From Chaos to Clarity: Comparative Analysis of Systemic Banking Crisis Data Sets

Victoria Landaberry

Nº 001 - 2025

Documento de trabajo
ISSN 1688-7565

From Chaos to Clarity: Comparative Analysis of Systemic Banking Crisis Data Sets[☆]

Victoria Landaberry[a*]

a *Banco Central del Uruguay*

Documento de trabajo del Banco Central del Uruguay 001-2025

Autorizado por: Jorge Ponce
Disponible en línea: 19/03/2025

Resumen

Este estudio examina nueve conjuntos de datos principales sobre crisis bancarias sistémicas y fronterizas, centrándose en las diferencias y similitudes en sus definiciones y métodos de identificación. Las variaciones significativas en las definiciones y reglas operativas conducen a discrepancias en la identificación de crisis, evaluadas utilizando el Kappa de Fleiss y el índice Generalizado de Jaccard. Se emplean métodos como minería de texto, análisis de clústeres y medidas estadísticas para cuantificar estas diferencias. Aunque los datos mejorados han reducido las discrepancias a lo largo del tiempo, las fechas de inicio se alinean más frecuentemente que las fechas de finalización, lo que conduce a inconsistencias en la duración de la crisis. Estas variaciones afectan la caracterización de la crisis, lo que influye en las evaluaciones de severidad, duración y recuperación. Para abordar estos desafíos, se aplica una regla de voto mayoritario para unificar las fechas de inicio y finalización en los conjuntos de datos. Este enfoque estandarizado mejora la consistencia y la utilidad de los datos de crisis para los investigadores y responsables de políticas, proporcionando un marco unificado para la identificación y análisis confiable de las crisis.

JEL: G01, G21, G28, C80
Palabras clave: crisis bancarias sistémicas, riesgo sistémico, estabilidad financiera, duración de la crisis, identificación de la crisis

Abstract

This study examines nine major data sets of systemic and borderline banking crises, focusing on differences and similarities in their definitions and identification methods. Significant variations in definitions and operational rules lead to discrepancies in crisis identification, evaluated using Fleiss's Kappa and the Generalized Jaccard Index. Methods such as text mining, cluster analysis, and statistical measures are employed to quantify these differences. Although improved data have reduced discrepancies over time, start dates align more frequently than end dates, leading to inconsistencies in crisis duration. These variations affect the characterization of the crisis, which influences the evaluations of severity, duration, and recovery. To address these challenges, a majority voting rule is applied to unify start and end dates across data sets. This standardized approach improves the consistency and usability of crisis data for researchers and policymakers, providing a unified framework for the reliable identification and analysis of crises.

Keywords: systemic banking crises, systemic risk, financial stability, duration of the crisis, identification of the crisis

☆ The views and conclusions in this paper are those of the author and may not necessarily reflect the positions of their affiliated institutions.

* Email: *mlanbdaberry@bcu.gub.uy*

1 Introduction

Systemic banking crises (SBC) are critical events in financial history, with their relevance underscored by the global financial crisis of 2008, which profoundly affected economic stability and growth. This crisis not only highlighted the vulnerability of financial systems, but also accelerated substantial growth in academic research on systemic risk, reflecting the increasing importance of this topic [Silva et al., 2017]. Accepted definitions, such as those of Laeven and Valencia [2020] and Reinhart and Rogoff [2009], describe these crises as severe disruptions in the banking sector that lead to insolvency or significant capital losses in multiple financial institutions, with spillover effects on the real economy.

Identifying systemic banking crises is crucial for assessing their severity, duration, and recovery processes. Accurate identification forms the basis for the development of prediction models and early warning systems, facilitates the analysis of the determinants of such crises, and informs the design of stress testing scenarios by supervisors and central banks, which often rely on historical crisis data. It is also essential to estimate the economic and social costs of SBC and to formulate macroprudential policies to prevent or mitigate their impact. For central banks, supervisors, governments, and international financial institutions, precise identification is especially critical, as these entities are responsible for implementing policy interventions to stabilize the financial system and manage recovery efforts, often incurring significant costs.

Despite the crucial importance of understanding systemic banking crises, significant challenges remain, particularly inconsistencies in definitions, methodologies, and crisis dating in existing data sets. Unlike economic recessions, which are typically defined as two consecutive quarters of negative real GDP growth, SBC lack a universally accepted definition, further complicating their identification [Chaudron and de Haan, 2014]. Moreover, debates persist regarding the theoretical and operational definitions of SBC, as well as the most effective approaches to measure their impact and determine their onset and resolution.

Previous studies, such as Chaudron and de Haan [2014], Boyd et al. [2019], Baron et al. [2021], and Sufi and Taylor [2022], have emphasized the discrepancies in the identification of crises across major data sets, underscoring the consequences of inconsistent definitions for empirical research. Existing data sets often vary in their definitions, operational criteria for identifying these episodes, and coverage in terms of countries and periods, resulting in significant differences in the identification of crises. For example, while Laeven and Valencia [2020] relies on objective financial thresholds and policy responses, Caprio and Klingebiel [2002] employs expert judgment to classify crises. Variations in operational definitions contribute to divergences in the identification of crisis episodes, even for the same countries and periods, leading to inconsistencies in start and end dates across data sets. Additionally, some data sets on banking crises distinguish between systemic and borderline non-systemic crises (e.g., [Caprio and Klingebiel, 2002]; Reinhart and Rogoff [2009]), adding complexity and introducing another layer of discrepancies.

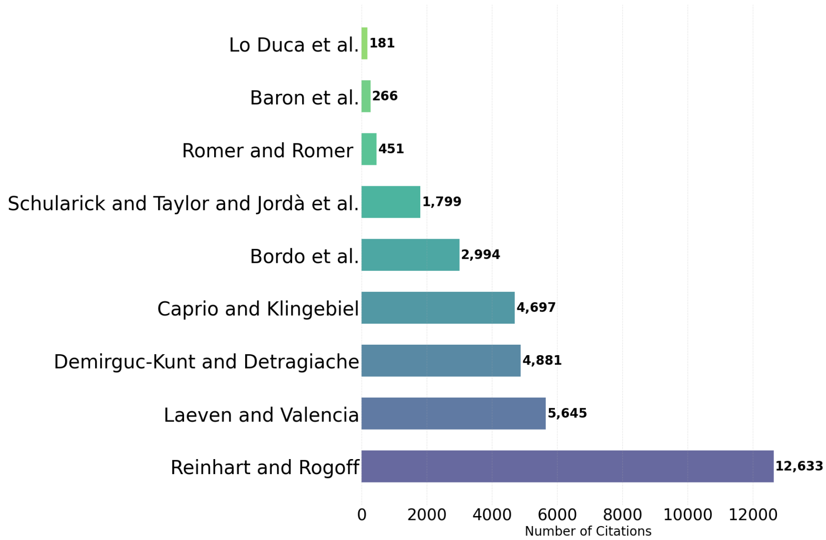
These differences, noted by Chaudron and de Haan [2014], Sufi and Taylor [2022], and Boyd et al. [2019] in a smaller sample of data sets than the one surveyed in this work, have a substantial impact on the general characterization of crises, including their frequency, duration, and the statistical behavior of macroeconomic and financial variables during these episodes. As emphasized by Jing et al. [2015], accurate identification of crises is essential for advancing research in this field. In its absence, analyses of the determinants or effects of crises become unreliable, and early warning models trained on inconsistent data sets can produce misleading signals. According to Boyd et al. [2019], the reliance on information from central banks or regulators introduces additional challenges. First, policy measures are often implemented only after a crisis has begun, leading to delays in event dating. Second, countries with weaker supervisory institutions or lower data quality may experience systemic banking crises that remain unidentified due to inaccurate reporting or the absence of policy interventions. As a result, certain crises may be overlooked. In particular, Caprio and Klingebiel [1996] and Reinhart and Rogoff [2009] found that non-performing loans data, which could be used to evaluate crises arising from asset deterioration, are sometimes unavailable or inaccurate due to banks' incentives to mask their situation, especially in environments with weak supervisory oversight.

Recent literature has attempted to address these differences by proposing alternative market-based measures to identify crisis episodes. For example, Von Hagen and Ho [2007] and Jing et al. [2015] use money market pressure indicators to identify banking crises. However, the methodology they propose results in a significant number of crises being classified as "false alarms," leading to the exclusion of their data set from this survey. Baron et al. [2021] identify banking crises using bank equity return data and compare their findings, in terms of the start dates of the episodes, with existing data sets. A notable disadvantage of this approach is that some countries lack bank equity market data and the methodology only provides information on start dates without covering end dates. Furthermore, Chaudron and de Haan [2014] propose using the number and size of bank failures, but their analysis is limited to four episodes. As noted by the authors, collecting data on these variables poses significant challenges when dealing with a large sample of countries. Although more financial data are now available than in the past, contributing to the development of better data sets in recent years and for the future, most historical episodes lack this information.

This study addresses these challenges by conducting a comprehensive comparison of nine widely used data sets on systemic banking crises, including Bordo et al. [2001], Reinhart and Rogoff [2009] (updated to 2016 [Reinhart et al., 2016]), Caprio and Klingebiel [2002], Demirgüç-Kunt and Detragiache [2005], Schularick and Taylor [2012] and Jordà et al. [2017], Romer and Romer [2017], Lo Duca et al. [2017], and Laeven and Valencia [2020], which was further extended by Nguyen et al. [2022] and Baron et al. [2021]. Evaluate inconsistencies in definitions, operational criteria, and geographical and temporal coverage while analyzing differences in the frequency, duration, and dates of crises between data sets. Finally, the study proposes a unified approach to crisis dating based on a majority voter rule to enhance the reliability and usability of SBC data for researchers and policymakers. The selection of data sets was determined by the number of citations in Google Scholar up to December 2024 (Figure 1[1]).

[1]Reinhart and Rogoff citations include Reinhart and Rogoff [2008] and Reinhart and Rogoff [2009]; Laeven and Valencia includes Laeven and Valencia [2013], Laeven and Valencia [2018], Laeven and Valencia [2020], and the update up to 2017 from Nguyen et al. [2022]; Demirgüç-Kunt and Detragiache includes Demirgüç-Kunt and Detragiache [1998] and Demirgüç-Kunt and Detragiache [2005]; Caprio and Klingebiel includes Caprio and Klingebiel [1996] and Caprio and Klingebiel [2002]; Schularick, Taylor, and Jordà includes Jordà et al. [2017]; Romer and Romer includes Romer and Romer [2015] and Romer and Romer [2017]; Baron et al. includes Baron et al. [2018] and Baron et al. [2021]; and Lo Duca et al. refers to Lo Duca et al. [2017]

Figure 1: Data set citations by author



Note: Citations as of December 2024, calculated by the author based on information from Google Scholar. As some of these articles have been updated or published in multiple versions, all citations referring to any version are included in this total.

To better understand the differences and similarities between data sets, this study employs text mining techniques combined with cluster analysis. The Term Frequency-Inverse Document Frequency (TF-IDF) method was used to quantify the importance of terms within the definitions. Principal Component Analysis (PCA) was applied for dimensionality reduction, followed by k-means clustering, an unsupervised machine learning method. Using the elbow method, five clusters were identified, each corresponding to one of the nine definitions used in the data sets. Although some definitions within each cluster share similarities, significant differences remain, reflecting variations in the definition of systemic banking crises.

To assess agreements and disagreements in crisis episode detection, the generalized Jaccard Index and Fleiss's Kappa were calculated for different time periods across the shared countries in the data sets. The results reveal significant variations in the identification of episodes, influenced by differences in definitions and identification rules. The agreement improves in more recent periods, probably due to better data availability; however, substantial differences persist.

On a second level of analysis, this study examines the agreement in terms of dates, focusing on the proportion of data sets where there is a coincidence in start and end dates, as well as exact matches. The results show that while data sets tend to be more consistent in identifying crisis start dates, greater variability and uncertainty are observed in determining end dates. Full alignment on both start and end dates occurs in only 29% of episodes identified by at least two data sets. These inconsistencies are influenced by biases related to income levels and regional factors.

In addition, this study provides further evidence of statistical differences in the durations and dates of crises, complementing previous findings (Chaudron and de Haan [2014], Boyd et al. [2019], Sufi and Taylor [2022]). Most data sets indicate that banking crises typically last between 2 and 6 years, with statistically significant differences in their distributions according to the Kolmogorov-Smirnov test. Finally, the paper proposes a unified approach to crisis dating, based on a majority voter rule, to improve the reliability and usability of SBC data for researchers and policymakers. Instead of introducing a new measure to determine the start and end dates of crises, it uses and integrates the information already available in the data sets surveyed. This approach selects the start and end dates that occur most frequently in all data sets. According to these unified criteria, financial crises last an average of 3.48 years, with a median duration of 3 years, indicating that half of the crises are resolved within this period.

This study makes several key contributions to the literature on financial crises. First, it provides the first comprehensive comparison of all major data sets on systemic banking crises, offering valuable insights into their differences and commonalities. This effort goes beyond previous studies, such as those by Boyd et al. [2019], Chaudron and de Haan [2014], and Sufi and Taylor [2022], which examined only a subset of these data sets. Second, it employs quantitative tools to assess the consistency of the data, enabling a detailed evaluation of the variations in definitions, crisis dates, and durations between data sets. Third, it presents new evidence on statistical differences among data sets, particularly concerning crisis durations, and offers a detailed comparison of these discrepancies. Finally, the study proposes a unified approach to crisis dating based on a majority voter rule, improving the reliability and usability of SBC data for researchers and policymakers alike.

This document is organized as follows. Section 2 outlines the methodology used to compare and evaluate similarities and differences between data sets. Section 3 provides a detailed description of the surveyed data sets. Section 4 investigates the similarities and differences in the definitions of systemic and borderline banking crises, together with the operational rules used to identify these episodes. Section 5 presents the results on the differences in definitions or rules in the identification of episodes between the data sets. Section 6 analyzes the differences in the dates of the crisis episodes in the data sets and explores potential biases in these differences based on the regions and income levels of the countries considered. Section 7 analyzes the difference in the duration of the crises and their distribution between the different data sets. In Section 8 a majority vote rule is applied to determine the start and end dates of crisis episodes, with a brief analysis of the duration of the crisis based on this rule. Finally, Section 9 presents the concluding remarks.

2 Methodology

To compare the nine data sets surveyed in this paper, various dimensions are analyzed. Section 3 presents a description of the time period, country coverage, and episodes identified by each data set.

The differences and similarities in the definitions and rules used to identify SBC episodes are explored, including the implications of incorporating a secondary classification of systemic and borderline banking crises in some data sets. To compare the definitions of SBC, the definitions provided by the authors and the rules used to identify an SBC are extracted from each reference document. These definitions are analyzed using text mining techniques, starting with lowercasing, tokenization, and stop-word removal to highlight the most common terms used across all definitions. To further evaluate similarities and differences, the Term Frequency-Inverse Document Frequency (TF-IDF) method is applied. TF-IDF quantifies the importance of terms within a definition relative to their frequency in all definitions, emphasizing distinctive terms while minimizing the impact of common and non-informative words [Schütze et al., 2008].

Following TF-IDF vectorization, Principal Component Analysis (PCA) is performed to reduce the dimensionality of the data, enabling a two-dimensional visualization of the TF-IDF vectors [Mekala and Rani, 2018, Drikvandi and Lawal, 2023]. The first two principal components are then used to group the definitions using the K-means algorithm [Springer, 2006]. This methodology, grounded in text mining techniques, has been employed in previous work to group text and documents (Kim and Gil [2019], Rejito et al. [2021], Kumar et al. [2021]) and identifies definitions that are similar and those that differ.

To assess agreements and disagreements in episode detection based on the definitions used, the Generalized Jaccard Index and Fleiss' Kappa are calculated [Fleiss et al., 2013, Costa, 2021]. This approach follows the methodology of Chaudron and de Haan [2014], who evaluated differences between Caprio and Klingebiel [1996], Laeven and Valencia [2013], and Reinhart and Rogoff [2009], but with key distinctions. Although the original study used Cohen's Kappa to compare pairs of data sets, Fleiss's Kappa was employed to compare multiple data sets simultaneously.

The Generalized Jaccard Index quantifies the similarity of crisis identification across overlapping data sets for the same countries and periods. It is defined as:

$$J(A_1, A_2, \ldots, A_k) = \frac{|A_1 \cap A_2 \cap \ldots \cap A_k|}{|A_1 \cup A_2 \cup \ldots \cup A_k|} \tag{1}$$

where $|A_1 \cap A_2 \cap \ldots \cap A_k|$ represents the number of events classified as crises by all overlapping data sets, and $|A_1 \cup A_2 \cup \ldots \cup A_k|$ represents the total number of events classified as crises by at least one overlapping data set. The Generalized Jaccard Index takes values between 0 and 1, where 0 indicates no agreement (no overlapping crises identified across data sets) and 1 indicates complete agreement (all data sets identify the same crises).

Fleiss's Kappa measures agreement among the data sets while accounting for chance agreement. For an event i with k_i overlapping data sets, the agreement proportion is calculated as:

$$P_i = \frac{1}{k_i(k_i - 1)} \sum_{j=1}^{2} n_{ij}(n_{ij} - 1) \tag{2}$$

where n_{ij} is the number of data sets that classify the event i into category j ($1 =$ crisis or $0 =$ no crisis). Fleiss' Kappa is then defined as:

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

where $\bar{P}$ is the mean observed agreement across all events, and $\bar{P}_e$ is the expected agreement by chance:

$$\bar{P}_e = \sum_{j=1}^{2} \left(\frac{\sum_{i=1}^{n} n_{ij}}{\sum_{i=1}^{n} k_i} \right)^2$$

Fleiss's Kappa ranges from -1 to 1, where 1 indicates perfect agreement, 0 indicates agreement equivalent to chance, and negative values indicate disagreement.

To assess differences in the distributions of the duration of the crisis, the Kolmogorov-Smirnov (KS) test is applied to each pair of data sets. The KS test is a non-parametric statistical method that compares the cumulative distribution functions (CDFs) of two data sets to determine whether they come from the same distribution. The test statistic D is defined as:

$$D = \sup_x |F_1(x) - F_2(x)| \tag{3}$$

where $F_1(x)$ and $F_2(x)$ are the empirical cumulative distribution functions of the two data sets being compared, and $\sup_x$ denotes the supremum (maximum) difference between the two CDFs. The KS test provides a p value to evaluate the statistical significance of D [Smirnov, 1948]. A p value less than 0.05 indicates that the two distributions differ significantly, meaning that they do not follow the same pattern in the way the lengths of the crises are distributed.

To derive a unified version of the start and end dates for each crisis episode, a majority voting rule is proposed. This approach selects the start and end dates that appear most frequently in the data sets. In cases where multiple start or end dates receive the same number of votes, priority is given to the earliest start date and the latest end date, ensuring a broader range is retained. This broader interval is chosen to avoid falling into the 5-year rule proposed by Laeven and Valencia [2013], which could exclude important aspects of some crises. In addition, this approach provides flexibility for researchers to identify other indicators that may reveal aspects of interest over a broader horizon. By combining all the information available across data sets in a consistent way, this methodology also provides a unified measure of the duration of the crisis, rather than multiple inconsistent measures. The results of this new data set and the consequences for characterizing the duration of the crisis are presented in Section 8.

3 Data

The survey comprises the most widely used data sets on systemic banking crises. These data sets differ in time coverage, the countries considered, the definitions of systemic banking crises, and the methods applied to identify these episodes. As a result, the number of identified crises varies, as do the start and end dates of episodes commonly identified by them, as discussed in Sections 4 to 6. Table 1 summarizes the main differences between the data sets[2].

Table 1: Overview of Systemic Banking Crisis Data Sets

Data Set	Total Crises	Countries Covered	Global Coverage (%)	Period Covered
Laeven and Valencia (2020) and Nguyen et al. (2022)	151	164	83.2	1970–2019
Caprio et al. (2002)	110	126	64.3	1970–1999
Reinhart and Rogoff (2009)	185	70	35.7	1800–2016
Demirgüç-Kunt and Detragiache (2005)	84	87	44.4	1980–1994
Bordo et al. (2001)	140	56	28.6	1880–1997
Baron et al. (2021)	222	46	23.5	1870–2016
Lo Duca et al. (2017)	46	28	14.3	1970–2016
Romer and Romer (2017)	33	24	12.2	1967–2012
Jordà et al. (2017)	88	18	9.2	1870–2008

Note: Total crises reflect systemic banking crises identified in each data set. Global coverage is calculated as the percentage of countries included in the data set relative to the total number of countries globally. Source: Author's calculations.

As shown in Table 1, Laeven and Valencia [2020] provides the broadest coverage of countries, including 164 countries, followed by Caprio and Klingebiel [2002]. Data sets with fewer countries generally adopt more restrictive analytical scopes. For example,

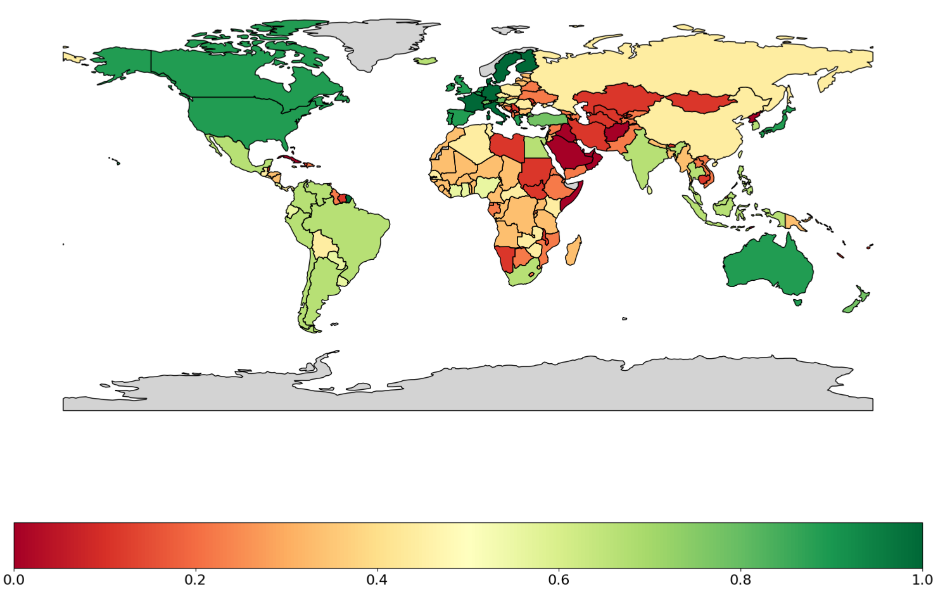
[2] This survey considers the latest version available of each data set. In the case of the Reinhart and Rogoff data set, the information corresponds to the data set updated up to 2016 from the Harvard Business School Behavioral Finance Financial Stability Project [Reinhart et al., 2016].

Romer and Romer [2017] covers 24 countries, specifically members of the Organization for Economic Cooperation and Development (OECD), while Lo Duca et al. [2017] focuses on the countries of the European Union and Norway. The data set with the fewest countries is Jordà et al. [2017], which includes only 18 advanced economies. A detailed list of countries covered by each data set is provided in Appendix A.

Using the United Nations list of countries as a reference, together with additional territories appearing in at least one data set,[3] Laeven and Valencia [2020] achieves global coverage of 83.2%. In contrast, Jordà et al. [2017] has the lowest coverage, including only 9.2% of the countries.

Figure 2 illustrates the proportion of data sets in which each country is included, with the countries colored according to their participation rate in the data sets surveyed. Developed and major economies, such as the United States, Canada, and much of Western Europe, exhibit the highest inclusion rates, reflecting their predominant presence in studies of systemic banking crises. In contrast, regions such as Sub-Saharan Africa, Central Asia, and smaller island nations are less represented.

Figure 2: Countries colored by proportion of data sets in which they are included

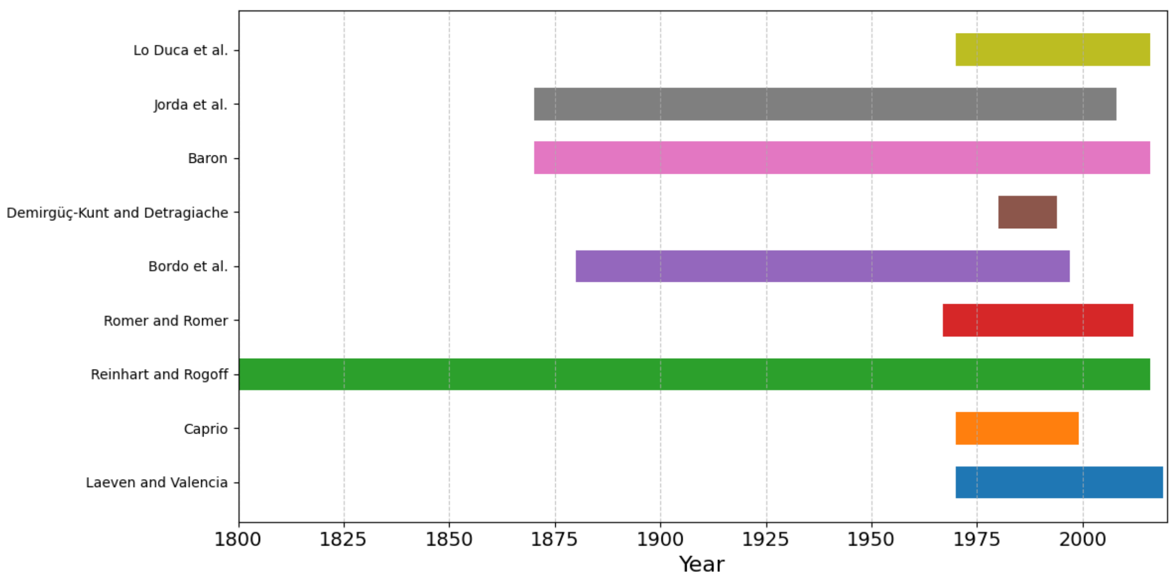


Source: Author's calculations.

[3]In addition to the 193 UN-recognized countries, territories such as New Caledonia, Hong Kong, and Taiwan are included, while Yugoslavia is excluded from the Laeven and Valencia data set due to its disintegration in the early 1990s and official dissolution in 2003.

When considering all data sets collectively, the overall time span covered extends from 1800 to 2019. The early period, from 1800 to 1869, is covered exclusively by Reinhart et al. [2016]. From 1870 onward, the coverage expands with significant overlap between Reinhart et al. [2016], Baron et al. [2021], Jordà et al. [2017], and Bordo et al. [2001], which begins in 1880. The most densely covered period is 1970 to 1999, during which the data sets of Laeven and Valencia [2020], Caprio and Klingebiel [2002], Reinhart et al. [2016], Romer and Romer [2017], Baron et al. [2021], Schularick and Taylor [2012], Lo Duca et al. [2017], and Bordo et al. [2001] overlap significantly. In contrast, Demirgüç-Kunt and Detragiache [2005] represents the smallest data set in terms of temporal coverage, ranging from 1980 to 1994 (Figure 3).

Figure 3: Time coverage by data set



Source: Author's calculations.

The overlapping periods across data sets are particularly valuable for cross-validating episode identification, allowing for an assessment of differences in crisis definitions across sources. The results of this comparison are presented in Section 4. Temporal coverage can introduce biases in the analysis of systemic and borderline banking crises. Limited data from earlier periods, coupled with a denser representation from 1970 to 1999, may lead to an emphasis on modern financial crises[4].

[4] A detailed table showing data set overlaps is provided in Appendix B.

4 Similarities and Differences in Crisis Definitions

Definitions of systemic and borderline banking crises vary between data sets. One reason, as noted by Chaudron and de Haan [2014], is the lack of a universally accepted operational definition for systemic banking crises, unlike the clear criteria used to identify economic recessions.[5] This section examines theoretical definitions, distinctions between systemic and borderline crises, and operational definitions across data sets. A classification based on episode identification criteria is also provided, along with a discussion of the implications for coverage and dating. Detailed definitions, identification criteria, and episode dates appear in Appendix C.

Common elements are present across all definitions. Systemic banking crises, in particular, involve severe financial distress within the banking sector, leading to insolvency or significant capital losses among multiple financial institutions [Caprio and Klingebiel, 1996, Bordo et al., 2001, Reinhart and Rogoff, 2009, Laeven and Valencia, 2013, 2020] or declines in bank equity [Baron et al., 2021]. Most definitions emphasize that these crises result in disruptions to credit availability, bank runs, or major bankruptcies, which prompt policy interventions to stabilize the system [Demirgüç-Kunt and Detragiache, 1998, Schularick and Taylor, 2012, Laeven and Valencia, 2018, 2020]. Furthermore, many definitions highlight spillover effects on the real economy, such as declines in investment, consumption, and economic activity, often accompanied by amplification mechanisms where financial instability worsens economic outcomes [Laeven and Valencia, 2020, Reinhart and Rogoff, 2009, Lo Duca et al., 2017]. Common words used in the definitions are presented in Figure 4.

To analyze the differences and commonalities among the definitions, the clustering analysis described in Section 2 is applied to the definitions presented in Appendix C. The results are illustrated in Figure 5, where each point represents a definition and the clusters are visually distinguished by color. The optimal number of clusters, determined using the Elbow method, is five.[6]

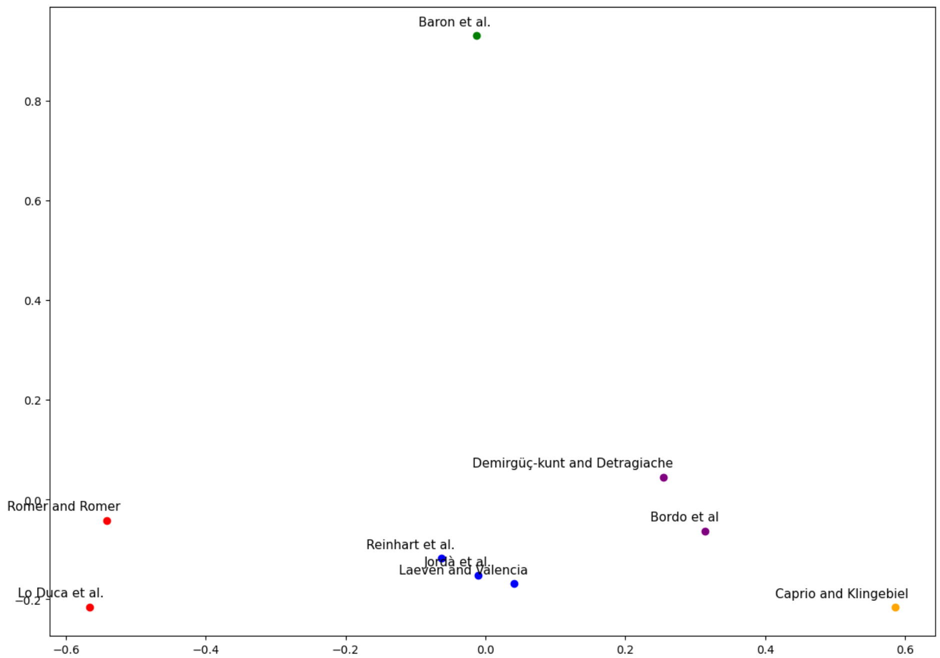
[5] An economic recession is generally defined as two consecutive quarters of negative real GDP growth, ending after two consecutive quarters of positive growth.

[6] Using BERT embeddings combined with K-means produces the same clustering results in terms of definition groups, underscoring the consistency of clusters across both methods.

Figure 4: Common Terms in Definitions of Systemic Banking Crises

Source: Author's calculations.

Figure 5: Clustering of Definitions: 2D PCA of TF-IDF Vectors



Source: Author's calculations. Colors indicate distinct clusters.

The relative distances and clustering indicate that, while many authors use common terms to define systemic banking crises, distinct differences exist. Baron et al. [2021] stands apart, describing systemic crises in terms of measurement rather than theory. Demirgüç-Kunt and Detragiache [2005] and Bordo et al. [2001] share a cluster, focusing on large-scale banking disruptions that erode confidence and capital. Similarly, Reinhart and Rogoff [2009], Jordà et al. [2017], and Laeven and Valencia [2020] emphasize financial distress, institutional failures, and policy interventions, highlighting events like bank runs. In another cluster, Romer and Romer [2017] and Lo Duca et al. [2017] focus on systemic crises with economy-wide impacts. Caprio and Klingebiel [2002] offers a clear and concise definition centered on capital depletion and sector-wide strain. Word clouds for each cluster are shown in Appendix 4.

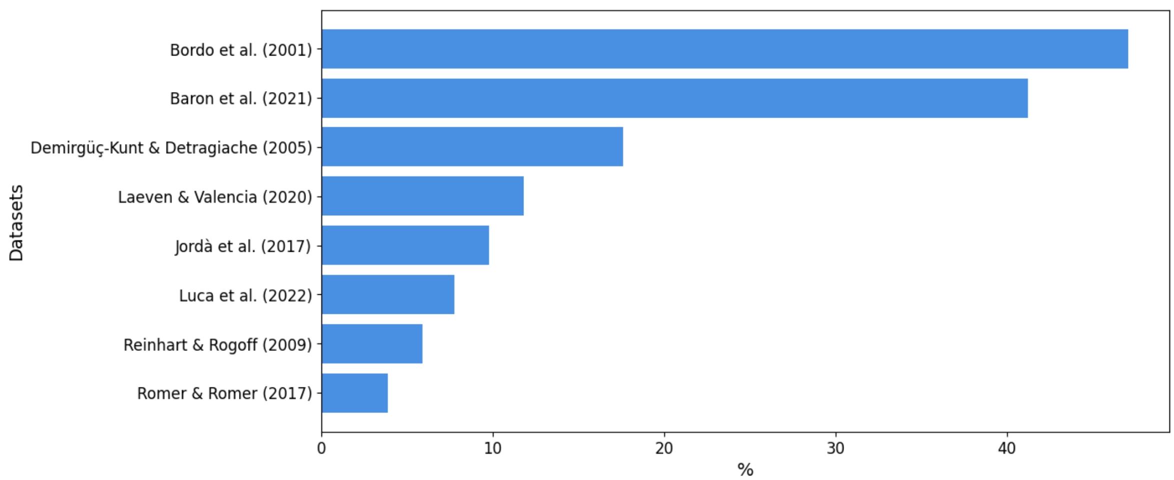
Differences in theoretical definitions result in variations in the operational definitions of systemic banking crises and their identification methods. A key observation is that, while economic impact is relevant in many theoretical definitions, only Lo Duca et al. [2017] and Romer and Romer [2017] explicitly incorporate negative real economic outcomes into their identification criteria. Within this framework, Romer and Romer [2017] classifies events according to the extent of economic impact, with generalized effects indicating moderate crises. Other authors, such as Laeven and Valencia [2020] and Bordo et al. [2001], use economic indicators, including GDP growth rates, to determine the timing or end date of crises.

Operational definitions and identification methods for determining crisis episodes can be classified into three categories. The first relies solely on objective measures or indicators, such as variables exceeding thresholds or events such as bank runs. The second employs qualitative assessments or expert judgment. The third combines both objective measures and qualitative evaluations. Among the data sets considered, only Laeven and Valencia [2020] and Baron et al. [2021] are based exclusively on objective measures. In contrast, Caprio and Klingebiel [1996], Caprio and Klingebiel [2002], Schularick and Taylor [2012], Jordà et al. [2017], and Bordo et al. [2001] use qualitative assessments, while the remaining data sets integrate objective measures with expert judgment or other qualitative evaluations.

When considering the distinction between systemic and borderline or non-systemic banking crises, Caprio and Klingebiel [1996, 2002], and Demirgüç-Kunt and Detragiache [2005], Romer and Romer [2017], Reinhart and Rogoff [2009], Laeven and Valencia [2020] identify different levels of crises with varying degrees of specificity. Although Laeven and Valencia [2020] and Demirgüç-Kunt and Detragiache [2005] acknowledge this distinction, their final data sets contain only systemic banking crises. The primary differences between these episodes focus on the extent of financial distress, policy interventions, and economic impact.

This distinction introduces discrepancies across data sets. For example, Caprio and Klingebiel [2002] identifies 51 borderline banking crises, but the percentage classified as systemic varies widely between data sets, ranging from 3.9% to 47.1%. These differences, illustrated in Figure 6, reflect variations in operational definitions and methodologies, leading to inconsistencies in the identification of crises.

Figure 6: Percentage of Borderline Crises According to Caprio and Klingebiel [2002] Classified as Systemic by Other Data Sets



Source: Author's calculations.

5 Impact of Definition Variability on Episode Identification

Variations in operational definitions lead to differences in the identification of crisis episodes. To isolate the effects of definitions and identification criteria, data sets with consistent country and year coverage are compared.[7]

The period from 1870 to 1879 is analyzed using Reinhart et al. [2016], Baron et al. [2021], and Jordà et al. [2017], focusing on 18 common countries.[8] During this period,

[7]The dataset from Romer and Romer [2017] is excluded from this comparison due to its inclusion of different types of crisis episodes that could distort the analysis.

[8]There are 18 countries common across these data sets.

Reinhart et al. [2016] identifies six banking crises (four systemic and two non systemic), Jordà et al. [2017] identifies nine episodes, and Baron et al. [2021] identifies seven. Only three episodes are consistently classified across the data sets, highlighting discrepancies in the identification criteria. The union of all episodes yields 13 unique crises, with 11 of the 18 countries experiencing at least one crisis. A Venn diagram illustrating the overlap is presented in Appendix 12, panel a. Fleiss's Kappa of -0.5 and a Generalized Jaccard Index of 0.22 reflect significant disagreements during this period.

For 1880-1966, the events identified by Reinhart et al. [2016], Bordo et al. [2001], Baron et al. [2021], and Jordà et al. [2017] are compared in the same 18 countries. Each country experienced at least one banking crisis in at least one data set, with the number of episodes ranging from 68 in Reinhart et al. [2016] (33 non systemic and 35 systemic crises) to 50 in Jordà et al. [2017]. The Generalized Jaccard Index for this period is 0.39, and Fleiss's Kappa is 0.24. The agreement among data sets improves compared to the previous period, likely due to better data availability, but discrepancies persist due to differing definitions and identification methods.

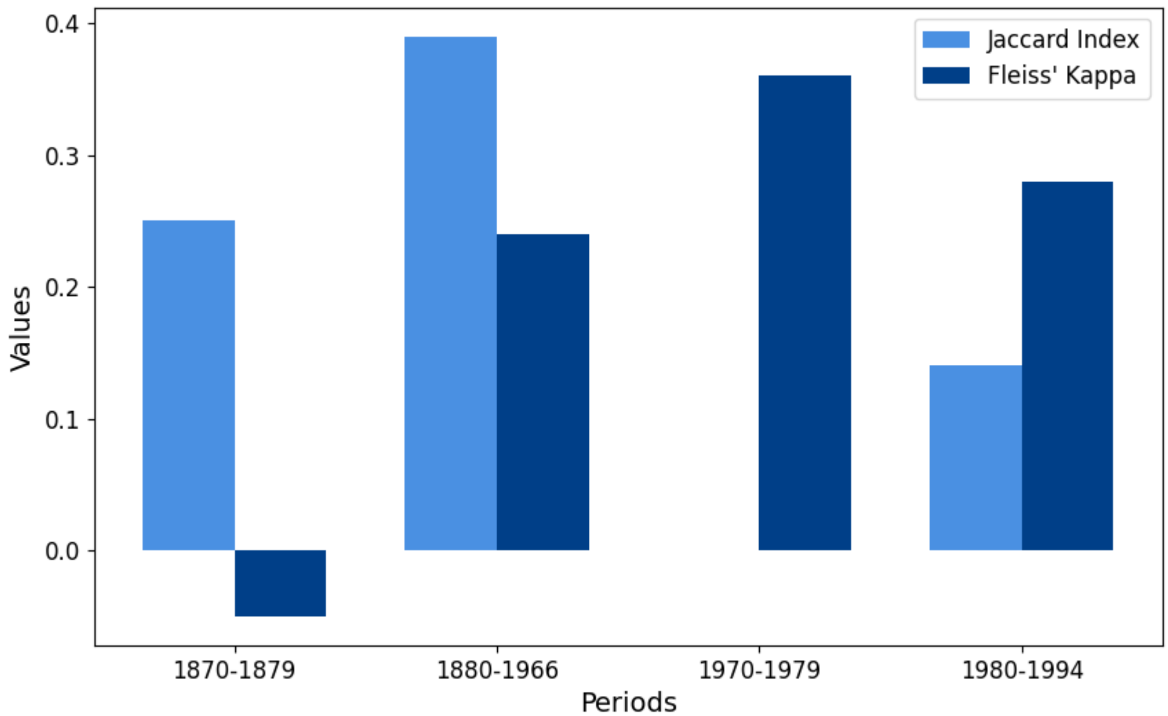
No crises are identified for the period 1967–1969 among the countries common to all data sets.

For 1970–1979, data sets from Laeven and Valencia [2020], Caprio and Klingebiel [2002], Lo Duca et al. [2017], and Romer and Romer [2017] are compared across eight common countries. Only three episodes from two countries are identified, none of which are recognized by all data sets. The Generalized Jaccard Index is 0, while Fleiss's Kappa is 0.36, indicating limited agreement.

From 1980–1994, the inclusion of Demirgüç-Kunt and Detragiache [2005] further increases the heterogeneity of the data set. The Generalized Jaccard Index for this period decreases to 0.14, and Fleiss' Kappa declines to 0.28, reflecting diminished agreement as the number of data sets and definitions increases.

For 1995-1998, no crises were identified among countries common to all data sets. Comparing data sets from Laeven and Valencia [2020], Caprio and Klingebiel [2002], Reinhart et al. [2016], and Baron et al. [2021] for 1970-1999 yields a generalized Jaccard index of 0.41 and Fleiss's kappa of 0.15.

Figure 7: Comparison of generalized jaccard index and fleiss' kappa across periods



Source: Author's calculation.

These results indicate that there is significant variability in the identification of crises in data sets, influenced by differing definitions and identification criteria. Although agreement improves in more recent periods due to enhanced data availability and quality, substantial discrepancies persist.

6 Discrepancies in Crisis Dating

The precision of crisis dating is a significant challenge widely discussed in the literature, with discrepancies in dates and episode lengths frequently observed across data sets (see, for example, Chaudron and de Haan [2014], Boyd et al. [2019] and Sufi and Taylor [2022]).

Among the data sets initially considered, Baron et al. [2021] and Jordà et al. [2017] are excluded from the analysis because they only provide the start date of each crisis. Furthermore, Caprio and Klingebiel [2002] identifies some crises using ranges or approximate periods rather than precise years, complicating comparisons of crisis dates.[9]

To evaluate temporal discrepancies, each episode is analyzed individually, focusing on the data sets that identify it. The analysis determines the number of data sets in which the start dates coincide,[10] the number in which the end dates coincide, and the number in which the start and end dates align. These counts are used to calculate the proportion of temporal agreement, serving as a measure of consistency in the timing between the data sets.

Data sets tend to align more closely on the start dates of crises, whereas greater variability is observed in their end dates. Approximately 70% of events exhibit perfect alignment in start dates in all data sets, compared to only 36% for end dates. The complete alignment of the start and end dates is observed in only 29% of the events identified by at least two data sets, while 6% of the events show no alignment in the dates.

Figure 8 examines the correlation between Region, Continent, Income Level, and the proportion of data sets in which episodes are dated within the same period. The results indicate a positive correlation between high-income countries and consistency in crisis dating, suggesting that crises in wealthier nations are documented more systematically and with greater agreement between sources. In contrast, negative correlations are observed for lower-middle-income countries and for nations in Africa, particularly in the Middle East and North Africa (MENA) region. This indicates a lower consistency in the identification of crises in these regions and income groups, reflecting potential biases in the dating of crises between data sets.

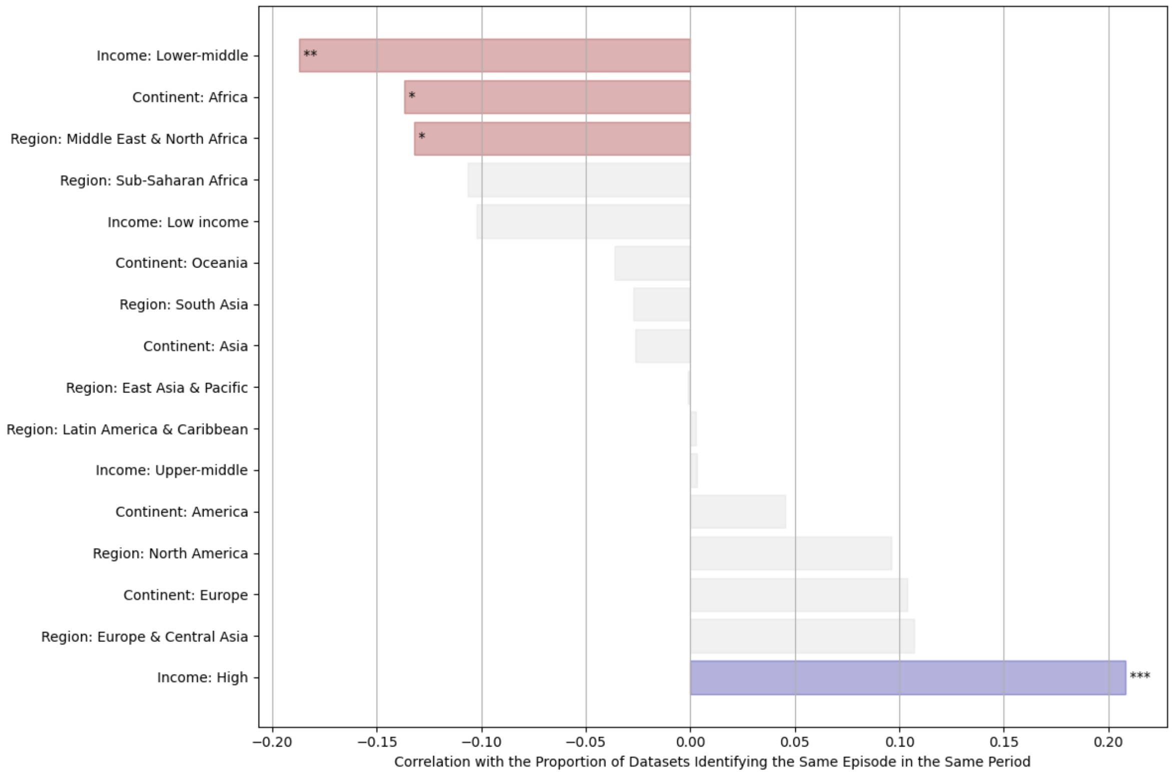
7 Effects of Date Discrepancies on Crisis Duration

Date discrepancies between data sets have a direct impact on the characterization of crisis durations. Variations in start and end dates influence the perceived duration of a crisis, affecting analyses of the severity of the crisis and recovery timelines. The

[9] For episodes without precise dates, the observations are excluded.

[10] Since partial alignment may occur among data sets, the analysis considers the maximum number of data sets with aligned dates.

Figure 8: Correlation Between Data set Alignment and Regional or Income Characteristics



Source: Author's calculations. Significance levels reported are $p < .05$, **** $p < .01$, ***** $p < .001$.

summary statistics for the duration of the crisis in each data set are presented in Table 2.

The duration of the crisis varies significantly across the data sets. Most data sets suggest that crises typically last between 2 and 6 years, but there are notable differences in their distributions. For example, the Laeven and Valencia [2010] data set enforces a maximum duration rule of five years, resulting in a conservative approach with the lowest variability. In contrast, Reinhart and Rogoff [2009] and Demirgüç-Kunt and Detragiache [2005] allow for longer crisis periods, with maximum durations reaching 15 and 13 years, respectively.

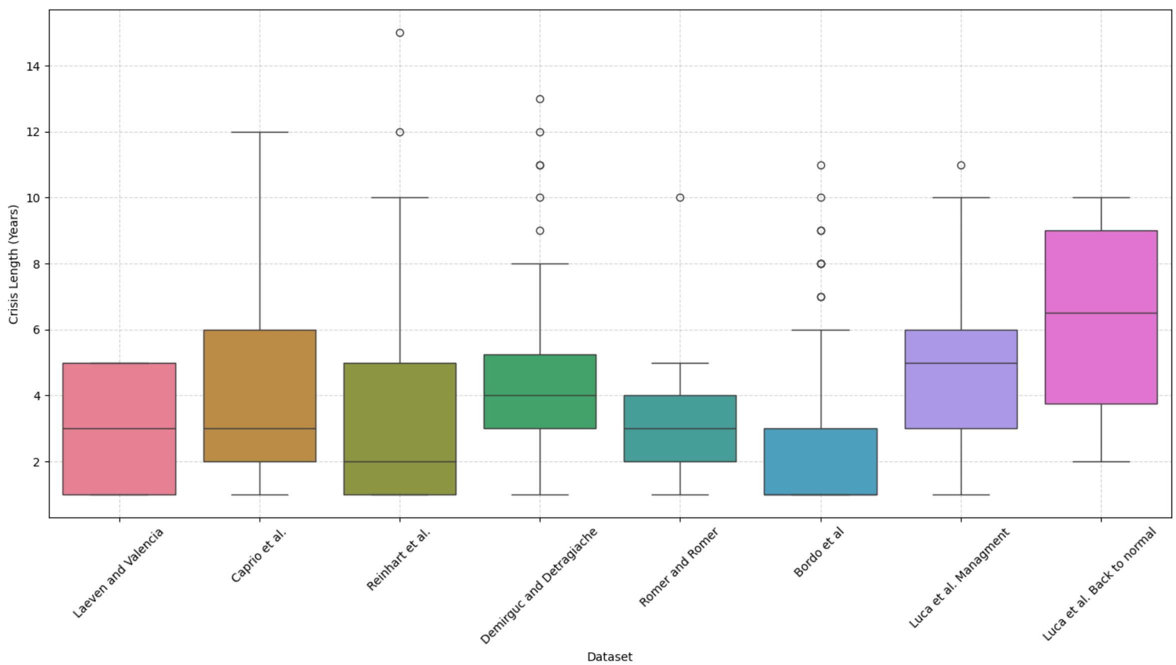
The data set Lo Duca et al. [2017] presents the longest mean duration of the crisis, at 6.25 years when considering macrofinancial recovery and 4.91 years when considering policy management. In contrast, Bordo et al. [2001] reports the shortest average

Table 2: Summary Statistics of Crisis Lengths Across Data sets

Data set	Mean	Std Dev	Min	25%	50%	75%	Max
Laeven and Valencia [2020]	3.80	1.65	1	1	3	5	5
Caprio and Klingebiel [2002]	4.60	2.68	1	2	3	6	12
Reinhart and Rogoff [2009]	3.17	2.60	1	1	2	5	15
Demirgüç-Kunt and Detragiache [2005]	4.44	2.66	1	3	4	5.25	13
Romer and Romer [2017]	3.20	1.98	1	2	3	4	10
Bordo et al. [2001]	2.42	2.20	1	1	1	3	11
Lo Duca et al. [2017] (mgmt.)	4.91	2.45	1	3	5	6	11
Lo Duca et al. [2017] (recovery)	6.25	2.82	2	3	6	9	10

Source: Author's calculations.

Figure 9: Length of Crisis by Data set

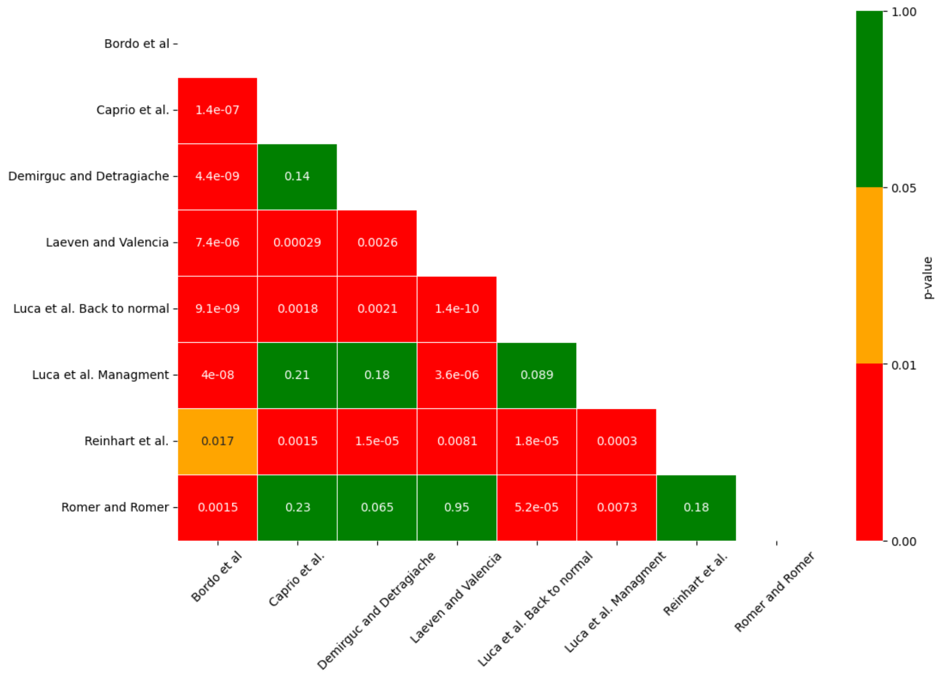


Source: Author's calculations.

duration of the crisis, at 2.42 years. These differences are also reflected in standard deviations, ranging from 1.65 to 2.82 years. All data sets exhibit positive skewness, with means consistently exceeding medians, indicating the presence of prolonged crises that skew the average durations upward. A box plot of the duration of the crises in the data sets is shown in Figure 9.

Pairwise comparisons of crisis length distributions across data sets, using the Kolmogorov-Smirnov (KS) test, reveal significant differences. The results of the KS test are visualized in Figure 10 as a heat map of the p-values, highlighting the variability in the distributions of the duration of the crisis.

Figure 10: Heat Map of KS Test p-Values for Pairwise Comparisons of Crisis Lengths



Source: Author's calculations.

The heat map indicates that many data sets exhibit statistically significant differences in the crisis length distributions. Red cells ($p < 0.01$) denote highly significant differences, while orange cells ($0.01 \leq p < 0.05$) indicate moderate significance. Green cells ($p \geq 0.05$) reflect non-significant differences, suggesting similar crisis length distributions. The high prevalence of red cells underscores substantial heterogeneity in the way data sets measure and define the duration of crises.

8 Data set integration and its impact on crisis duration

The results of the unified approach to start and end dates for identified crises are presented in Table 3. This methodology combines information from multiple data sets using a majority vote rule to establish a single set of unified crisis dates. For crises identified by only one data set, the start and end dates remain as originally reported. For crises identified by more than one data set, the start and end dates with the most votes among the data sets are selected. When multiple start or end dates receive the same number of votes, the earliest start date and the latest end date are chosen to ensure the inclusion of the entire relevant period. This approach avoids imposing the Laeven and Valencia [2020] five-year maximum criterion, allowing for a longer crisis duration that researchers can further explore using additional indicators.

Table 3: Summary Statistics of Unified Crisis Lengths

Statistic	Count	Mean	Std. Dev.	Min	25%	Median	Max
Unified Crisis Length	251	3.48	2.66	1	1	3	12

According to unified criteria, the average duration of the crisis is 3.48 years, with a median of 3 years, indicating that half of the crises are resolved within this period. However, there is notable variability, with a minimum crisis duration of 1 year and a maximum of 12 years. The high maximum reflects the broader time horizon captured by the unified approach in cases where the data sets do not converge on a single start or end date.

The unified data set comprises 251 unique episodes of systemic banking crises. Among these, 108 episodes retain their original dates without any changes. For 95 episodes, only one date changes, with 15 cases involving a start date change and 80 involving an end date change. In addition, 48 episodes experience simultaneous changes on the start and end dates. This process balances the need for consistency with the flexibility to accommodate different perspectives on the duration of crises between data sets. The resulting dates for each episode are detailed in Appendix.

9 Conclusions

The comprehensive analysis of systemic and borderline banking crisis data sets reveals considerable variations in definitions, temporal coverage, and episode identification methods. These differences significantly affect the characterization of crises, influencing the number of detected events, their duration, and the observed macroeconomic and financial dynamics. Although some alignment is observed in the identification of the onset of crises, substantial discrepancies persist, particularly in determining end dates and the overall duration of crisis episodes. Such inconsistencies are not uniformly distributed, as they are influenced by regional and income-level biases, with high-income countries showing greater consistency in crisis identification compared to lower-middle-income countries and regions such as Africa and the MENA area.

These variations pose challenges for researchers and policy makers who rely on these data sets to model crises or formulate policy conclusions. The use of the generalized Jaccard index and Fleiss's Kappa highlight significant disagreements in episode detection, driven by discrepancies in crisis definitions or identification methods. This underscores the need for standardized approaches to crisis definitions and dating criteria to ensure more consistent and reliable analyzes. The proposed majority voting rule to unify the date of the crisis in all data sets offers a pathway to harmonizing the findings, providing a clearer picture of the average durations of the crisis and the periods to consider when evaluating episodes of systemic banking crisis.

Ultimately, this analysis emphasizes the importance of recognizing data set-specific biases and limitations when researching financial crises. Careful consideration of the coverage, definitions, and methodologies employed in different data sets is essential to draw robust conclusions about these episodes.

References

Matthew Baron, Emil Verner, and Wei Xiong. Identifying banking crises. *Princeton university manuscript.–2018.–63 p*, 2018.

Matthew Baron, Emil Verner, and Wei Xiong. Banking crises without panics. *The Quarterly Journal of Economics*, 136(1):51–113, 2021.

Michael Bordo, Barry Eichengreen, Daniela Klingebiel, and Maria Soledad Martinez-Peria. Is the crisis problem growing more severe? *Economic policy*, 16(32):52–82, 2001.

John H Boyd, Gianni De Nicolo, and Tatiana Rodionova. Banking crises and crisis dating: Disentangling shocks and policy responses. *Journal of Financial Stability*, 41:45–54, 2019.

Gerard Caprio and Daniela Klingebiel. Bank insolvency: bad luck, bad policy, or bad banking? In *Annual World Bank conference on development economics*, volume 79, pages 1–26, 1996.

Gerard Caprio and Daniela Klingebiel. Episodes of systemic and borderline financial crises. *World Bank. October.-1999 [In English]*, 2002.

Raymond Chaudron and Jakob de Haan. Dating banking crises using incidence and size of bank failures: Four crises reconsidered. *Journal of Financial Stability*, 15:63–75, 2014.

Luciano da F Costa. Further generalizations of the jaccard index. *arXiv preprint arXiv:2110.09619*, 2021.

Asli Demirgüç-Kunt and Enrica Detragiache. The determinants of banking crises in developing and developed countries. *Staff Papers*, 45(1):81–109, 1998.

Asli Demirgüç-Kunt and Enrica Detragiache. Cross-country empirical studies of systemic bank distress: a survey. *National institute economic review*, 192:68–83, 2005.

Reza Drikvandi and Olufemi Lawal. Sparse principal component analysis for natural language processing. *Annals of Data Science*, 10:25–41, 2023. doi: 10.1007/s40745-020-00277-x. URL https://doi.org/10.1007/s40745-020-00277-x.

Joseph L Fleiss, Bruce Levin, and Myunghee Cho Paik. *Statistical methods for rates and proportions*. john wiley & sons, 2013.

Zhongbo Jing, Jakob de Haan, Jan Jacobs, and Haizhen Yang. Identifying banking crises using money market pressure: New evidence for a large set of countries. *Journal of Macroeconomics*, 43:1–20, 2015.

Òscar Jordà, Moritz Schularick, and Alan M. Taylor. Macrofinancial history and the new business cycle facts. In Martin Eichenbaum and Jonathan A. Parker, editors, *NBER Macroeconomics Annual 2016*, volume 31. University of Chicago Press, Chicago, 2017.

Seung-Woo Kim and Jae-Min Gil. Research paper classification systems based on tf-idf and lda schemes. *Human-Centric Computing and Information Sciences*, 9 (30):1–15, 2019. doi: 10.1186/s13673-019-0192-7. URL https://doi.org/10.1186/s13673-019-0192-7.

Naveen Kumar, Sanjay Kumar Yadav, and Divakar Singh Yadav. An approach for documents clustering using k-means algorithm. In Pradeep Kumar Singh, Zdzislaw Polkowski, Sudeep Tanwar, Sunil Kumar Pandey, Gheorghe Matei, and Daniela Pirvu, editors, *Innovations in Information and Communication Technologies (IICT-2020)*, pages 453–460, Cham, 2021. Springer International Publishing. ISBN 978-3-030-66218-9.

Luc Laeven and Fabian Valencia. Systemic banking crises database. *IMF Economic Review*, 61(2):225–270, 2013.

Luc Laeven and Fabian Valencia. Systemic banking crises database ii. *IMF Economic Review*, 68:307–361, 2020.

Mr Luc Laeven and Mr Fabian Valencia. *Resolution of banking crises: The good, the bad, and the ugly*. International Monetary Fund, 2010.

Mr Luc Laeven and Mr Fabian Valencia. *Systemic banking crises revisited*. International Monetary Fund, 2018.

Marco Lo Duca, Anne Koban, Marisa Basten, Elias Bengtsson, Benjamin Klaus, Piotr Kusmierczyk, Jan Hannes Lang, Carsten Detken, and Tuomas Peltonen. A new database for financial crises in european countries. *ECB occasional paper*, (2017/194), 2017.

Srinivas Mekala and B. Padmaja Rani. Dimensionality reduction in natural language text document using pca techniques. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 5(12):636–642, 2018. ISSN 2349-5162. URL http://www.jetir.org/papers/JETIR1812088.pdf. December 2018.

Thanh Cong Nguyen, Vítor Castro, and Justine Wood. A new comprehensive database of financial crises: Identification, frequency, and duration. *Economic*

Modelling, 108:105770, 2022. ISSN 0264-9993. doi: https://doi.org/10.1016/j.econmod.2022.105770. URL https://www.sciencedirect.com/science/article/pii/S0264999322000165.

Carmen Reinhart, Ken Rogoff, Christoph Trebesch, and Vincent Reinhart. Banking crisis dates and other financial crisis data series. https://www.hbs.edu/behavioral-finance-and-financial-stability/data/Pages/global.aspx, 2016. Data available as of 2024 from the Behavioral Finance Financial Stability Project at Harvard Business School. Last updated in 2016.

Carmen M Reinhart and Kenneth S Rogoff. This time is different: A panoramic view of eight centuries of financial crises. Technical report, National Bureau of Economic Research, 2008.

Carmen M Reinhart and Kenneth S Rogoff. *This time is different: Eight centuries of financial folly*. princeton university press, 2009.

J. Rejito, A. Atthariq, and A. S. Abdullah. Application of text mining employing k-means algorithms for clustering tweets of tokopedia. *Journal of Physics: Conference Series*, 1722:012019, 2021. doi: 10.1088/1742-6596/1722/1/012019. Published under licence by IOP Publishing Ltd. Tenth International Conference and Workshop on High Dimensional Data Analysis (ICW-HDDA-X) 12-15 October 2020 in Sanur-Bali, Indonesia.

Christina D Romer and David H Romer. New evidence on the impact of financial crises in advanced countries. Technical report, National Bureau of Economic Research, 2015.

Christina D Romer and David H Romer. New evidence on the aftermath of financial crises in advanced countries. *American Economic Review*, 107(10):3072–3118, 2017.

Moritz Schularick and Alan M Taylor. Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870–2008. *American Economic Review*, 102(2):1029–1061, 2012.

Hinrich Schütze, Christopher D Manning, and Prabhakar Raghavan. *Introduction to information retrieval*, volume 39. Cambridge University Press Cambridge, 2008.

Walmir Silva, Herbert Kimura, and Vinicius Amorim Sobreiro. An analysis of the literature on systemic financial risk: A survey. *Journal of Financial Stability*, 28: 91–114, 2017.

N. V. Smirnov. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, 19(2):279–281, 1948. doi: 10.1214/aoms/1177730256. URL https://doi.org/10.1214/aoms/1177730256.

Springer. *Introduction to data mining*. Springer, 2006.

Amir Sufi and Alan M Taylor. Financial crises: A survey. *Handbook of international economics*, 6:291–340, 2022.

Jürgen Von Hagen and Tai-kuang Ho. Money market pressure and the determinants of banking crises. *Journal of Money, Credit and Banking*, 39(5):1037–1066, 2007.

A Country coverage across data sets

Country	Continent	LV	CK	BEKM	RRTR	DD	JST	RR	BVX	LD
Afghanistan	Asia	0		0	0	0	0	0	0	0
Albania	Europe	1	1	0	0	0	0	0	0	0
Algeria	Africa	1	1	0	1	1	0	0	0	0
Andorra	Europe	0		0	0	0	0	0	0	0
Angola	Africa	1	1	0	1	0	0	0	0	0
Antigua And Barbuda	America	0		0	0	0	0	0	0	0
Argentina	America	1	1	1	1	1	0	0	1	0
Armenia	Asia	1	1	0	0	0	0	0	0	0
Australia	Oceania	1	1	1	1	1	1	1	1	0
Austria	Europe	1		1	1	1	0	1	1	1
Azerbaijan	Asia	1	1	0	0	0	0	0	0	0
Bahamas	America	0		0	0	0	0	0	0	0
Bahrain	Asia	0		0	0	1	0	0	0	0
Bangladesh	Asia	1	1	1	0	0	0	0	0	0
Barbados	America	1		0	0	0	0	0	0	0
Belarus	Europe	1	1	0	0	0	0	0	0	0
Belgium	Europe	1		1	1	1	1	1	1	1
Belize	America	1		0	0	0	0	0	0	0
Benin	Africa	1	1	0	0	1	0	0	0	0
Bhutan	Asia	1		0	0	0	0	0	0	0
Bolivia	America	1	1	0	1	1	0	0	0	0
Bosnia And Herzegovina	Europe	1	1	0	0	0	0	0	0	0
Botswana	Africa	1	1	0	0	0	0	0	0	0
Brazil	America	1	1	1	1	1	0	0	1	0
Brunei	Asia	1	1	0	0	0	0	0	0	0
Bulgaria	Europe	1	1	0	0	0	0	0	0	1
Burkina Faso	Africa	1	1	0	0	1	0	0	0	0
Burundi	Africa	1	1	0	0	1	0	0	0	0
Cambodia	Asia	1		0	0	0	0	0	0	0
Cameroon	Africa	1	1	0	0	1	0	0	0	0
Canada	America	1	1	1	1	1	1	1	1	0
Cape Verde	Africa	1	1	0	0	0	0	0	0	0
Central African Republic	Africa	1	1	0	1	1	0	0	0	0
Chad	Africa	1	1	0	0	1	0	0	0	0
Chile	America	1	1	1	1	1	0	0	1	0
China	Asia	1	1	1	1	0	0	0	0	0
Colombia	America	1	1	1	1	1	0	0	1	0
Comoros	Africa	1		0	0	0	0	0	0	0
Congo (congo-brazzaville)	Africa	1	1	0	0	1	0	0	0	0
Costa Rica	America	1	1	1	1	1	0	0	0	0
Croatia	Europe	1	1	0	0	0	0	0	0	1

Continued on next page

Country Coverage Across Various Datasets (continued)

Country	Continent	LV	CK	BEKM	RRTR	DD	JST	RR	BVX	LD
Cuba	America	0		0	0	0	0	0	0	0
Cyprus	Europe	1		0	0	1	0	0	0	1
Czechia (czech Republic)	Europe	1	1	0	0	0	0	0	1	1
Côte D'ivoire	Africa	1	1	1	1	1	0	0	0	0
Democratic Republic Of The Congo	Africa	1	1	0	0	1	0	0	0	0
Denmark	Europe	1	1	1	1	1	1	1	1	1
Djibouti	Africa	1	1	0	0	0	0	0	0	0
Dominica	America	1		0	0	0	0	0	0	0
Dominican Republic	America	1		0	1	0	0	0	0	0
Ecuador	America	1	1	1	1	1	0	0	0	0
Egypt	Africa	1	1	1	1	1	0	0	1	0
El Salvador	America	1	1	0	1	1	0	0	0	0
Equatorial Guinea	Africa	1	1	0	0	0	0	0	0	0
Eritrea	Africa	1	1	0	0	0	0	0	0	0
Estonia	Europe	1	1	0	0	0	0	0	0	1
Ethiopia	Africa	1	1	0	0	0	0	0	0	0
Fiji	Oceania	1		0	0	0	0	0	0	0
Finland	Europe	1	1	1	1	1	1	1	1	1
France	Europe	1	1	1	1	1	1	1	1	1
Gabon	Africa	1	1	0	0	0	0	0	0	0
Gambia	Africa	1	1	0	0	0	0	0	0	0
Georgia	Asia	1	1	0	0	0	0	0	0	0
Germany	Europe	1	1	1	1	1	1	1	1	1
Ghana	Africa	1	1	1	1	1	0	0	0	0
Greece	Europe	1	1	1	1	1	0	1	1	1
Grenada	America	1		0	0	0	0	0	0	0
Guatemala	America	1	1	0	1	1	0	0	0	0
Guinea	Africa	1	1	0	0	1	0	0	0	0
Guinea-bissau	Africa	1	1	0	0	1	0	0	0	0
Guyana	America	1		0	0	1	0	0	0	0
Haiti	America	1		0	0	0	0	0	0	0
Honduras	America	1		0	1	1	0	0	0	0
Hong Kong	Asia	0	1	1	0	0	0	0	1	0
Hungary	Europe	1	1	0	1	0	0	0	1	1
Iceland	Europe	1	1	1	1	0	0	1	1	0
India	Asia	1	1	1	1	1	0	0	1	0
Indonesia	Asia	1	1	1	1	1	0	0	1	0
Iran	Asia	1		0	0	0	0	0	0	0
Iraq	Asia	0		0	0	0	0	0	0	0
Ireland	Europe	1		1	1	1	1	1	1	1
Israel	Asia	1	1	1	0	1	0	0	1	0
Italy	Europe	1	1	1	1	1	1	1	1	1

Continued on next page

Country Coverage Across Various Datasets (continued)

Country	Continent	LV	CK	BEKM	RRTR	DD	JST	RR	BVX	LD
Jamaica	America	1	1	1	0	1	0	0	0	0
Japan	Asia	1	1	1	1	1	1	1	1	0
Jordan	Asia	1	1	0	0	1	0	0	0	0
Kazakhstan	Asia	1		0	0	0	0	0	0	0
Kenya	Africa	1	1	0	1	1	0	0	0	0
Kiribati	Oceania	0		0	0	0	0	0	0	0
Kuwait	Asia	1	1	0	0	0	0	0	0	0
Kyrgyzstan	Asia	1	1	0	0	0	0	0	0	0
Laos	Asia	1	1	0	0	0	0	0	0	0
Latvia	Europe	1	1	0	0	0	0	0	0	1
Lebanon	Asia	1	1	0	0	1	0	0	0	0
Lesotho	Africa	1	1	0	0	0	0	0	0	0
Liberia	Africa	1	1	0	0	1	0	0	0	0
Libya	Africa	1		0	0	0	0	0	0	0
Liechtenstein	Europe	0		0	0	0	0	0	0	0
Lithuania	Europe	1	1	0	0	0	0	0	0	1
Luxembourg	Europe	1		0	0	0	0	1	1	1
Macedonia	Europe	1	1	0	0	0	0	0	0	0
Madagascar	Africa	1	1	0	0	1	0	0	0	0
Malawi	Africa	1		0	0	0	0	0	0	0
Malaysia	Asia	1	1	1	1	1	0	0	1	0
Maldives	Asia	1		0	0	0	0	0	0	0
Mali	Africa	1	1	0	0	1	0	0	0	0
Malta	Europe	0		0	0	0	0	0	0	1
Marshall Islands	Oceania	0		0	0	0	0	0	0	0
Mauritania	Africa	1	1	0	0	1	0	0	0	0
Mauritius	Africa	1	1	0	1	0	0	0	0	0
Mexico	America	1	1	1	1	1	0	0	1	0
Micronesia	Oceania	0		0	0	0	0	0	0	0
Moldova	Europe	1		0	0	0	0	0	0	0
Monaco	Europe	0		0	0	0	0	0	0	0
Mongolia	Asia	1		0	0	0	0	0	0	0
Montenegro	Europe	0		0	0	0	0	0	0	0
Morocco	Africa	1	1	0	1	0	0	0	0	0
Mozambique	Africa	1	1	0	0	0	0	0	0	0
Myanmar	Asia	1	1	0	1	0	0	0	0	0
Namibia	Africa	1		0	0	0	0	0	0	0
Nauru	Oceania	0		0	0	0	0	0	0	0
Nepal	Asia	1	1	0	0	1	0	0	0	0
Netherlands	Europe	1		1	1	1	1	1	1	1
New Caledonia	Oceania	1		0	0	0	0	0	0	0
New Zealand	Oceania	1	1	1	1	1	0	1	1	0

Continued on next page

Country Coverage Across Various Datasets (continued)

Country	Continent	LV	CK	BEKM	RRTR	DD	JST	RR	BVX	LD
Nicaragua	America	1	1	0	1	0	0	0	0	0
Niger	Africa	1	1	0	0	1	0	0	0	0
Nigeria	Africa	1	1	1	1	1	0	0	0	0
North Korea	Asia	0		0	0	0	0	0	0	0
Norway	Europe	1	1	1	1	1	1	1	1	1
Oman	Asia	0		0	0	0	0	0	0	0
Pakistan	Asia	1		1	0	0	0	0	0	0
Palau	Oceania	0		0	0	0	0	0	0	0
Panama	America	1	1	0	1	1	0	0	0	0
Papua New Guinea	Oceania	1	1	0	0	1	0	0	0	0
Paraguay	America	1	1	1	1	1	0	0	0	0
Peru	America	1	1	1	1	1	0	0	1	0
Philippines	Asia	1	1	1	1	1	0	0	1	0
Poland	Europe	1	1	0	1	0	0	0	0	1
Portugal	Europe	1		1	1	1	1	1	1	1
Qatar	Asia	0		0	0	0	0	0	0	0
Romania	Europe	1	1	0	1	0	0	0	0	1
Russia	Europe	1	1	0	1	0	0	0	1	0
Rwanda	Africa	1	1	0	0	0	0	0	0	0
Saint Kitts And Nevis	America	1		0	0	0	0	0	0	0
Saint Lucia	America	0		0	0	0	0	0	0	0
Saint Vincent And The Grenadines	America	0		0	0	0	0	0	0	0
Samoa	Oceania	0		0	0	0	0	0	0	0
San Marino	Europe	0		0	0	0	0	0	0	0
Sao Tome And Principe	Africa	1	1	0	0	0	0	0	0	0
Saudi Arabia	Asia	0		0	0	0	0	0	0	0
Senegal	Africa	1	1	1	0	1	0	0	0	0
Serbia	Europe	1		0	0	0	0	0	0	0
Seychelles	Africa	1		0	0	1	0	0	0	0
Sierra Leone	Africa	1	1	0	0	1	0	0	0	0
Singapore	Asia	1	1	1	1	1	0	0	1	0
Slovakia	Europe	1	1	0	0	0	0	0	0	1
Slovenia	Europe	1	1	0	0	0	0	0	0	1
Solomon Islands	Oceania	0		0	0	0	0	0	0	0
Somalia	Africa	0		0	0	0	0	0	0	0
South Africa	Africa	1	1	1	1	1	0	0	1	0
South Korea	Asia	1	1	1	1	1	0	0	1	0
South Sudan	Africa	1		0	0	0	0	0	0	0
Spain	Europe	1	1	1	1	0	1	1	1	1
Sri Lanka	Asia	1	1	1	1	1	0	0	0	0
Sudan	Africa	1		0	0	0	0	0	0	0
Suriname	America	1		0	0	0	0	0	0	0

Continued on next page

Country Coverage Across Various Datasets (continued)

Country	Continent	LV	CK	BEKM	RRTR	DD	JST	RR	BVX	LD
Swaziland	Africa	1	1	0	0	1	0	0	0	0
Sweden	Europe	1	1	1	1	1	1	1	1	1
Switzerland	Europe	1		1	1	1	1	1	1	0
Syria	Asia	1		0	0	1	0	0	0	0
Taiwan	Asia	0	1	1	1	1	0	0	1	0
Tajikistan	Asia	1	1	0	0	0	0	0	0	0
Tanzania	Africa	1	1	0	0	1	0	0	0	0
Thailand	Asia	1	1	1	1	1	0	0	1	0
Timor-leste	Asia	0		0	0	0	0	0	0	0
Togo	Africa	1	1	0	0	1	0	0	0	0
Tonga	Oceania	0		0	0	0	0	0	0	0
Trinidad And Tobago	America	1	1	0	0	0	0	0	0	0
Tunisia	Africa	1	1	0	1	1	0	0	0	0
Turkey	Asia	1	1	1	1	1	0	1	1	0
Turkmenistan	Asia	1		0	0	0	0	0	0	0
Tuvalu	Oceania	0		0	0	0	0	0	0	0
Uganda	Africa	1	1	0	0	1	0	0	0	0
Ukraine	Europe	1	1	0	0	0	0	0	0	0
United Arab Emirates	Asia	0		0	0	0	0	0	0	0
United Kingdom	Europe	1	1	1	1	1	1	1	1	0
United States Of America	America	1	1	1	1	1	1	1	1	0
Uruguay	America	1	1	1	1	1	0	0	0	0
Uzbekistan	Asia	1		0	0	0	0	0	0	0
Vanuatu	Oceania	0		0	0	0	0	0	0	0
Venezuela	America	1	1	1	1	1	0	0	1	0
Vietnam	Asia	1	1	0	0	0	0	0	0	0
Yemen	Asia	1	1	0	0	0	0	0	0	0
Zambia	Africa	1	1	0	1	1	0	0	0	0
Zimbabwe	Africa	1	1	1	1	0	0	0	0	0

References: LV: Laeven and Valencia [2020], CK: Caprio and Klingebiel [2002], BEKM: Bordo et al. [2001], RRTR: Reinhart and Rogoff [2009], DD: Demirgüç-Kunt and Detragiache [2005], JST: Jordà et al. [2017], RR: Romer and Romer [2017], BVX: Baron et al. [2021], LD: Lo Duca et al. [2017]. 1 if coutry is in the dataset and 0 if it is not considered.

B Overlap of Data set Coverage by Period

Start Year	End Year	Datasets
1800	1869	Reinhart et al. [2016]
1870	1879	Reinhart et al. [2016], Baron et al. [2021], Jordà et al. [2017]
1880	1966	Reinhart et al. [2016], Bordo et al. [2001], Baron et al. [2021], Jordà et al. [2017]
1967	1969	Reinhart et al. [2016], Romer and Romer [2017], Bordo et al. [2001], Baron et al. [2021], Jordà et al. [2017]
1970	1979	Reinhart et al. [2016], Romer and Romer [2017], Bordo et al. [2001], Baron et al. [2021], Jordà et al. [2017], Lo Duca et al. [2017],Laeven and Valencia [2020] and Nguyen et al. [2022], Caprio and Klingebiel [2002]
1980	1994	Reinhart et al. [2016], Romer and Romer [2017], Bordo et al. [2001],Demirgüç-Kunt and Detragiache [2005], Baron et al. [2021], Jordà et al. [2017], Lo Duca et al. [2017],Laeven and Valencia [2020] and Nguyen et al. [2022], Caprio and Klingebiel [2002]
1995	1997	Laeven and Valencia [2020] and Nguyen et al. [2022] Caprio and Klingebiel [2002], Reinhart et al. [2016], Romer and Romer [2017], Baron et al. [2021], Jordà et al. [2017], Lo Duca et al. [2017]
1998	1999	Caprio and Klingebiel [2002], Laeven and Valencia [2020] and Nguyen et al. [2022], Reinhart et al. [2016], Romer and Romer [2017], Baron et al. [2021], Lo Duca et al. [2017]
2017	2019	Laeven and Valencia [2020] and Nguyen et al. [2022]

C Definitions

Table 6: Definitions and identification criteria for systemic banking crises

Author(s)	**Definition**	**Identification Criteria**	**Start and End Date Considerations**	**Classification**
Laeven and Valencia [2020]	Defines systemic banking crises as highly disruptive events that "lead to sustained declines in economic activity, financial intermediation, and ultimately in welfare." Two conditions must be met for an event to qualify as a banking crisis: 1. "Significant signs of financial distress in the banking system (evidenced by bank runs, losses in the banking system, and/or bank liquidations)." 2. "Significant banking policy interventions in response to significant losses in the banking system." When one of these criteria is severe enough, it acts as a sufficient condition to identify an episode.	Policy interventions and financial thresholds. Policy measures considered include deposit freezes and/or bank holidays, significant bank nationalizations, bank restructuring with fiscal costs of at least 3% of GDP, extensive liquidity support of at least 5% of deposits and liabilities to non-residents, guarantees, and asset purchases of at least 5% of GDP. Losses are severe if the non-performing loan ratio is at least 20%.	Start: Identified by policy interventions or financial variables. End: The year before both real GDP and credit growth are positive for at least two consecutive years.	Objective Measure

Continued on next page

Table 6 – continued from previous page

Author(s)	Definition	Identification Criteria	Start and End Date Considerations	Classification
Caprio and Klingebiel [1996, 2002]	Identifies episodes where much or all of the banking capital is depleted. Although systemic and borderline crises are distinguished, no detailed criteria for differentiation are provided. Episodes are identified based on negative net capital assessed by experts, even when official data indicate positive net capital. The identification relies on expert evaluations and published reports.	Expert assessments and published reports are prioritized over official data, recognizing both overt and underlying financial distress.	Start: Determined by expert consensus. End: Difficult to determine as financial distress may persist without overt signs.	Qualitative

Continued on next page

Table 6 – continued from previous page

Author(s)	Definition	Identification Criteria	Start and End Date Considerations	Classification
Demirgüç-Kunt and Detragiache [1998, 2005]	Defines systemic banking crises as large-scale disruptions in the banking sector. These crises have significant impacts on the real economy, disrupting credit availability and leading to reductions in investment and consumption, firm bankruptcies, and reduced confidence in domestic financial institutions. Such crises may cause capital outflows and affect payment systems, potentially leading to the insolvency of initially sound banks. A systemic crisis arises when a significant portion of the banking system experiences loan losses that exceed its capital.	Fragility episodes identified by previous authors; systemic crises are marked by specific financial distress criteria: NPL $>$ 10%, rescue costs $>$ 2% of GDP, or extensive emergency measures.	Dates are referenced from previous studies.	Combined

Continued on next page

Table 6 – continued from previous page

Author(s)	Definition	Identification Criteria	Start and End Date Considerations	Classification
Reinhart and Rogoff [2008]	Defines a major, systemic banking crisis as an event in which "a significant portion of a nation's banking sector becomes insolvent following heavy investment losses or banking panics." The authors distinguish between systemic crises, involving "bank runs that result in the closure, merger, or takeover by the public sector of one or more financial institutions," and borderline (non-systemic) crises, which occur when "the closure, merger, takeover, or large-scale government intervention in a significant financial institution (or group of institutions) triggers similar outcomes in others." The dataset includes both types of episodes.	Identified using existing banking crisis literature and financial media. If a decline in bank equity begins before the qualitative assessment's identified date, the crisis start date is adjusted.	Start and end dates are based on qualitative assessments and equity declines.	Combined

Continued on next page

Table 6 – continued from previous page

Author(s)	Definition	Identification Criteria	Start and End Date Considerations	Classification
Romer and Romer [2017]	Uses the OECD Economic Outlook to identify levels of financial stress. Mild crises are associated with borderline or high financial stress, while moderate and severe levels are considered systemic. Such crises represent widespread financial difficulties that impact overall economic performance but do not lead to a total breakdown of the financial system. Major and extreme crises involve substantial obstacles to financial intermediation, affecting credit supply and macroeconomic outcomes. Other datasets are also considered in the analysis.	Based on the assessment provided in the OECD Economic Outlook.	Based on the assessment provided in the OECD Economic Outlook.	Qualitative

Continued on next page

Table 6 – continued from previous page

Author(s)	Definition	Identification Criteria	Start and End Date Considerations	Classification
Bordo et al. [2001]	Defines a banking crisis as a period of financial distress leading to the erosion of most or all aggregate banking system capital. The identification criteria align with those outlined by Caprio and Klingebiel [1996, 2002].	Based on Caprio and Klingebiel [1996, 2002] and expert assessments of capital deterioration.	Based on qualitative assessment.	Qualitative
Schularick and Taylor [2012], Jordà et al. [2017]	Defines financial crises as events where a country's banking sector experiences bank runs, sharp increases in default rates, and significant capital losses, resulting in public intervention, bankruptcy, or forced mergers of financial institutions.	Based on qualitative assessment and existing datasets, including Bordo et al. [2001] and Reinhart and Rogoff [2009].	Based on qualitative assessment and existing datasets.	Qualitative

Continued on next page

Table 6 – continued from previous page

Author(s)	Definition	Identification Criteria	Start and End Date Considerations	Classification
Baron et al. [2021]	Does not provide an explicit definition of systemic banking crises but considers a potential banking crisis as an event marked by a "large crash in a country's bank equity index," indicating widespread distress affecting bank equity prices.	A threshold is established at a 30% drop in bank equity.	Start date based on the point when equity drops exceed the threshold. Only start dates are provided.	Objective

Continued on next page

Table 6 – continued from previous page

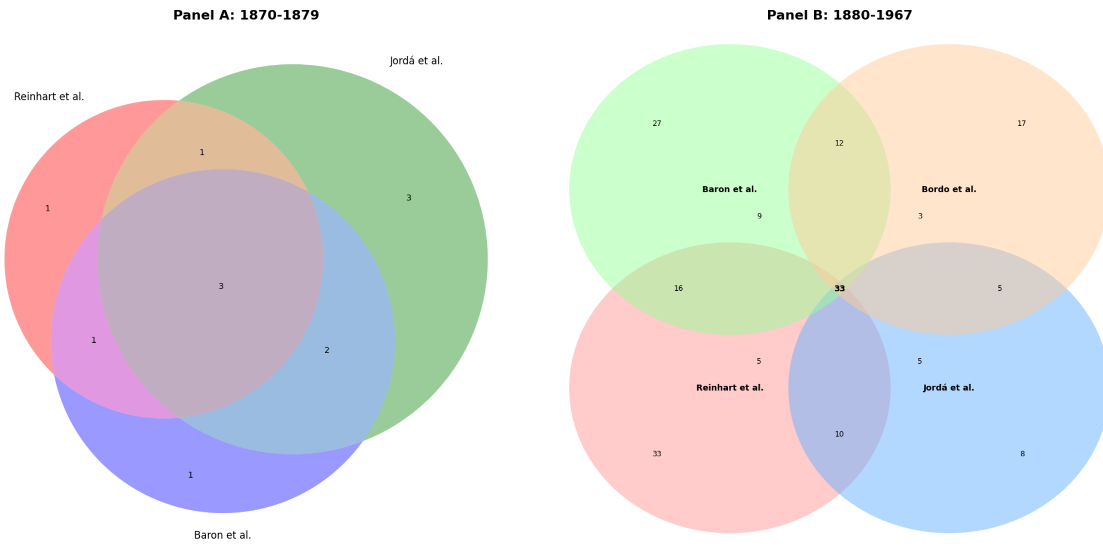
Author(s)	Definition	Identification Criteria	Start and End Date Considerations	Classification
Lo Duca et al. [2017]	Does not provide a specific definition of systemic banking crises, but based on their identification methods, a crisis involves "financial stress associated with negative real economic outcomes." These events include "the financial system amplifying shocks, major bankruptcies, and policy interventions to mitigate issues." Identification uses a two-step approach: quantitative measures of a financial stress index and qualitative assessments.	Combines a financial stress index with expert judgment.	Based on the financial stress index.	Combined

D Word Clouds by cluster

Figure 11: Word Cloud of frequent words by cluster

E Venn diagram for selected periods

Figure 12: Comparison of Financial Crisis Identifications Across Datasets



F Systemic banking crisis dates with majority voter´s rule

article longtable

Country	Unified start date	Unified end date
Albania	1992	1994
Algeria	1990	1992
Angola	1991	1999
Argentina	1890	1891
Argentina	1914	1914
Argentina	1931	1931
Argentina	1934	1934
Argentina	1980	1982
Argentina	1989	1990
Argentina	1995	1995
Argentina	2001	2003
Armenia	1994	1996
Australia	1893	1893
Australia	1989	1992
Austria	2008	2012
Azerbaijan	1995	1995
Bangladesh	1987	1996
Belarus	1995	1995
Belgium	1914	1914
Belgium	1925	1926
Belgium	1931	1931
Belgium	1934	1934
Belgium	1939	1939
Belgium	2008	2014
Benin	1988	1990
Bolivia	1986	1988
Bolivia	1994	1997
Bosnia and Herzegovina	1992	1999
Brazil	1890	1891
Brazil	1897	1897
Brazil	1900	1901
Brazil	1914	1914

Country	Unified Start Date	Unified End Date
Brazil	1923	1923
Brazil	1963	1963
Brazil	1990	1990
Brazil	1994	1999
Bulgaria	1996	1997
Burkina Faso	1988	1994
Burundi	1994	1997
Cameroon	1987	1993
Cameroon	1995	1998
Canada	1923	1923
Canada	1983	1985
Cape Verde	1993	1993
Central African Rep.	1976	1982
Central African Rep.	1988	1999
Chad	1983	1983
Chad	1992	1992
Chile	1889	1890
Chile	1898	1899
Chile	1907	1908
Chile	1914	1914
Chile	1925	1926
Chile	1976	1976
Chile	1981	1983
China	1997	1999
China, P.R.: Hong Kong	1982	1983
China, P.R.: Hong Kong	1983	1986
Colombia	1982	1987
Colombia	1998	2000
Congo, Dem. Rep. of	1983	1983
Congo, Dem. Rep. of	1991	1994
Congo, Dem. Rep. of	1994	2002
Congo, Rep. of	1992	2002
Costa Rica	1987	1997
Costa Rica	1994	1997
Côte d'Ivoire	1988	1991
Croatia	1998	1999
Cyprus	2011	2015

Country	Unified Start Date	Unified End Date
Czech Republic	1991	2000
Denmark	1885	1885
Denmark	1907	1907
Denmark	1914	1914
Denmark	1921	1921
Denmark	1931	1931
Denmark	1987	1992
Denmark	2008	2009
Djibouti	1991	1995
Dominican Republic	2003	2004
Ecuador	1981	1981
Ecuador	1996	2002
Egypt	1981	1983
Egypt	1991	1995
El Salvador	1989	1990
Equatorial Guinea	1983	1985
Eritrea	1993	1993
Estonia	1992	1995
Estonia	1998	1998
Finland	1900	1900
Finland	1921	1921
Finland	1931	1931
Finland	1939	1939
Finland	1991	1994
France	1881	1882
France	1889	1889
France	1907	1907
France	1930	1932
France	1994	1995
France	2008	2009
Georgia	1991	1995
Germany	1901	1901
Germany	1977	1979
Germany	2008	2010
Ghana	1982	1989
Ghana	1997	1999
Greece	1931	1932

Country	**Unified Start Date**	**Unified End Date**
Greece	1991	1995
Greece	2008	2012
Guatemala	1990	1990
Guinea	1985	1985
Guinea	1993	1994
Guinea-Bissau	1995	1998
Guyana	1993	1995
Hungary	1991	1995
Hungary	2008	2014
Iceland	1985	1986
Iceland	2008	2014
India	1993	1999
Indonesia	1992	1995
Indonesia	1994	1994
Indonesia	1997	2002
Ireland	2008	2012
Israel	1977	1983
Italy	1891	1891
Italy	1893	1893
Italy	1907	1907
Italy	1914	1914
Italy	1921	1922
Italy	1930	1931
Italy	1935	1935
Italy	1990	1995
Italy	2008	2014
Jamaica	1994	2000
Japan	1901	1901
Japan	1907	1907
Japan	1917	1917
Japan	1927	1927
Japan	1992	2002
Japan	1997	2001
Jordan	1989	1990
Kenya	1985	1989
Kenya	1992	1995
Korea	1997	2002

Country	Unified Start Date	Unified End Date
Kuwait	1982	1985
Kyrgyz Republic	1995	1999
Latvia	1995	1999
Latvia	2008	2012
Lebanon	1988	1990
Liberia	1991	1995
Lithuania	1995	1996
Luxembourg	2008	2012
Macedonia	1993	1995
Madagascar	1988	1988
Malaysia	1985	1988
Malaysia	1997	2001
Mali	1987	1989
Mauritania	1984	1993
Mauritius	1996	1996
Mexico	1981	1982
Mexico	1995	1997
Morocco	1980	1984
Mozambique	1987	1995
Myanmar	1996	1997
Nepal	1988	1988
Netherlands	1897	1897
Netherlands	1914	1914
Netherlands	1921	1922
Netherlands	1939	1939
Netherlands	2008	2009
New Zealand	1987	1990
Nicaragua	1987	1996
Nicaragua	2000	2001
Niger	1983	1986
Nigeria	1997	1997
Nigeria	1991	1995
Nigeria	2009	2014
Norway	1921	1923
Norway	1987	1993
Norway	2008	2009
Panama	1988	1989

Country	Unified Start Date	Unified End Date
Papua New Guinea	1989	1992
Paraguay	1995	1999
Peru	1983	1990
Peru	1998	2002
Philippines	1981	1987
Philippines	1997	2001
Poland	1992	1995
Portugal	1890	1891
Portugal	1920	1920
Portugal	1923	1923
Portugal	1931	1932
Portugal	2008	2012
Romania	1990	1999
Russia	1995	1995
Russia	1998	1998
Russia	2008	2014
São Tomé and Príncipe	1992	1992
Senegal	1988	1991
Sierra Leone	1990	1999
Singapore	1982	1982
Slovak Republic	1991	2002
Slovenia	1992	1994
Slovenia	2008	2012
South Africa	1977	1977
South Africa	1989	1989
Spain	1920	1925
Spain	1931	1931
Spain	1977	1985
Spain	2008	2012
Sri Lanka	1989	1993
Swaziland	1995	1995
Sweden	1897	1897
Sweden	1907	1907
Sweden	1931	1932
Sweden	1991	1994
Sweden	2008	2009
Switzerland	2008	2009

Country	Unified Start Date	Unified End Date
Taiwan	1997	1998
Taiwan	1983	1984
Taiwan	1995	1995
Tanzania	1987	1991
Thailand	1983	1987
Thailand	1997	2000
Togo	1993	2002
Tunisia	1991	1995
Turkey	1994	1994
Turkey	1982	1985
Turkey	1991	1991
Turkey	2000	2001
Uganda	1994	1999
Ukraine	1997	1999
United Kingdom	1890	1890
United Kingdom	1974	1976
United Kingdom	1984	1984
United Kingdom	1991	1991
United Kingdom	2007	2014
United States	1884	1884
United States	1893	1893
United States	1907	1907
United States	1914	1914
United States	1929	1933
United States	1984	1991
United States	2007	2010
Uruguay	1981	1985
Uruguay	2002	2005
Venezuela	1978	1986
Venezuela	1994	1997
Vietnam	1997	1999
Yemen	1996	1996
Zambia	1995	1998
Zimbabwe	1995	1999