risks

MDPI

Article

Turning Points in the Core–Periphery Displacement of Systemic Risk in the Eurozone: Constrained Weighted Compositional Clustering

Anna Maria Fiori [1,*,†] and Germà Coenders [2,*]

[1] Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano, Italy
[2] Department of Economics, Faculty of Economics and Management, University of Girona, C/Universitat 10, 17003 Girona, Spain
* Correspondence: anna.fiori@unimib.it (A.M.F.); germa.coenders@udg.edu (G.C.)
† Member of Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA), Italy.

Abstract: Investigating how systemic risk originates and spreads across the financial system poses an inherently compositional question, i.e., a question concerning the joint distribution of relative risk share across several interdependent contributors. To address this question, we propose a weighted compositional clustering approach aimed at tackling the trajectories and turning points of systemic risk in the Eurozone, from both a chronological and a geographical perspective. The cluster profiles emerging from our analysis indicate a progressive shift from Northern Europe towards the Euro-Mediterranean region in the coordinate center of systemic risk compositions. This shift matures as the outcome of complex interactions between core and peripheral EU countries that compositional methods have the merit of capturing and unifying in a self-contained multivariate framework.

Keywords: Compositional Data (CoDa); adjacency-constrained hierarchical clustering; CONISS algorithm; Systemic Risk Index (SRISK); weighted centered log-ratios

check for updates

Academic Editor: Nicu Sprincean

Received: 15 November 2024
Revised: 8 January 2025
Accepted: 20 January 2025
Published: 24 January 2025

Citation: Fiori, Anna Maria, and Germà Coenders. 2025. Turning Points in the Core–Periphery Displacement of Systemic Risk in the Eurozone: Constrained Weighted Compositional Clustering. *Risks* 13: 21. https://doi.org/10.3390/risks13020021

1. Introduction

Systemic risk has been defined as "the risk of disruption to the flow of financial services that (i) is caused by an impairment of all or parts of the financial system; and (ii) has the potential to have serious negative consequences for the real economy" (Financial Stability Board 2009). Due to linkages within the financial system, all types of intermediaries, markets, infrastructures, and even whole countries can be systemically at risk or can potentially become so in specific time periods, depending on the economic environment. As pointed out by Andrieş et al. (2024), risk at the country level is made up of the aggregation of firm-level risks, but country-level solvency also has an impact on firm-level risk. Against this backdrop, a key question relates to where and when potential imbalances between parts of the system occur or, in other words, which parts of the system dominate the overall systemic risk in the face of events such as the European Sovereign Debt Crisis (SDC) or COVID-19.

The European Union is particularly susceptible to systemic risk due to its unique economic and institutional structure, where banks in one country are heavily exposed to sovereign debt in others (see, e.g., Acharya et al. 2022; Foglia et al. 2023) and the lack of a fully integrated fiscal union limits the possibility for member states to address financial shocks in a coordinated manner (Borri and Di Giorgio 2022; Shambaugh 2012). These fragilities have clearly emerged during the SDC, when financial instability in one country quickly spread to others through sovereign debt linkages and interbank exposures, with spillover effects across the entire region (Betz et al. 2016; Bratis et al. 2018; Shahzad et al. 2019). Studies of systemic risk in the Eurozone have often employed network analysis to investigate the critical role of endogenous complexity and trace contagion patterns within the EU's financial framework (see, e.g., Aldasoro and Alves 2018; Paltalidis et al. 2015).

The systemic riskiness of banks and other financial institutions has also been quantified with different methodologies, ranging from the regulatory framework established by the Basel Committee on Banking Supervision (https://www.bis.org/basel_framework, accessed on 13 November 2024) to a variety of academic measures, including—but not limited to—the distressed insurance premium introduced by Huang et al. (2009), the marginal expected shortfall and systemic expected shortfall in Acharya et al. (2012), the Systemic Risk Index (SRISK) from Brownlees and Engle (2017), and the CoVaR in Adrian and Brunnermeier (2016). See Stolbov and Shchepeleva (2024) for a comprehensive study on the evolution of systemic risk research during the period 2007–2021. All these methods have progressively enabled a deeper understanding of shock propagation channels and policy responses in times of systemic crises.

In addition to the variety of mechanisms and measurement tools investigated in the literature, in this article, we show that an important dimension of systemic risk in the Euro area is linked to compositional variation in the relative risk share pertaining to individual member states. This finding is the outcome of a novel application of Compositional Data (CoDa) analysis, in which we propose a weighted chronological and geographical clustering of systemic risk compositions in the Eurozone. In our CoDa analysis, "individuals" (the single member states) are not considered independently but parts of a "whole" (the EU financial system), with a focus on risk proportions—rather than monetary, or absolute, metrics—that distinguishes our article from the vast body of research on the topic.

Despite growing interest in CoDa methods (see, e.g., Coenders et al. 2023; Egozcue and Pawlowsky-Glahn 2019; Greenacre 2021; Greenacre et al. 2023, and references therein), only a few studies have applied them to finance and risk management problems so far. Vega Baquero and Santolino (2022), in particular, constructed a concentration index for the Colombian banking system via compositional analysis to detect potentially "too big to fail" financial entities, while Porro (2022) and Fiori and Porro (2023) proposed the adoption of CoDa visualization tools to explore some regional disparities in the distribution of systemic risk share. Overall, these works indicate a clear potential of CoDa methods to uncover some less known aspects of complex financial phenomena thanks to the use of a relative scale, which prioritizes the information content of the part proportions over their absolute values or sums.

The present article proposes a new framework to tackle the evolution of systemic risk in the Eurozone as a compositional time series, with constituent parts corresponding to the relative risk contributions of individual member states during the period 2009–2022. To the best of our knowledge, this approach has not been used so far in the literature, where the time dynamics of systemic risk have been mainly studied regarding monetary units or scores (see, e.g., Brownlees et al. 2020; Magkonis and Tsopanakis 2020; Stolbov and Shchepeleva 2018, and references therein) . A relevant concern in the systemic risk literature is the so-called overshadowing effect of size, meaning that common indicators are inclined to overstate big financial institutions as risky, while small institutions tend to be overdefined as safe (Varotto and Zhao 2018). As CoDa analysis operates on relative information, it can be expected to reveal shifts and co-movements among larger and smaller risk contributors, which may be masked when focusing solely on absolute metrics (see also the discussion on small parts and their role in a composition by Pawlowsky-Glahn et al. 2015, Example 1.2).

Based on adjacency-constrained Ward clustering with the CONstrained Incremental Sum of Squares (CONISS) algorithm from Grimm (1987), we develop a chronological clustering approach for CoDa that gives new insights into the time trajectories and turning points of systemic risk across the Eurozone. To calibrate the influence of low-mean/high-dispersion parts associated with some peripheral countries (e.g., Ireland or Portugal), we introduce a system of differential weights for the parts as proposed by Greenacre (2018) and used by Jofre-Campuzano and Coenders (2022), and we integrate them into the CONISS algorithm for hierarchical constrained clustering. The outcome of the proposed approach is a structured periodization of relative systemic risk trajectories into five clearly separated time segments, indicating a progressive shift from Northern Europe to the Euro-Mediterranean region in the compositions of systemic risk. We show how this shift matures as a consequence of complex interactions between core and peripheral countries that the CoDa approach has the merit of detecting and unifying in a self-contained multivariate framework. Time series classification with the differential weighting of parts is indeed a key contribution that characterizes the present article and distinguishes it, in particular, from the graphical CoDa analysis of systemic risk presented in Porro (2022) and Fiori and Porro (2023).

A critical dimension of the European Union is the exposure to geographically asymmetric shocks and changes in macro-economic conditions that need to be addressed within a single-currency framework (see, e.g., Shambaugh 2012; and Karimalis and Nomikos 2018). In order to investigate this aspect, we contribute a geographical cluster analysis of systemic risk compositions combining Ward's hierarchical procedure with the same differential weighting of parts. The proposed approach reveals a mix of grouping structure and country-specific behaviors; while the former captures the relative balance between the proportional risk contributions of "similar" subsets of countries, the latter underscore the presence of a few highly volatile parts that tend to drive apart from the rest of the Eurozone.

Whereas a large portion of the literature has focused on the aftermath of the Global Financial Crisis (GFC) and on specific episodes in subsequent years (e.g., SDC, Brexit, the Italian banking crisis in 2016–17, the outbreak of the pandemic in 2020), our compositional time series embraces a longer period—from February 2009 to November 2022—thus enabling a comparative investigation of multiple interrelated crises (GFC, SDC, COVID-19, war in Ukraine) that have challenged the resilience of the EU's financial system. This time extension and the considerable heterogeneity between the systemic risk magnitude of core and peripheral EU countries makes our compositional time series a complex combination of major and minor constituents, with significant disproportions in their scales of variation. We show that the application of weighting methods from CoDa in conjunction with chronological and geographical clustering can be used to identify the time and space trajectories of systemic risk compositions, highlighting the instability threats posed by relative imbalances among ratios of major and minor parts in a complex entity such as the European Union. Our conclusion is that CoDa analysis should be viewed not as a replacement for other well-established methodologies that have been developed and tested in the recent literature but as a complementary tool in the study of the multifaceted nature of systemic risk whenever the research questions concern the distribution of risk among countries or financial institutions.

The rest of this article is organized as follows. In Section 2, we recall the definition of the SRISK that will be considered and we describe the aggregation mechanism underlying the construction of the relative SRISK share for individual countries. Section 3 focuses on compositional methods, starting with a brief introduction (Section 3.1). Then, we present the weighted CoDa clustering procedures for chronological segmentation of a compositional time series and for the geographical aggregation of its parts in Sections 3.2 and 3.3, respectively. The empirical analysis is contained in Section 4 and followed by a discussion of the results (Section 5).

2. Systemic Risk Index

The Systemic Risk Index (SRISK) introduced by Brownlees and Engle (2017) is a monetary indicator of the expected capital shortfall of a financial intermediary, f, at a given time, i, conditional on a systemic event implying a prolonged market decline. In particular,

$$\mathrm{SRISK}_{f,i} = \underbrace{k\Big[D_{f,i} + (1 - \mathrm{LRMES}_{f,i})W_{f,i}\Big]}_{\textit{Required Capital}} - \underbrace{(1 - \mathrm{LRMES}_{f,i})W_{f,i}}_{\textit{Available Capital}}, \tag{1}$$

where k is a prudential capital ratio established by regulatory authorities, $D_{f,i}$ and $W_{f,i}$ denote, respectively, the book value of total liabilities and the market value of equity, and LRMES is the Long-Run Marginal Expected Shortfall, which corresponds to the expected drop in equity value that the firm would experience if the global equity market were to decline by more than 40% within the next six months. This distress scenario is indeed consistent with historical data. In particular, during the dot-com bubble burst and the subprime crisis, the global equity market declined by around 40% (see Engle and Zazzara 2018, and references therein).

As mentioned in the Introduction, a number of well-established metrics for systemic risk analysis are available in the literature. However, a unique feature of the SRISK is that the measure is regularly computed and published by The Volatility Institute (https://vlab.stern.nyu.edu, accessed on 13 November 2024) for more than 1000 financial firms in the world and by the Center for Risk Management at the University of Lausanne (http://www.crml.ch/systemic-risks, accessed on 13 November 2024) for European institutions (see also Engle et al. 2014, for a detailed description of the methodology). Additionally, and more importantly for the present study, the individual measure (1) can be aggregated across all the undercapitalized firms in a country, yielding a nationwide SRISK measure:

$$\mathrm{SRISK}_i^{(j)} = \sum_{f=1}^{N_j} \max\{\mathrm{SRISK}_{f,i}, 0\}, \tag{2}$$

which represents the total amount of capital needed to bail out the national financial system in a country, j, consisting of N_j institutions, conditional on a crisis scenario. The ratios of country-level measures (2) to the overall SRISK of a reference financial system represent the risk share associated with each of the countries (parts), information that can be properly processed by the compositional methods that we describe in the next section.

It is worth mentioning that the contributions of overcapitalized firms (i.e., those with a negative SRISK) are not considered in (2) because it is unlikely that surplus capital would be easily mobilized from prudent to weak firms in the event of a downturn (Brownlees and Engle 2017; Engle 2018).

3. Compositional Analysis

CoDa are a special type of multivariate observations with the peculiarity that relative rather than absolute information is relevant to the analysis (Filzmoser et al. 2018; Pawlowsky-Glahn et al. 2015). A specific methodology for CoDa began to emerge in the 1980s, when the pioneering work of John Aitchison was systematized in the monograph *The Statistical Analysis of Compositional Data* (Aitchison 1986). Since then, the CoDa approach has made substantial progress (see Bacon-Shone 2011, for a historical retrospective) and is currently regarded as a preferred modeling choice to uncover the relative structure of multivariate datasets, prioritizing the information content of ratios between their constituent parts.

3.1. Preliminaries of CoDa Modeling

Let $\mathbb{R}^J$ denote the set of all ordered J-tuples of real numbers and $\mathbb{R}^J_+$ the set of J-tuples with strictly positive elements. Following Egozcue and Pawlowsky-Glahn (2019), the general concept of CoDa can be formalized as follows.

Definition 1. *A compositional vector with J parts, $\mathbf{x} = [x_1, \ldots, x_J]$, is a J-dimensional (row) vector with strictly positive elements that carry relative information. Two vectors, $\mathbf{x}$ and $\mathbf{y} \in \mathbb{R}^J_+$, are compositionally equivalent if there exists a positive constant, λ, such that $\mathbf{y} = \lambda\mathbf{x}$. A J-part composition is therefore an equivalence class of proportional vectors with positive elements.*

The *closure* operation,

$$\mathcal{C}(\mathbf{x}) = \left[\frac{x_1}{\sum_{j=1}^J x_j}, \ldots, \frac{x_J}{\sum_{j=1}^J x_j}\right], \tag{3}$$

converts a compositional vector from $\mathbb{R}^J_+$ into its representation in the J-part simplex,

$$\mathcal{S}^J = \left\{\mathbf{x} = [x_1, \ldots, x_J] \in \mathbb{R}^J_+ : \sum_{j=1}^J x_j = 1\right\}, \tag{4}$$

traditionally regarded as a natural sample space for Compositional Data. Log-transformed ratios of parts (*log-ratios*) are fundamental to map Compositional Data to real vectors for which the usual Euclidean structure is suitable (e.g., Pawlowsky-Glahn et al. 2015). Among various families of log-ratios, the following is particularly useful in our analysis.

Definition 2. *The centered log-ratio (clr) of a J-part composition, $\mathbf{x} = [x_1, \ldots, x_J]$, is a vector in $\mathbb{R}^J$ defined by*

$$\mathbf{v} = clr(\mathbf{x}) = [clr_1(\mathbf{x}), \ldots, clr_J(\mathbf{x})] = \left[\log\frac{x_1}{g(\mathbf{x})}, \ldots, \log\frac{x_J}{g(\mathbf{x})}\right], \tag{5}$$

where $g(\cdot)$ denotes the geometric mean of parts:

$$g(\mathbf{x}) = \prod_{j=1}^J x_j^{1/J} = \exp\left(\frac{1}{J}\sum_{j=1}^J \log x_j\right). \tag{6}$$

This transformation establishes an isometry between $\mathcal{S}^J$ and the $(J-1)$-dimensional subspace $\mathcal{V} = \{\mathbf{v} \in \mathbb{R}^J : v_1 + \ldots + v_J = 0\}$, enabling the computation of compositional distances through standard Euclidean distances between *clr*-vectors.

Definition 3. *Let $\mathbf{x} = [x_1, \ldots, x_J]$ and $\mathbf{x}' = [x'_1, \ldots, x'_J]$ denote two compositional vectors in S^J. The Aitchison distance between $\mathbf{x}$ and $\mathbf{x}'$ is the Euclidean distance between the clr-transformed compositions:*

$$d_a(\mathbf{x}, \mathbf{x}') = d_e(clr(\mathbf{x}), clr(\mathbf{x}')) = \sqrt{\sum_{j=1}^J [clr_j(\mathbf{x}) - clr_j(\mathbf{x}')]^2}. \tag{7}$$

3.2. Weighted Chronological Clustering

In this section, we propose a new combination of hierarchical clustering and weighting methods from CoDa analysis that will be instrumental in identifying relevant time trajectories and turning points of systemic risk in the Euro area. The special nature of CoDa and the principles of Aitchison's geometry on the simplex (see Pawlowsky-Glahn and Egozcue 2001) guide our formulation of the weighted chronological clustering algorithm.

Chronological clustering generally refers to the segmentation of a time series into a sequence of consecutive sub-periods (clusters), which are intended to detect significant change points in the evolution of the underlying phenomenon. Unlike ordinary cluster analysis, this problem requires that only time-adjacent samples are considered for merging, provided that they fulfill a reasonable similarity criterion with other observations in the same group. The task can be performed by adjacency-constrained Ward clustering, using the CONstrained Incremental Sum of Squares (CONISS) algorithm, originally introduced for the stratigraphical classification of zones in the geosciences (Grimm 1987; see also Di Donato et al. 2009). The method is agglomerative and hierarchical, joining elements in accordance with Ward's criterion of the minimization of the total within-cluster dispersion while taking into account contiguity information along the series.

We propose here an implementation of adjacency-constrained clustering to the compositional time series of the systemic risk share $X_1, \ldots, X_J$ (corresponding to J different countries), additionally allowing for the differential weighting of parts associated with minor and major contributors. A compositional time series with J parts, $X_1, \ldots, X_J$, is usually represented as a matrix, $\mathbf{X}_{(I \times J)}$, where each row, $\mathbf{x}_i = [x_{i1}, \ldots, x_{iJ}] \in \mathcal{S}^J$, is a (closed) CoDa sample of J parts observed at a specific time point, $i = 1, \ldots I$. The CoDa literature has repeatedly reported imbalance in the contribution of parts when some very large parts coexist with very small ones, the latter tending to have a disproportionate effect on the clustering solution. Unlike standard cluster analysis, compositional cluster analysis does not make variable standardization possible, but column weighting can be used instead. As recommended by Greenacre (2018, 2021); Greenacre and Lewi (2009), we introduce a weighting scheme based on the arithmetic means of the closed parts:

$$w_j = \frac{1}{I} \sum_{i=1}^{I} x_{ij}, \quad \text{for } j = 1, \ldots, J, \quad \text{with } \sum_{j=1}^{J} w_j = 1. \tag{8}$$

Following Jofre-Campuzano and Coenders (2022) and Dao et al. (2024), we adopt the following definition of weighted *clr*-transformation:

Definition 4. *Consider a compositional time series, $\mathbf{X}_{(I \times J)}$, and a set of positive real numbers, w_j, for $j = 1, \ldots, J$, such that $\sum_{j=1}^{J} w_j = 1$. For each row, $\mathbf{x}_i = [x_{i1}, \ldots, x_{iJ}]$, of $\mathbf{X}$, the part-weighted (or column-weighted) clr-coefficients are defined as*

$$\begin{aligned} clr^{(w)}(\mathbf{x}_i) &= [clr_1^{(w)}(\mathbf{x}_i), \ldots, clr_J^{(w)}(\mathbf{x}_i)] \\ &= \left[\sqrt{w_1} \log \frac{x_{i1}}{g^{(w)}(\mathbf{x}_i)}, \ldots, \sqrt{w_J} \log \frac{x_{iJ}}{g^{(w)}(\mathbf{x}_i)}\right], \end{aligned} \tag{9}$$

where $g^{(w)}(\mathbf{x}_i) = \prod_{j=1}^{J} x_{ij}^{w_j}$ is the weighted geometric mean of parts in $\mathbf{x_i}$, for $i = 1, \ldots, I$.

Other weighting schemes are possible, for instance, based on inverse *clr* variances (Hron et al. 2017). These weights were compared in a financial context with those suggested in Greenacre (2018, 2021); Greenacre and Lewi (2009) by Jofre-Campuzano and Coenders (2022) and no significant differences were found in the analysis results.

The definition of Aitchison's distance (7) between compositional vectors is easily modified to account for variable weights.

Definition 5. *The column-weighted compositional distance between two samples,* $\mathbf{x}_i, \mathbf{x}'_i$*, in a compositional time series,* $\mathbf{X}_{(I \times J)}$*, is defined as the Euclidean distance between their respective weighted clr-transformations:*

$$d_a^{(w)}(\mathbf{x}_i, \mathbf{x}_{i'}) = d_e\left(clr^{(w)}(\mathbf{x}_i), clr^{(w)}(\mathbf{x}_{i'})\right) = \sqrt{\sum_{j=1}^{J}\left[clr_j^{(w)}(\mathbf{x}_i) - clr_j^{(w)}(\mathbf{x}_{i'})\right]^2}. \tag{10}$$

We introduce the following version of the CONISS algorithm for a compositional time series with weighted parts. The procedure begins with each CoDa sample in a separate cluster and a dissimilarity matrix,

$$D = \left\{d(i, i') = d_a^{(w)}(\mathbf{x}_i, \mathbf{x}_{i'}) : i, i' = 1, \ldots, I\right\}, \tag{11}$$

consisting of weighted compositional distances (10) between all CoDa samples (rows) in $\mathbf{X}$. The matrix (11) is searched for a minimum between adjacent periods, say, $d(p, q)$, and the corresponding samples are merged into a new cluster, $C_{\{p,q\}}$. This reduces the total number of clusters by one, and all dissimilarities involving either the element p or q have to be updated. In each successive stage, a general scheme for evaluating the dissimilarity between a newly formed cluster, $C_{\{p,q\}}$, and any other cluster, C_r, is given by the updating equation:

$$d(C_r, C_{\{p,q\}}) = \frac{(n_r + n_p)d(C_r, C_p) + (n_r + n_q)d(C_r, C_q) - n_r d(C_p, C_q)}{n_r + n_p + n_q}, \tag{12}$$

where n_k is the number of elements in C_k, for $k = r, p, q$ (see Grimm 1987, for details). The process is iterated until all samples are found in a unique group of the size I.

Whereas conventional (unconstrained) cluster analysis searches the entire dissimilarity matrix for a minimum in each stage, the chronological CONISS method allows only time-adjacent elements to be considered for joining. The final outcome is a sequence of hierarchically nested partitions, which can be displayed in the well-known form of a clustering tree (dendrogram) preserving the natural order of samples in the compositional time series.

3.3. Weighted Clustering of Parts

A critical dimension of the European Union is exposure to geographically asymmetric shocks, which cause significant imbalances in systemic risk proportions across different countries. We here propose a weighted compositional clustering of parts in order to investigate the existence of grouping structures and country-specific behaviors in the multivariate distribution of the systemic risk share X_j, $j = 1, \ldots, J$, associated with EU member states. This clustering problem concerns the columns of a compositional time series, $\mathbf{X}_{(I \times J)}$. This involves transposing the data matrix. For coherence, the weights in Section 3.2 have to be applied to the rows of the transposed matrix and the definitions of the *clr*-coefficients and compositional distance need to be reformulated accordingly.

Definition 6. *Consider a CoDa matrix,* $\mathbf{X}_{(I \times J)}$*, and its transpose,* $\mathbf{Y} = \mathbf{X}^{(t)}$*, where each row,* $\mathbf{y}_j = [y_{j1}, \ldots, y_{jI}]$*, contains I observations of the j-th part* X_j*, for* $j = 1, \ldots, J$*.*

- *The columnwise clr-coefficients of* $\mathbf{X}$ *are defined as*

$$clr^{(t)}(\mathbf{y}_j) = [clr_1^{(t)}(\mathbf{y}_j), \ldots, clr_I^{(t)}(\mathbf{y}_j)] = \left[\log \frac{y_{j1}}{g^{(t)}(\mathbf{y}_j)}, \ldots, \log \frac{y_{jI}}{g^{(t)}(\mathbf{y}_j)}\right], \quad (13)$$

where $g^{(t)}(\cdot)$ *denotes the geometric mean of samples:*

$$g^{(t)}(\mathbf{y}_j) = \prod_{i=1}^{I} y_{ji}^{1/I} = \exp\left(\frac{1}{I}\sum_{i=1}^{I} \log y_{ji}\right). \quad (14)$$

- *The compositional distance between any two parts in* $\mathbf{X}$ *can be defined as the Euclidean distance between their columnwise clr-coefficients:*

$$d_a(\mathbf{y_j}, \mathbf{y_{j'}}) = d_e\left(clr^{(t)}(\mathbf{y_j}), clr^{(t)}(\mathbf{y_{j'}})\right) = \sqrt{\sum_{i=1}^{I}\left[clr_i^{(t)}(\mathbf{y}_j) - clr_i^{(t)}(\mathbf{y}_{j'})\right]^2}. \quad (15)$$

The compositional clustering of parts can now be performed with the agglomerative hierarchical procedure described in Greenacre (2018), which takes differential weighting of variables into account during the aggregation process. The initial input to the clustering algorithm is the colunmwise dissimilarity matrix:

$$D^{(t)} = \left\{d_a(\mathbf{y_j}, \mathbf{y_{j'}}), j, j' = 1, \ldots, J\right\}, \quad (16)$$

where (15) is used to compute the distance between pairs of variables, X_j and $X_{j'}$, in the original CoDa set, $\mathbf{X}$, now associated with the columnwise *clr*-vectors $\mathbf{y_j}$ and $\mathbf{y_{j'}}$ in the transposed matrix $\mathbf{Y}$. The partitioning process begins by joining the two variables X_p and X_q that minimize the weighted Ward criterion:

$$\frac{w_j w_{j'}}{w_j + w_{j'}} d_a^2(\mathbf{y_j}, \mathbf{y_{j'}}), \quad (17)$$

where j and j' are any two columns in the CoDa matrix and w_j and $w_{j'} > 0$ are the respective weights, with $\sum_{j=1}^{J} w_j = 1$. The resulting group, $C_{\{p,q\}}$, appears as a new column variable with *clr*-coefficients,

$$\frac{w_p clr^{(t)}(\mathbf{y}_p) + w_q clr^{(t)}(\mathbf{y}_q)}{w_p + w_q}, \quad (18)$$

and the associated weight $w_{\{p,q\}} = w_p + w_q$. The dissimilarity matrix, $D^{(t)}$, is updated by recomputing the columnwise compositional distances (15), and the process is iterated by searching, in each stage, for the minimum of (17), until all elements are grouped into one all-encompassing cluster. Further insights into the hierarchical clustering of variables for CoDa and relationships to the particular geometry of the simplex can be found in Martín-Fernández et al. (2024), who do not differentially weight the points.

4. Empirical Analysis

Based on the methodology described in the previous Section, we investigate the compositional structure of systemic risk in the Euro area using country-level SRISK measures for each of the founding member states, namely, Austria (AT), Belgium (BE), France (FR), Germany (DE), Ireland (IE), Italy (IT), the Netherlands (NL), Portugal (PT), and Spain (ES). Greece (EL), which joined the Euro area in January 2001, is also included in our dataset in view of the key role it played in the so-called 'diabolic loop', as Brunnermeier et al. (2016) labeled the dangerous nexus between sovereign and bank credit risk that characterized the escalation of the SDC in the Euro area. Finland and Luxembourg, in spite of being initial Eurozone members, are excluded due to limited data availability.

From the website of the Center for Financial Risk Management at Lausanne (2023) (http://www.crml.ch/systemic-risks, access on 13 November 2024), we collected monthly SRISK measures covering the period from February 2009 to November 2022. The website includes a specific section with the latest systemic risk measures for individual institutions and countries in Europe. By selecting a given country, it is possible to visualize the temporal evolution of the corresponding variables, including numerical values of monthly SRISK measures expressed in billions of EUR. This information was used to define a compositional time series of $J = 10$ parts (countries) with $I = 166$ monthly observations, represented in the CoDa matrix $\mathbf{X}_{(I\times J)}$ where each row,

$$\mathbf{x}_i = [x_{i1}, \ldots, x_{ij}, \ldots, x_{iJ}], \qquad \text{for} \quad i = 1, \ldots, I, \tag{19}$$

is a (closed) compositional sample with J variables, corresponding to the SRISK share of Eurozone countries:

$$x_{ij} = \frac{\mathrm{SRISK}_i^{(j)}}{\sum_{j=1}^{J} \mathrm{SRISK}_i^{(j)}}, \qquad \text{for} \quad j = 1, \ldots J. \tag{20}$$

Figure 1 shows the time dynamics of country-level SRISK measures both in their absolute scale (billions of EUR, top panel) and in percentage compositions (bottom panel). The compositional display highlights a peculiarity of the group formed by Greece–Ireland–Italy–Portugal–Spain (GIIPS), which was at the epicenter of the European financial crisis. Although the SRISK share of these countries appears remarkably lower in comparison with core Euro members like Germany and France, they are characterized by higher volatility with frequent spikes. This combination underscores the relevance of compositional variability in uncovering the threats of potential instability posed by smaller SRISK contributors within the Euro area.

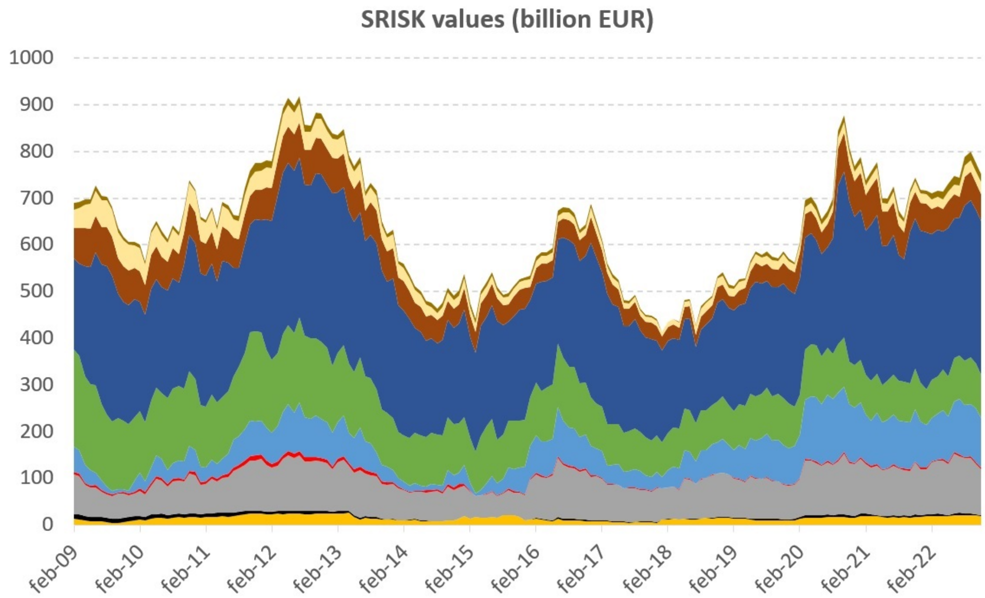

Figure 1. *Cont*.

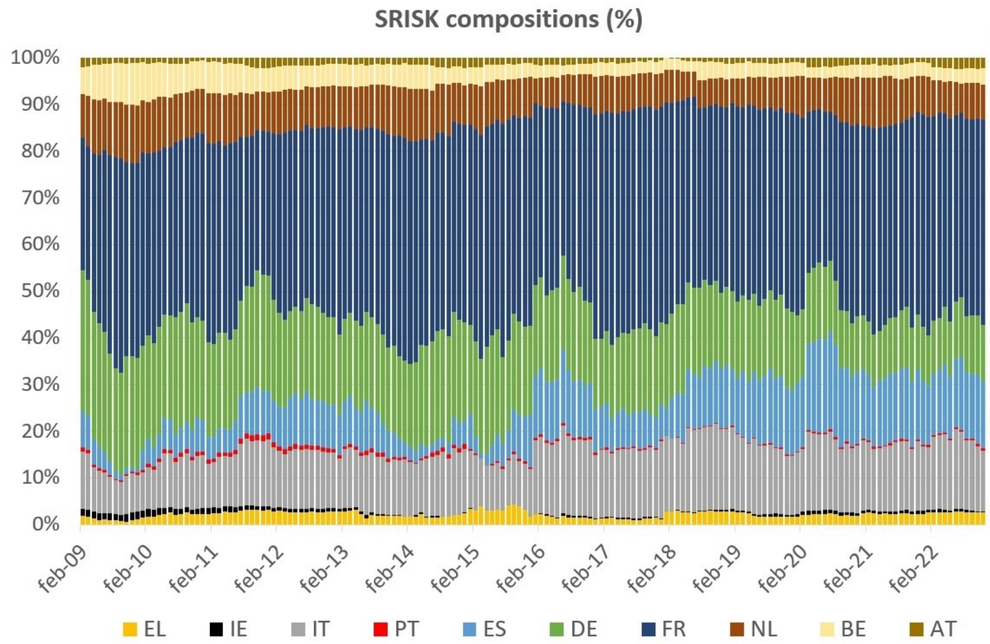

Figure 1. Time series of monthly SRISK in Eurozone countries; monetary values, in billions of EUR (**top**), and percentage compositions (**bottom**). Country codes: AT (Austria), BE (Belgium), DE (Germany), EL (Greece), ES (Spain), FR (France), IE (Ireland), IT (Italy), NL (Netherlands), and PT (Portugal).

4.1. Data Pre-Processing and Weighting

The compositional time series contained 14 samples (8.4% of the total) that had zero values in the SRISK part pertaining to Ireland and thus were not directly amenable to the log-ratio transformations required for CoDa analysis. Following Palarea-Albaladejo and Martín-Fernández (2008), we implemented a zero replacement strategy using the Expectation–Maximization (EM) imputation algorithm. Zero values were not missing at random and, accordingly, the algorithm constrained replaced values to be below a certain limit. We specified the limit as 0.002473%, corresponding to the minimum value of positive SRISK parts observed for Ireland. The EM algorithm adapted to Compositional Data is available in the Rpackage `zCompositions` (Palarea-Albaladejo and Martín-Fernández 2015).

Table 1 summarizes the main descriptive statistics for the compositional time series of the SRISK share. In particular, the "mean" column shows the arithmetic means of parts, which are the basic ingredients of the weighting scheme described in (8). The reference measure of central tendency for a CoDa set is the closed geometric mean, calculated according to (14) and reported in the "center" column in Table 1. France, Germany, and Italy were the main contributors to the SRISK over the sample period, with respective center coordinates of 0.4117, 0.1840, and 0.1329 relative to the total SRISK of the Eurozone. Time discrepancies with respect to the center will be considered in the next section to characterize the compositional profiles of chronological clusters. The minimum, maximum, and quartile statistics, also reported in Table 1, suggest considerable variation in the CoDa set. The interquartile range (IQR = Q_3-Q_1) is particularly high for Germany, Spain, and France.

Using the R package `easyCODA` (Greenacre 2018), we estimated the global dispersion of the CoDa set by its total log-ratio variance,

$$totvar(\mathbf{X}) = \sum_{j=1}^{J} var\left[clr_j^{(w)}(\mathbf{X})\right], \tag{21}$$

assuming first an equally weighted scheme, $w_j = 1/J$, for all $j = 1, \dots, J$ and then using the set of differential weights specified in (8). As shown in Table 2, the equally weighted *clr*-variances and their percentage contributions to $totvar(\mathbf{X})$ are largely dominated by Ireland, which combines near-zero SRISK proportions in several years with high dispersion relative to all other parts. This situation is indeed common in complex CoDa sets including major and minor elements, where the variation in the latter is amplified by the relative scale of the data (Greenacre et al. 2023). The analysis with unequal weights resulted in a more balanced decomposition of $totvar(\mathbf{X})$, attenuating the *clr*-variance contribution of Ireland from 63.71% to 21.14% and underscoring the relevance of Spain with a weighted *clr*-variance contribution of 37.69%. The contributions of Germany and France were also rescaled in accordance with their relative SRISK magnitude, respectively, from 2.60% and 2.39% (unweighted) to 10.80% and 5.74% (weighted). The adoption of a differential weighting scheme for parts thus appeared to be a reasonable choice to describe the interplay between larger and smaller SRISK contributors while simultaneously controlling for possible distortion effects induced by extreme variation in the rarer parts (cf. Dao et al. 2024; Jofre-Campuzano and Coenders 2022).

Table 1. Descriptive statistics for the compositional time series of the SRISK share. Country codes: AT (Austria), BE (Belgium), DE (Germany), EL (Greece), ES (Spain), FR (France), IE (Ireland), IT (Italy), NL (the Netherlands), and PT (Portugal).

Country	Mean	Center	Min.	Q_1	Median	Q_3	Max.	IQR
AT	0.0138	0.0131	0.0001	0.0107	0.0137	0.0161	0.0243	0.0054
BE	0.0413	0.0395	0.0191	0.0280	0.0335	0.0517	0.0899	0.0237
DE	0.1840	0.1859	0.1070	0.1635	0.1864	0.2057	0.3010	0.0422
EL	0.0225	0.0221	0.0064	0.0171	0.0234	0.0271	0.0434	0.010
ES	0.0965	0.0849	0.0079	0.0584	0.0951	0.1355	0.2088	0.0771
FR	0.4117	0.4243	0.2852	0.3908	0.4136	0.4375	0.5004	0.0467
IE	0.0057	0.0029	0.0000	0.0025	0.0052	0.0076	0.0162	0.0051
IT	0.1329	0.1350	0.0713	0.1168	0.1349	0.1500	0.1864	0.0332
NL	0.0860	0.0873	0.0514	0.0709	0.0864	0.1007	0.1243	0.0298
PT	0.0056	0.0049	0.0003	0.0036	0.0047	0.0081	0.0123	0.0045

Table 2. Percentage *clr*-variance contributions by country, with equal weighting (left side) and with differential weighting based on the arithmetic means of the closed parts in the mean column in Table 1 (right side). Country codes: AT (Austria), BE (Belgium), DE (Germany), EL (Greece), ES (Spain), FR (France), IE (Ireland), IT (Italy), NL (the Netherlands), and PT (Portugal).

	w_j	***clr*-Variance Contrib. (%)**	w_j	***clr*-Variance Contrib. (%)**
AT	0.1	5.84	0.0138	4.23
BE	0.1	3.23	0.0413	6.65
DE	0.1	2.60	0.1840	10.80
EL	0.1	3.86	0.0225	2.46
ES	0.1	8.48	0.0965	37.69
FR	0.1	2.39	0.4117	5.74
IE	0.1	63.71	0.0057	21.14
IT	0.1	2.20	0.1329	4.73
NL	0.1	2.00	0.0860	4.46
PT	0.1	5.60	0.0056	2.11

4.2. Weighted Chronological Clustering

Integrating the set of differential weights reported in Table 2 and the weighted distances in (10) into the CONISS algorithm available in the R package `rioja` (Juggins 2023), we implemented the proposal outlined in Section 3.2 for a weighted chronological clustering of the SRISK compositions. A preliminary analysis was carried out to verify that the time series of the weighted *clr*-coefficients associated with each part were not influenced by seasonal factors which might have induced similarities between non-adjacent CoDa samples. The analysis was performed with a seasonal and trend decomposition using locally estimated scatterplot smoothing. The corresponding measures of seasonal strength available in the R package `tsfeatures` (Hyndman et al. 2023) confirmed the absence of seasonal components in all the series.

The CoDa dendrogram in Figure 2 provides a structured periodization of relative SRISK trajectories in the Eurozone and an insightful view of when the major turning points occurred. To determine the number of contiguous clusters that could be reasonably distinguished in this representation, we compare the decrease in the within-cluster sum of squares (WSS) associated with each stage, $k \geq 2$, of the partition to that obtained from a broken stick model representing a random distribution of k sub-periods within the time sequence (see, e.g., Bennett 1996). As shown in Figure 3, splits for the first twelve clusters account for WSS reductions that exceed those expected from the broken stick model, while splits for all subsequent clusters lead to WSS reductions that are lower than in the broken stick model. We thus have an indication that twelve is the maximum number of reliably recognizable clusters based on the structure in the CoDa set. However, after the fifth split, the pattern of decrease in the WSS becomes very similar between CONISS clustering and the broken stick model. Therefore, we retain a five-group periodization as a reasonable compromise between statistical significance and the complexity of the clustering solution.

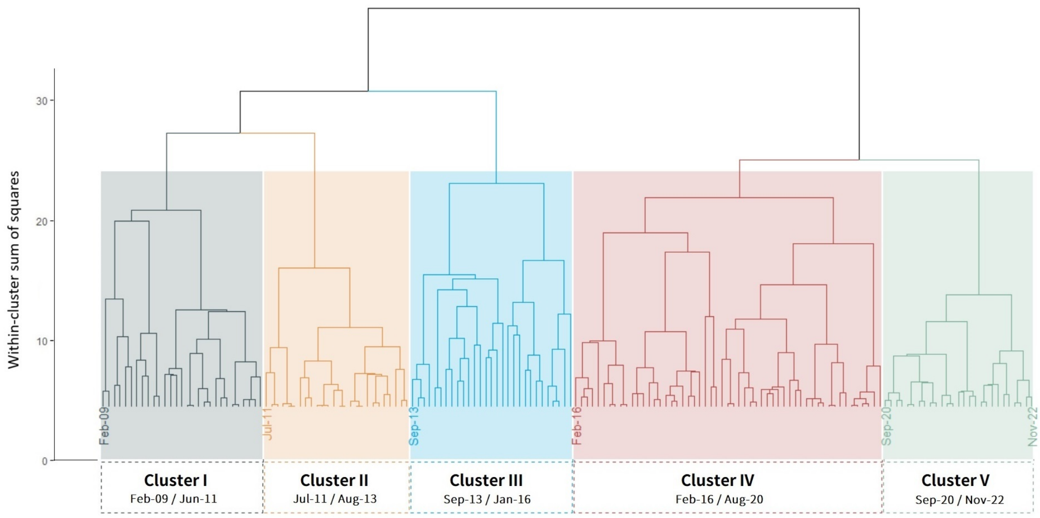

Figure 2. Weighted chronological clustering of SRISK compositions in the Eurozone, with a five-group partition.

According to the database of financial crises developed by the European Systemic Risk Board (see Lo Duca et al. 2017 for the methodology; the database is accessible at https://www.esrb.europa.eu/pub/financial-crises/html/index.en.html, accessed on 13 November 2024), cluster I in our CoDa dendrogram covers the initial spread of the GFC to Europe and the consequent worsening of the European SDC, with a number of countries beginning to fear the size of their budget deficits (February 2009–June 2011). Interestingly, the turning point that marks the passage from cluster I to cluster II occurs in July 2011, concomitant with the publication of results of the first EU-wide stress test executed by the European Banking Authority and the announcements concerning Italy's first austerity package and the second Greek bail-out (see also Cotter and Suurlaht 2019 for a detailed analysis of this period). The time segment that we label cluster II (July 2011–August 2013) is widely recognized as a phase of great uncertainty for the performance of the European banking system and the debt situation of some member states, raising real questions about the survival of the monetary union (Lannoo 2011). The split between clusters II and III, identified as September 2013 in the CoDa dendrogram, closely follows the announcement of the official exit of the Eurozone from recession with a 0.3% GDP growth in the second quarter of 2013 (Le Blond et al. 2013). Then, cluster III (September 2013–January 2016) sees a consistent decline in interest rates and a progressive improvement in financial stability in a number of EU countries, as new policy measures were designed and implemented under the macro-prudential framework of the Single Supervisory Mechanism (Kok et al. 2022; see also Nkwaira and Van der Poll 2023 for a broader discussion of macro-prudential policies). The transition between clusters III and IV takes place in February 2016, when the EU Bank Recovery and Resolution Directive and the Single Resolution Mechanism Regulation had just become fully operational with a complete set of powers, including the "bail-in-tool", to ensure the orderly resolution of bank crises in the EU (cf. Philippon and Salord 2017). Notwithstanding the consolidation of the European Banking Union, cluster IV (February 2016–August 2020) is marked by a sequence of interconnected stress events including Brexit and the Italian banking crisis, followed by the outbreak of the COVID-19 pandemic in February–March 2020. Finally, the shift from cluster IV to V coincides with the second COVID-19 wave in Europe (September 2020). The Russian invasion of Ukraine in February 2022 and the subsequent energy crisis are the main stress event that characterize the last cluster (September 2020–November 2022).

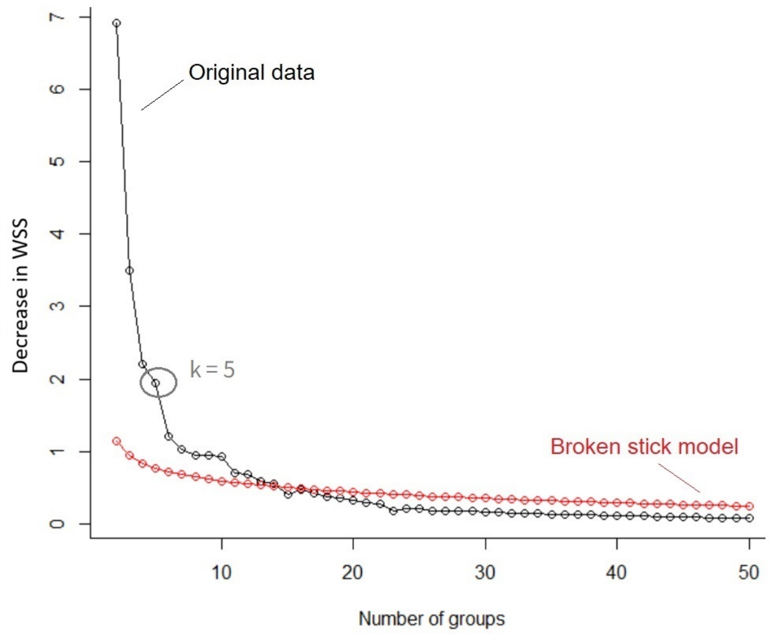

Figure 3. The decrease in the WSS (within-cluster sum of squares) at different fusion levels, $k = 2, \ldots, 50$, for the weighted chronological clustering classification (black line), compared to a broken stick model of randomly arranged samples (red line).

Table 3 illustrates the evolution—along the time sequence highlighted in the CoDa dendrogram—of the cluster means of the SRISK compositions:

$$cen^{(k)}(\mathbf{X}) = \mathcal{C}[g_1^{(k)}, \ldots, g_J^{(k)}], \tag{22}$$

where $g_j^{(k)} = \prod_{i=1}^{n_k} x_{ij}^{1/n_k}$ is the geometric mean of the j-th part in the k-th cluster, formed by n_k elements, for $k = \mathrm{I}, \ldots, \mathrm{V}$. This information is complemented by the cluster averages of the total SRISK (in billions of EUR), reported in the last column, which highlight the parallel between compositional shifts and the evolution of the absolute SRISK values in our time series clustering of systemic risk in the EU.

Table 3. Compositional profiles of chronological clusters based on their center coordinates (closed geometric means of SRISK parts) and cluster averages of the total SRISK in billions of EUR (last column). The full series row is the center column in Table 1. Bold type indicates center coordinates and total SRISK averages that exceed the corresponding values in the full series row. In brackets: the number of samples for each cluster and for the whole time series. Country codes: AT (Austria), BE (Belgium), DE (Germany), EL (Greece), ES (Spain), FR (France), IE (Ireland), IT (Italy), NL (Netherlands), and PT (Portugal).

	DE	PT	AT	NL	BE	FR	EL	IT	IE	ES	Avg SRISK
Cluster I (29)	**0.2340**	**0.0071**	0.0115	**0.1096**	**0.0729**	0.3943	0.0176	0.0986	**0.0134**	0.0411	**660.4**
Cluster II (26)	**0.2002**	**0.0096**	**0.0161**	**0.0886**	**0.0498**	0.3789	**0.0274**	0.1276	**0.0069**	**0.0950**	**808.9**
Cluster III (29)	**0.2066**	**0.0052**	**0.0158**	**0.0953**	**0.0396**	**0.4529**	**0.0240**	0.1171	0.0001	0.0434	527.9
Cluster IV (55)	0.1743	0.0029	0.0096	0.0684	0.0277	0.4170	0.0191	**0.1576**	**0.0036**	**0.1199**	559.3
Cluster V (27)	0.1185	0.0041	**0.0164**	0.0864	0.0289	**0.4247**	**0.0246**	**0.1461**	**0.0050**	**0.1453**	**747.6**
Full series (166)	0.1859	0.0049	0.0131	0.0873	0.0395	0.4243	0.0221	0.1350	0.0029	0.0849	641.2

To gain a deeper insight into these aspects, we also investigated the interdependence among the Eurozone countries using an empirical measure of commonality derived from weighted Principal Component Analysis (PCA) of the SRISK compositions (see, e.g., chap. 5 in Greenacre 2018, for the methodology). In Figure 4, we display the proportions of the total log-ratio variance (21) which are explained by eigenvalues associated with the first three principal components (labeled PC1, PC2, PC3) in each chronological cluster. According to Billio et al. (2012), periods in which the first few eigenvalues capture a large portion of total volatility may be considered indicative of increased interconnectedness in the system, which is often associated with times of distress and crisis.

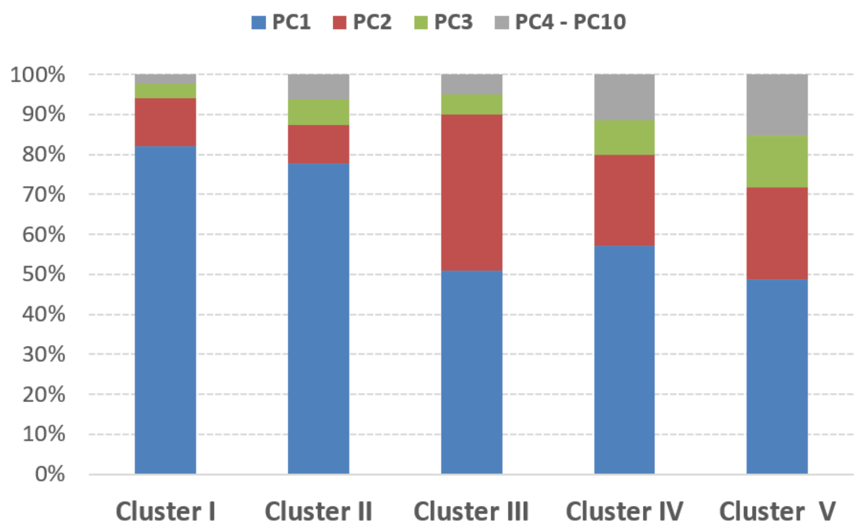

Figure 4. Percentage contributions of weighted CoDa principal components to the total log-ratio variance of the SRISK compositions in the Eurozone, by chronological clusters.

Integrating the information displayed in Figure 4 with the description of cluster centers and average SRISK totals in Table 2, we arrive at the following characterization of cluster profiles from a compositional perspective. In cluster I, the first principal component accounts for more than 80% of the total log-ratio variance, reflecting the progressing path of the financial and sovereign debt crises towards a truly systemic dimension across the Eurozone. At the same time, the (closed) geometric means of the SRISK share associated with Germany, the Netherlands, Belgium, and Ireland assume their maximal values, indicating a significant contribution from Northern Europe to SRISK compositions in the immediate aftermath of the GFC. In cluster II, the SRISK proportions of Portugal and Greece reach their maxima, and remarkable growth is also observed for Spain and Austria. These elements signal a progressive displacement of systemic imbalances towards different areas of the Eurozone, concomitant with a sharp rise in the total SRISK and persistent commonality in the compositional variation. In cluster III, the average share of France is maximal and the SRISK parts associated with Germany, Portugal, Austria, the Netherlands, Belgium, and Greece persist above the corresponding full-series averages. However, compared to the previous phase, both the total SRISK and the percentage variance explained by PC1 appear considerably lower. Indeed, as shown in Table 3, a number of countries begin to display a gradual decrease in their SRISK contributions, unraveling some discrepancies among individual EU members along the recovery path from the SDC. Clusters IV and V exhibit an important shift in the CoDa center towards Southern EU countries (particularly Italy, Spain, and Greece, but also France and Austria) and Ireland. In cluster IV, this shift appears a likely consequence of the Italian banking crisis and the post-Brexit scenario, resulting in renewed concern about the future of the Eurozone and increased volatility in financial markets. In cluster V, the redistribution of the SRISK share possibly reflects the vulnerability of Southern Europe to the severe consequences of COVID-19, the rise in inflation, and the shock to energy prices following the Russian invasion of Ukraine. Interestingly, these events triggered the return of the total SRISK to above its GFC average level but did not induce a corresponding increase in commonality in the compositional variation; in cluster V, in particular, the proportion of the total log-ratio variance associated with the first eigenvalue drops below 50%, and three principal components are required to capture the same proportion of variance explained by PC1 in cluster I.

4.3. Weighted Geographical Clustering

Based on the set of differential weights reported in Table 2 and the Ward clustering procedure for parts described in Section 3.3, we performed a weighted geographical clustering of the SRISK compositions in the transposed data table. The corresponding CoDa dendrogram, illustrated in Figure 5, shows a hierarchical grouping structure in which the strongest similarity is found in a cluster of countries from continental Europe (Austria, Belgium, the Netherlands, and Germany) plus Portugal; a second and distinct SRISK profile is associated with a smaller group of Euro-Mediterranean countries (France, Greece, and Italy). Finally, Ireland and Spain exhibit country-specific behaviors, with larger dissimilarities in their relative SRISK patterns compared to the rest of the Eurozone.

In order to investigate the links between chronological and geographical CoDa clustering results, we simplified the multivariate structure of the SRISK compositions by amalgamating (i.e., merging by summation) some parts in accordance with the partitioning suggested by the dendrogram in Figure 5 (Greenacre 2020). In particular, we proposed a three-part amalgamation with the following:

$$
\begin{aligned}
a_1 &= \mathrm{DE} + \mathrm{PT} + \mathrm{AT} + \mathrm{NL} + \mathrm{BE} \qquad \text{(continental Europe and Portugal)},\\
a_2 &= \mathrm{FR} + \mathrm{EL} + \mathrm{IT} \qquad \text{(Euro-Mediterranean component)},\\
a_3 &= \mathrm{ES} + \mathrm{IE} \qquad \text{(high-volatility component)}.
\end{aligned}
$$

As shown in Table 4, the newly formed composition, $\mathcal{C}[a_1, a_2, a_3]$, has center coordinates that reflect the overall prevalence of the Euro-Mediterranean component (57.72%), followed by continental Europe and Portugal (33.13%) and, with a much smaller proportion, by the high-volatility component including Ireland and Spain (9.15%).

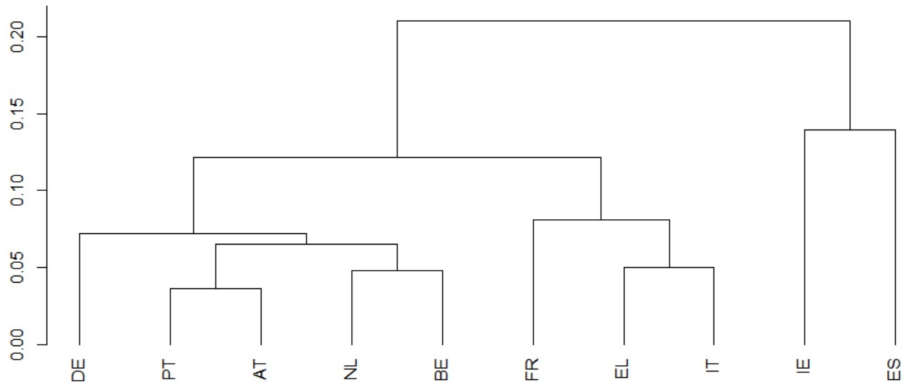

Figure 5. Weighted geographical clustering of SRISK compositions. Country codes: AT (Austria), BE (Belgium), DE (Germany), EL (Greece), ES (Spain), FR (France), IE (Ireland), IT (Italy), NL (Netherlands), and PT (Portugal).

Table 4. Compositional profiles of geographical SRISK amalgamations, $\mathcal{C}[a_1, a_2, a_3]$, based on their closed geometric means (center coordinates). The five-cluster segmentation corresponds to the constrained chronological clustering solution described in Section 4.2.

	a_1	a_2	a_3
Cluster I	0.4332	0.5100	0.0568
Cluster II	0.3640	0.5344	0.1017
Cluster III	0.3629	0.5929	0.0443
Cluster IV	0.2844	0.5927	0.1229
Cluster V	0.2554	0.5945	0.1501
Full series	0.3313	0.5772	0.0915

Taking a closer look into the time dynamics of these CoDa centers along the chronological clusters obtained in Section 4.2, we immediately detect three distinct patterns of evolution. The (closed) geometric means of a_1 and a_2 show opposite trends, as the reduction in the SRISK proportion of continental Europe and Portugal progresses in parallel with a consistent increase in the relative share of the Euro-Mediterranean part. The highly volatile component a_3 exhibits peculiar dynamics, evolving from a center coordinate of 5.68% in cluster I to a maximum of 15.01% in cluster V, with considerable variation during intermediate time segments.

As a robustness check, we repeated both the chronological and geographical clusterings without two outliers identified with the Mahalanobis distance and corresponding to March 2015 and January 2018. The turning points for the five chronological clusters and the country classification into three clusters remained unaltered.

5. Discussion

This article proposed a new framework for tackling the evolution of systemic risk in the Eurozone as a compositional time series with multiple parts, corresponding to the relative SRISK contributions of different constituent members. In CoDa analysis, “individuals” (here, the single member states) were not considered independently but parts of a “whole” (the Eurozone financial system), so the focus was on relative information (percentage SRISK shares) rather than absolute measures expressed in monetary units. This approach is inherently consistent with the global nature of systemic risk, which arises from the combination of multiple interrelated sources. Compared to Porro (2022) and Fiori and

Porro (2023), which are—to the best of our knowledge—the only articles that have applied CoDa analysis to systemic risk so far, the novelty of the present study consists in proposing a compositional time series classification that not only detects groups of similarly behaved countries but also identifies turning points in the SRISK distribution over time.

Based on adjacency-constrained Ward clustering and adapting the CONISS algorithm of Grimm (1987) to a compositional time series with weighted parts, we indeed identified some important trajectories and shifts in the compositions of systemic risk in the EU. To control for possible distortion effects associated with extreme variation in one or few components (e.g., Ireland), we preliminarily introduced a system of differential weights that calibrated the influence of low-proportion/high-dispersion parts on the classification results. This approach, recommended in recent research of the compositional classification of financial distress profiles (Dao et al. 2024; Jofre-Campuzano and Coenders 2022), yielded a well-balanced decomposition of the compositional time series variance (Table 2), in which Ireland and Spain appeared as the main sources of compositional fluctuation, as opposed to much lower contributions from Greece and Portugal. These findings are rather new and different from the existing literature as they go beyond the traditional view of Greece–Ireland–Italy–Portugal–Spain (GIIPS countries) as a unified group of highly vulnerable entities and suggest distinct mechanisms in their relative SRISK trajectories. The weighting structure introduced in this paper additionally allowed the inclusion of smaller countries that were not analyzed in Fiori and Porro (2023). Based on the mean country contributions to the SRISK of the Eurozone, the differential weighting scheme shown in Table 2 assigns a higher weight to France (0.4117) compared to Germany (0.1840). Indeed, the French financial system has a relatively high risk concentration in few key institutions which have led market capital shortfall measures, both in absolute EUR amounts and relative to the country's GDP (Acharya et al. 2014; Engle et al. 2014). On the other hand, the German financial system is more diversified and includes a non-negligible fraction of unlisted banks, for which market-based measures of systemic risk cannot be computed (Engle 2018). Furthermore, following the outbreak of the SDC, German banks restructured their sovereign bond portfolios, reducing the exposure to peripheral Eurozone debt and, as a result, lowering their systemic risk (Buch et al. 2016).

The outcome of our weighted CONISS procedure sheds further light on the SRISK trajectories, indicating a structured periodization of the compositional time series into five time segments (chronological clusters) which are clearly recognizable from the CoDa dendrogram in Figure 2. Interestingly, the turning points that mark the separation between consecutive clusters have meaningful intersections with the database of financial crises developed by the European Systemic Risk Board (Lo Duca et al. 2017), as well as with the alternating sequence of stress events and supervisory responses from EU institutions presented by Kok et al. (2022).

A peculiar contribution of chronological CoDa clustering relates to the characterization of compositional profiles associated with the various sub-periods (Table 3). While in the aftermath of the GFC and the initial phase of the SDC (cluster I) the relative distribution of SRISK share is mainly centered on Northern Europe, in the last periods extending from February 2016 to November 2022 (clusters IV and V), the compositional center considerably shifts towards the Euro-Mediterranean region. This change progressively matures in intermediate clusters, following a sequence of peaks and troughs in the highly volatile SRISK share of countries like Spain, Ireland, and—to a lesser extent—Greece, in combination with a neatly decreasing trend for Germany, a notable growth for Italy, and a permanently relevant contribution for France. The multivariate nature of CoDa methods has the merit of capturing the complex combination of concurrent elements that determined the displacement of systemic risk from core to peripheral EU countries between 2009 and 2022 and summarizing it at a glance. Our work thus supports, unifies, and extends a number of important findings that have been characterized separately in previous research, including, in particular, the existence of regional asymmetries in the EU's response to the GFC-SDC (Shambaugh 2012), a notable increase in the systemic influence of Italian and Spanish banks starting from late 2011 (Black et al. 2016), a progressive improvement in the situation in Germany after the initial susceptibility to spillovers from the US, and a somehow strong role of Ireland as a financial shock transmitter (MacDonald et al. 2015). Additionally, empirical measures of commonality derived from weighted CoDa PCA (Figure 4) show that the core–periphery displacement of systemic risk compositions in the EU was accompanied by a consistent decline in the proportion of the total log-ratio variance explained by the first principal component (PC1), which dropped from over 80% in cluster I (financial and sovereign debt crises) to less than 50% in cluster V (during and after COVID-19, until November 2022). Combined with the results of chronological CoDa clustering, these findings suggest that the core and peripheral EU countries have become less interdependent in recent years, and the stress conditions following COVID-19 and the war in Ukraine hit the southern Euro area stronger than the northern area. In particular, the lower importance of PC1 reflects a progressive increase in the heterogeneity of the SRISK compositions, which may (in Clusters III and IV) or may not (in Cluster V) be accompanied by a decline in the magnitude of the average systemic risk measure (Table 3). This finding confirms the importance of monitoring the fragmentation of interbank and financial markets in the Eurozone, following the consequences of the SDC (see also Betz et al. 2016).

The above interpretations appear to be confirmed by the results of weighted geographical CoDa clustering. Based on the CoDa dendrogram in Figure 5, we are able to detect a grouping structure between two clearly identifiable subsets of parts: the former is centered in the Euro-Mediterranean region (including France, Italy, and Greece), whose aggregate SRISK contribution steadily increased post-GFC; the latter, including Portugal and the continental EU (Austria, Belgium, Germany, and the Netherlands), exhibits an opposite trend of consistent decline in its systemic relevance over time. Interestingly, the similarity between Portugal and the continental EU in the recovery path from the GFC-SDC has not received much attention in the literature, with a notable exception in a recent study by Alves et al. (2021) who documented the exit of Portugal from the "perfect storm" of 2010 in light of the key role played by the central bank during the liquidity crisis. Ireland and Spain are somehow excluded from the grouping structure visible in Figure 5, as they seem to have contributed and reacted to the spread of financial stress with country-specific behaviors characterized by a highly volatile SRISK share. Since a relative scale can give better information than an absolute one in the comparison of large and small proportions (Pawlowsky-Glahn et al. 2015), CoDa methods appear to be particularly effective in detecting not only the recovery paths but also the instability threats associated with larger and smaller countries in various EU regions.

The results of our weighted CoDa clustering for the SRISK compositions should be interpreted within the context of the vast literature on systemic risk, which has identified and extensively explored a variety of channels behind contagion mechanisms in the EU. It is important to remember that systemic risk can propagate through multilayer networks, which depend on sovereign, banking, and equity sectors' risk spillovers among countries (Foglia et al. 2023); macro-uncertainty and financial distress also play different roles (Cipollini and Mikaliunaite 2020). In particular, the time and spatial structure of our CoDa clusters is coherent with Cipollini and Mikaliunaite's (2020) finding that connectedness between core and periphery EU countries has diminished since the peak of the SDC, with an increasing role played by peripheral countries. We also observe that the turning points highlighted in Figure 2 are quite aligned with the sequence of peaks and troughs in connectedness measures for the sovereign and banking networks presented by Foglia et al. (2023).

A limitation of standard CoDa analysis is that, by focusing on its distribution, it ignores the total SRISK. As recommended by Ferrer-Rosell and Coenders (2018), in this article, we added total risk as an external variable and related it to the clustering in Table 3. In this manner, clusters were characterized with their overall risk levels besides the risk share distribution. This approach revealed an interesting parallel between compositional cluster centers and total SRISK information, shedding light on the reasons why "this time was different", as Borri and Di Giorgio (2022) observed in their comparison of the GFC to the recent COVID-19 shock in Europe. Indeed, the total SRISK was highest both at the beginning and the end of the period considered, but the compositional centers of gravity were markedly different. In relation to these aspects, an interesting question for further research concerns the use of compositional predictors in addition to common indicators to forecast the future dynamics of systemic risk. This question could be addressed with compositional regression and other prediction methods, which have been successfully employed in multiple contexts (see, e.g., Carreras-Simó and Coenders 2021).

Overall, our findings indicate that the composition of systemic risk has considerably evolved in time and space since the failure of Lehman Brothers plunged the global financial system into meltdown. The compositional perspective adopted in our work underscores the importance of continuously monitoring not only the accumulation path of the total SRISK in the Eurozone but also, and not less importantly, the relative contribution of each country in proportion to the contributions of other participants to the financial system. This knowledge is ultimately relevant to the broader scope of macro-prudential policies (see, e.g., Nkwaira and Van der Poll 2023, for an up-to-date discussion, also in relation to global challenges posed by climate-related risks), as these policies are designed with a global perspective to prevent an excessive build-up of total risk in the system but are implemented at the level of individual contributors to ensure that the regulatory burden is commensurate with the systemic threats posed by specific risk sources.

Author Contributions: Conceptualization, G.C.; methodology, G.C.; supervision: G.C; data curation: A.M.F.; validation, G.C. and A.M.F.; formal analysis, G.C. and A.M.F.; investigation, G.C. and A.M.F.; writing—original draft preparation, A.M.F.; writing—review and editing, G.C. and A.M.F. All authors have read and agreed to the published version of the manuscript.

Funding: The research of G.C. was supported by the Spanish Ministry of Science and Innovation and ERDF—a way of making Europe [grant number PID2021-123833OB-I00], the Spanish Ministry of Health [grant number CIBERCB06/02/1002], the Department of Research and Universities of Generalitat de Catalunya [grant numbers 2021SGR01197 and 2023-CLIMA-00037], and AGAUR and the Department of Climate Action, Food, and Rural Agenda of Generalitat de Catalunya [grant number 2023-CLIMA-00037]. The funders had no role in the research process from the study design to submission.

Data Availability Statement: The data used in this article are based on SRISK measures for European institutions and countries published by the Center for Risk Management at the University of Lausanne at http://www.crml.ch/systemic-risks (accessed date: 20 July 2023). The curated CoDa set is available from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AT	Austria
BE	Belgium
CoDa	Compositional Data
CONISS	CONstrained Incremental Sum of Squares
DE	Germany
EL	Greece
ES	Spain
EU	European Union
FR	France
GDP	Gross Domestic Product
GFC	Global Financial Crisis
GIIPS	Greece–Ireland–Italy–Portugal–Spain
IE	Ireland
IT	Italy
LRMES	Long-Run Marginal Expected Shortfall
NL	The Netherlands
PCA	Principal Component Analysis
PT	Portugal
SDC	Sovereign Debt Crisis
SRISK	Systemic Risk Index
US	United States
WSS	Within-cluster sum of squares

References

Acharya, Viral V., and Sascha Steffen. 2014. Falling short of expectations? Stress-testing the European banking system. *CEPS Policy Brief* 315: 21. [CrossRef]

Acharya, Viral V., Robert F. Engle, and Matthew Richardson. 2012. Capital shortfall: A new approach to ranking and regulating systemic risks. *American Economic Review* 102: 59–64. [CrossRef]

Acharya, Viral V., Yalin Gündüz, and Timothy C. Johnson. 2022. Bank use of sovereign CDS in the Eurozone crisis: Hedging and risk incentives. *Journal of Financial Intermediation* 50: 100964. [CrossRef]

Adrian, Tobias, and Markus K. Brunnermeier. 2016. CoVaR. *The American Economic Review* 106: 1705. [CrossRef]

Aitchison, John. 1986. *The Statistical Analysis of Compositional Data*. Reprinted in 2003 with additional material by The Blackburn Press. London: Chapman & Hall.

Aldasoro, Iñaki, and Ivan Alves. 2018. Multiplex interbank networks and systemic importance: An application to European data. *Journal of Financial Stability* 35: 17–37. [CrossRef]

Alves, Nuno, Diana Bonfim, and Carla Soares. 2021. Surviving the perfect storm: The role of the lender of last resort. *Journal of Financial Intermediation* 47: 100918. [CrossRef]

Andrieş, Alin Marius, Alexandra Maria Chiper, Steven Ongena, and Nicu Sprincean. 2024. External wealth of nations and systemic risk. *Journal of Financial Stability* 70: 101192. [CrossRef]

Bacon-Shone, John. 2011. A short history of compositional data analysis. In *Compositional Data Analysis: Theory and Applications*. Edited by Vera Pawlowsky-Glahn and Antonella Buccianti. Chichester: John Wiley & Sons, pp. 1–11.

Bennett, Keith D. 1996. Determination of the number of zones in a biostratigraphical sequence. *New Phytologist* 132: 155–70. [CrossRef]

Betz, Frank, Nikolaus Hautsch, Tuomas A. Peltonen, and Melanie Schienle. 2016. Systemic risk spillovers in the European banking and sovereign network. *Journal of Financial Stability* 25: 206–24. [CrossRef]

Billio, Monica, Mila Getmansky, Andrew W. Lo, and Loriana Pelizzon. 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics* 104: 535–59. [CrossRef]

Black, Lamont, Ricardo Correa, Xin Huang, and Hao Zhou. 2016. The systemic risk of European banks during the financial and sovereign debt crises. *Journal of Banking & Finance* 63: 107–25.

Borri, Nicola, and Giorgio Di Giorgio. 2022. Systemic risk and the COVID challenge in the European banking sector. *Journal of Banking & Finance* 140: 106073.

Bratis, Theodoros, Nikiforos T. Laopodis, and Georgios P. Kouretas. 2018. Contagion and interdependence in Eurozone bank and sovereign credit markets. *International Journal of Finance & Economics* 23: 655–74.

Brownlees, Christian, and Robert F. Engle. 2017. SRISK: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies* 30: 48–79. [CrossRef]

Brownlees, Christian, Ben Chabot, Eric Ghysels, and Christopher Kurz. 2020. Back to the future: Backtesting systemic risk measures during historical bank runs and the great depression. *Journal of Banking & Finance* 113: 105736.

Brunnermeier, Markus K., Luis Garicano, Philip R. Lane, Marco Pagano, Ricardo Reis, Tano Santos, David Thesmar, Stijn Van Nieuwerburgh, and Dimitri Vayanos. 2016. The sovereign-bank diabolic loop and ESBies. *American Economic Review* 106: 508–12. [CrossRef]

Buch, Claudia M., Micheal Koetter, and Jana Ohls. 2016. Banks and sovereign risk: A granular view. *Journal of Financial Stability* 25: 1–15. [CrossRef]

Carreras-Simó, Miquel, and Germà Coenders. 2021. The relationship between asset and capital structure: A compositional approach with panel vector autoregressive models. *Quantitative Finance and Economics* 5: 571–90. [CrossRef]

Center for Financial Risk Management at Lausanne. 2023. Systemic Risk: Latest Systemic Risk Measures in Europe. Available online: https://crml.ch/systemic-risks (accessed on 13 December 2022).

Cipollini, Andrea and Ieva Mikaliunaite. 2020. Macro-uncertainty and financial stress spillovers in the Eurozone. *Economic Modelling* 89: 546–558. [CrossRef]

Coenders, Germa, Juan José Egozcue, Kamila Fačevicová, Carolina Navarro-López, Javier Palarea-Albaladejo, Vera Pawlowsky-Glahn, and Raimon Tolosana-Delgado. 2023. 40 years after Aitchison's article "The statistical analysis of compositional data". Where we are and where we are heading. *SORT. Statistics and Operations Research Transactions 47*: 207–28.

Cotter, John, and Anita Suurlaht. 2019. Spillovers in risk of financial institutions. *The European Journal of Finance* 25: 1765–92. [CrossRef]

Dao, Binh Thi Thanh, Germa Coenders, Phuong Hoai Lai, Trang Thu Dam, and Huong Thi Trinh. 2024. An empirical examination of financial performance and distress profiles during COVID-19: The case of fishery and food production firms in Vietnam. *Journal of Financial Reporting and Accounting, ahead of print*. [CrossRef]

Di Donato, Valentino, Paola Esposito, Vittorio Garilli, Debora Naimo, Giuseppe Buccheri, Mauro Caffau, Giuliano Ciampo, Antonino Greco, and Damiano Stanzione. 2009. Surface–bottom relationships in the Gulf of Salerno (Tyrrhenian Sea) over the last 34 kyr: Compositional data analysis of palaeontological proxies and geochemical evidence. *Geobios* 42: 561–79. [CrossRef]

Egozcue, Juan José, and Vera Pawlowsky-Glahn. 2019. Compositional data: The sample space and its structure. *Test* 28: 599–638. [CrossRef]

Engle, Robert F. 2018. Systemic risk 10 years later. *Annual Review of Financial Economics* 10: 125–52. [CrossRef]

Engle, Robert F., and Cristiano Zazzara. 2018. Systemic risk in the financial system: Capital shortfalls under Brexit, the US elections and the Italian referendum. *Journal of Credit Risk* 14: 97–120. [CrossRef]

Engle, Robert F., Eric Jondeau, and Michael Rockinger. 2014. Systemic risk in Europe. *Review of Finance* 19: 145–90. [CrossRef]

Ferrer-Rosell, Berta, and Germà Coenders. 2018. Destinations and crisis. Profiling tourists' budget share from 2006 to 2012. *Journal of Destination Marketing & Management* 7: 26–35.

Filzmoser, Peter, Karel Hron, and Matthias Templ. 2018. *Applied Compositional Data Analysis: With Worked Examples in R*. Cham: Springer.

Financial Stability Board. 2009. Report to G20 Finance Ministers and Governors Guidance to Assess the Systemic Importance of Financial Institutions, Markets and Instruments: Initial Considerations. Available online: https://www.fsb.org/uploads/r_091107c.pdf (accessed on 13 August 2024).

Fiori, Anna Maria, and Francesco Porro. 2023. A compositional analysis of systemic risk in European financial institutions. *Annals of Finance* 19: 1–30. [CrossRef]

Foglia, Matteo, Vincenzo Pacelli, and Gang-Jin Wang. 2023. Systemic risk propagation in the Eurozone: A multilayer network approach. *International Review of Economics & Finance* 88: 332–346.

Greenacre, Michael. 2018. *Compositional Data Analysis in Practice*. New York: CRC Press.

Greenacre, Michael. 2020. Amalgamations are valid in compositional data analysis, can be used in agglomerative clustering, and their logratios have an inverse transformation. *Applied Computing and Geosciences* 5: 100017. [CrossRef]

Greenacre, Michael. 2021. Compositional data analysis. *Annual Review of Statistics and its Application* 8: 271–99. [CrossRef]

Greenacre, Michael, and Paul Lewi. 2009 Distributional equivalence and subcompositional coherence in the analysis of compositional data, contingency tables and ratio-scale measurements. *Journal of Classification* 26: 29–54. [CrossRef]

Greenacre, Michael, Eric Grunsky, John Bacon-Shone, Ionas Erb, and Thomas Quinn. 2023. Aitchison's compositional data analysis 40 years on: A reappraisal. *Statistical Science* 38: 386–410. [CrossRef]

Grimm, Eric C. 1987. CONISS: A FORTRAN 77 program for stratigraphically constrained cluster analysis by the method of incremental sum of squares. *Computers & Geosciences* 13: 13–35.

Huang, Xin, Hao Zhou, and Haibin Zhu. 2009. A framework for assessing the systemic risk of major financial institutions. *Journal of Banking & Finance* 33: 2036–49.

Hyndman, Rob, Yanfei Kang, Pablo Montero-Manso, Mitchell O'Hara-Wild, Thiyanga Talagala, Earo Wang, and Yangzhuoran Yang. 2023. tsfeatures: Time Series Feature Extraction. R package Version 1.1-1. Available online: https://pkg.robjhyndman.com/tsfeatures (accessed on 13 August 2024).

Hron, Karel, Peter Filzmoser, Patrice de Caritat, Eva Fišerová, and Alžběta Gardlo. 2017. Weighted pivot coordinates for compositional data and their application to geochemical mapping. *Mathematical Geosciences* 49: 797–814. [CrossRef]

Jofre-Campuzano, Pol, and Germà Coenders. 2022. Compositional classification of financial statement profiles: The weighted case. *Journal of Risk and Financial Management* 15: 546. [CrossRef]

Juggins, Steve. 2023. rioja: Analysis of Quaternary Science Data. R Package Version 1.0-6. Available online: https://cran.r-project.org/package=rioja (accessed on 13 August 2024).

Karimalis, Emmanouil N., and Nikos K. Nomikos. 2018. Measuring systemic risk in the European banking sector: A copula CoVaR approach. *The European Journal of Finance* 24: 944–75. [CrossRef]

Kok, Christoffer, Francesco Paolo Mongelli, and Karin Hobelsberger. 2022. A tale of three crises: Synergies between ECB tasks. In *ECB Occasional Paper*. 2022/305. Frankfurt: ECB.

Lannoo, Karel. 2011. The second EU Bank Stress Test: Further Work in Progress. *CEPS Commentary*, July 22.

Le Blond, Josie, Kim Willsher, Helena Smith, Henry McDonald, and Lizzy Davies. 2013. Eurozone exits recession: The view from the top . . . and bottom. *The Guardian*, August 14.

Lo Duca, Marco, Anne Koban, Marisa Basten, Elias Bengtsson, Benjamin Klaus, Piotr Kusmierczyk, Jan Hannes Lang, Carsten Detken, and Tuomas Peltonen. 2017. A New Database for Financial Crises in European Countries. In *ECB Occasional Paper*. 2017/194. Frankfurt: ECB.

MacDonald, Ronald, Vasilios Sogiakas, and Andreas Tsopanakis. 2015. An investigation of systemic stress and interdependencies within the Eurozone and Euro area countries. *Economic Modelling* 48: 52–69. [CrossRef]

Magkonis, Georgios, and Andreas Tsopanakis. 2020. The financial connectedness between Eurozone core and periphery: A disaggregated view. *Macroeconomic Dynamics* 24: 1674–99. [CrossRef]

Martín-Fernández, Josep Antoni, Valentino Di Donato, Vera Pawlowsky-Glahn, and Juan José Egozcue. 2024. Insights in hierarchical clustering of variables for compositional data. *Mathematical Geosciences* 56: 415–35. [CrossRef]

Nkwaira, Chekani, and Huibrecht Margaretha Van der Poll. 2023. Anticipating the unforeseen and expecting the unexpected: Effectiveness of macro-prudential policies in curbing the impact of stranded assets in the banking sector. *Risks* 11: 87. [CrossRef]

Palarea-Albaladejo, Javier, and Josep-Antoni Martín-Fernández. 2008. A modified EM alr-algorithm for replacing rounded zeros in compositional data sets. *Computers & Geosciences* 34: 902–17.

Palarea-Albaladejo, Javier, and Josep Antoni Martín-Fernández. 2015. zCompositions—R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems* 143: 85–96. [CrossRef]

Paltalidis, Nikos, Dimitrios Gounopoulos, Renatas Kizys, and Yiannis Koutelidakis. 2015. Transmission channels of systemic risk and contagion in the European financial network. *Journal of Banking & Finance* 61: S36–S52.

Pawlowsky-Glahn, Vera, and Juan José Egozcue. 2001. Geometric approach to statistical analysis on the simplex. *Stochastic Environmental Research and Risk Assessment 15*: 384–98. [CrossRef]

Pawlowsky-Glahn, Vera, Juan José Egozcue, and Raimon Tolosana-Delgado. 2015. *Modeling and Analysis of Compositional Data*. Chichester: John Wiley & Sons.

Philippon, Thomas, and Aude Salord. 2017. Bail-ins and bank resolution in Europe. *Geneva Reports on the World Economy, Special Report* 4.

Porro, Francesco. 2022. A geographical analysis of the systemic risk by a Compositional Data (CoDa) approach. In *Mathematical and Statistical Methods for Actuarial Sciences and Finance*. Edited by Marco Corazza, Cira Perna, Claudio Pizzi and Marilena Sibillo. Cham: Springer, pp. 383–389.

Shahzad, Syed Jawad Hussain, Thi Hong Van Hoang, and Jose Arreola-Hernandez. 2019. Risk spillovers between large banks and the financial sector: Asymmetric evidence from Europe. *Finance Research Letters* 28: 153–59. [CrossRef]

Shambaugh, Jay C. 2012. The Euro's three crises. *Brookings Papers on Economic Activity 12*: 157–211. [CrossRef]

Stolbov, Mikhail, and Maria Shchepeleva. 2018. Systemic risk in Europe: Deciphering leading measures, common patterns and real effects. *Annals of Finance* 14: 49–91. [CrossRef]

Stolbov, Mikhail, and Maria Shchepeleva. 2024. 15 years of research on systemic risk across the globe: The evolution of the field and its drivers. *Research in Globalization* 8: 100195. [CrossRef]

Varotto, Simone, and Lei Zhao. 2018. Systemic risk and bank size. *Journal of International Money and Finance* 82: 45–70. [CrossRef]

Vega Baquero, Juan David, and Miguel Santolino. 2022. Too big to fail? An analysis of the Colombian banking system through compositional data. *Latin American Journal of Central Banking (Previously Monetaria)* 3: 100060. [CrossRef]