Asia and the Americas. J Policy Reform 6, 35–49. Https://Doi.Org/10.2139/Ssrn.398061

Licandro, G., 2000. Is Mercosur an Optimal Currency Area ? A Shock Correlation Perspective A Shock Correlation Perspective, 004–2000. Montevideo.

Macqueen, J., 1967. Some Methods for Classification and Analysis of Multivariate Observations, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, Fifth Berkeley Symposium on Mathematical Statistics and Probability. University Of California Press, Berkeley, Calif., Pp. 281–297.

Márquez-Ramos, L., Florensa, L.M., Recalde, M.L., 2017. Understanding the Determinants of Economic Integration in Latin America. J Econ Integr 32, 558–585. Https://Doi.Org/10.11130/Jei.2017.32.3.558

Mckinnon, R.I., 1963. Optimum Currency Areas. Am Econ Rev 53, 717–725.

Mundell, R.A., 1961. A Theory of Optimum Currency Areas. Am Econ Rev Econ Rev 51, 657–665.

Ng, R.T., Han, J., 2002. CLARANS: A Method for Clustering Objects for Spatial Data Mining. IEEE Trans Knowl Data Eng 14, 1003–1016. Https://Doi.Org/10.1109/TKDE.2002.1033770

Obstfeld, M., Rogoff, K., 1996. Foundations of International Macroeconomics. MIT Press, Cambridge.

Padilla, L., Rodriguez García-Brazales, Á., 2021. Can South America form an Optimal Monetary Area? A Structural Vector Autoregression Analysis. Int Econ Econ Policy 18, 309–329. Https://Doi.Org/10.1007/S10368-020-00490-2

Pegkas, P., Staikouras, C., Tsamadias, C., 2020. On the Determinants of Economic Growth: Empirical Evidence from the Eurozone Countries. Int Area Stud Rev 23, 210–229. Https://Doi.Org/10.1177/2233865920912588

Regmi, K., Nikolsko-Rzhevskyy, A., Thornton, R., 2015. To Be Or Not To Be: An Optimum Currency Area for South Asia? J Policy Model 37, 930–944. Https://Doi.Org/10.1016/J.Jpolmod.2015.09.005

Sambandam, R., 2003. Cluster Analysis Gets Complicated. Mark Res.

Sancho, R., 2007. Innovación Industrial. Rev Española Doc Científica 30.

Segura, J., 1993. Teoría de la economía industrial. Rev econ apl número.

Sokal, R.R., Rohlf, F.J., 1962. The Comparison of Dendrograms by Objective Methods. Taxon. Https://Doi.Org/10.2307/1217208

Tibshirani, R., Walther, G., Hastie, T., 2001. Estimating the Number of Clusters in a Data Set Via the Gap Statistic. J R Stat Soc Ser B Stat Methodol. Https://Doi.Org/10.1111/1467-9868.00293

Tsangarides, C.G., Qureshi, M.S., 2008. Monetary Union Membership in West Africa: A Cluster Analysis. World Dev 36, 1261–1279. Https://Doi.Org/10.1016/J.Worlddev.2007.06.019

Visser, H., 2004. A Guide to International Monetary Economics.Exchange Rate Theories, Systems and Policies., Third Edit. Ed, Books. Edward Elgar Publishing.

Annexes

Table 13. Data sources and periodicity

Variable	Sources	Period
Nominal and/or traditional variables	SA countries	
(i) Inflation (end of period consumer prices)	IMF Outlook	2001-2017
(ii) Government balance (general government net lending/ borrowing)	IMF Outlook	2001-2017
(iii) Debt (general government gross debt)	IMF Outlook	2001-2017
(iv) Monetary policy rate	CEPAL Statistics	2001-2017
(v) Nominal exchange rate variation	CEPAL Statistics	2001-2017
(vi) Business cycle correlations	CEPAL Statistics	2001-2017
(vii) Regional trade intensity	CEPAL Statistics	2001-2017
(viii) Work flexibility*	World Economic Forum (GCI)	2006-2017
Industrial Indicators		
(a) Concentration*	CEPAL Statistics	2002-2017
(b) Innovation*	World Economic Forum (GCI)	2006-2017
(c) Market size*	World Economic Forum (GCI)	2006-2017
(d) Competition*	World Economic Forum (GCI)	2006-2017
(e) Real exchange rate variation	World Bank (Databank)	2001-2017
(f) Productivity growth	World Bank (Databank)	2001-2017
(g) Energy consumption of the industry	U.S. Energy Information Administration (EIA)	2001-2017

Note: * The imputation method used was SARIMA.