Advance Journal of Econometrics and Finance
Vol-4, Issue-2, 2026

Advance Journal of Econometrics and Finance

Online ISSN
2959-8990
Print ISSN
2959-8982
https://ajeaf.com/index.php/Journal/About

Name of Publisher: SCHOLAR CRAFT EDUCATION & RESEARCH HUB

Review Type: Double Blind Peer Review

Journal Frequency: Quarterly Research Journal

CREDIT RISK AND BANK PERFORMANCE IN EMERGING MARKETS. A DUAL PROTOCOL SYSTEMATIC REVIEW WITH EVIDENCE FROM PAKISTAN

Raza Ali*[1], Dr. Jameel Ahmed[2]

Raza Ali
PhD Candidate, Faculty of Management Sciences,
SZABIST University, Karachi
raza.ali@szabist.pk

Dr. Jameel Ahmed
Assistant Professor, Faculty of Management Sciences,
SZABIST University, Karachi
Additional Director, Financial Stability Department,
State Bank of Pakistan, Karachi

Abstract

This paper sets forth the methodological framework of a systematic literature review investigating the link between credit risk management and bank performance, with a special emphasis on emerging markets and the Islamic banking system within Pakistan. Based on the PRISMA 2020 reporting protocol (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), the literature review process extracted evidence from the initial 1350 identified literature databases, identifying 247 relevant studies after a detailed process of screening and deduplication. An update performed in 2025 and based on an integrative narrative strand, led to a final evidence base of 281 studies, with a chronological scope from the earliest model of credit risk measurement, the Altman Z-score model of 1968, to the use of machine learning techniques in the area in the latest period. This methodology merges systematicity and contextual sensitivity, utilizing Zotero reference management software to support deduplication, thematic labeling, and structured data extraction organized around 6 key themes: credit risk measurement techniques; the relationship between bank performance and risk; macroeconomic influences on the same topic; Islamic banking perspective; institutional specific studies focused on Pakistan; and technical/methodological developments in this topic. Quality assessment was applied along three aspects: technical/methodological robustness, context relevance, and influence in the academic field, adopted from existing management-related systematic review frameworks. This document specifies methods to account for publication bias, geographical bias, conceptual bias, researcher bias, and integrative bias. With documentation of the systematic literature search, selection criteria, screening process, quality assessment methodology, and theme extraction strategy, this paper contributes to the debate over the adaptability of medical systematic review methodologies to a context of information retrieval specific to banking and finance research.

Keywords: systematic literature review, PRISMA 2020, credit risk, bank performance, emerging markets, Pakistan, Islamic banking, thematic coding, quality assessment

1. Introduction

The process of conducting a literature review in a doctoral study cannot simply be a perfunctory exercise of compiling references. Rather, it is, in its purest sense, a methodological undertaking which is the basis of each subsequent empirical claim (Hart, 2018; Webster & Watson, 2002). In credit risk and banking performance literature, an area in which conclusions vary quite widely depending on the institution, regulation, and time period of the study, the task is less of accumulating evidence than of presenting the patterns, conflicts, and unknowns in the existing research in a way that cannot be challenged. This paper outlines in considerable detail the research design that was the basis of conducting a thorough review of 281 papers that cover six decades of research on credit risk measurement, banking performance determination, macro-economic link, and unique phenomena of Islamic banking in Pakistan and South Asia.

This paper has been prompted by the observation that since the global fiscal crisis of 2007-08, the field of credit risk studies on emerging markets has burgeoned, but many reviews of such literature fail to address the underlying methodology. Literature reviews in the field of finance more often tend to take narrative form, where, even though an abundance of useful context may be provided, there is a lack of clarity as to how the studies selected were chosen and weighed (Tranfield et al., 2003). This opacity is problematic, particularly in doctoral study, where the theoretical framework needs to be well-defended by the evidence produced through a credible literature review. Ideally, a researcher reading the thesis should be able to replicate exactly how the research evidence was built on which parameters studies were chosen, and how any biases were tackled.
To account for these concerns, this paper draws on the PRISMA 2020 guideline (Page et al. 2021) to organize the systematic component of the review. Although the PRISMA protocol was initially designed for systematic reviews of health literature, it is increasingly being adopted by authors in the social sciences, management, and finance literature, as a way to ensure a higher level of transparency and replicability in review methodology (Grant & Booth 2009; Greenhalgh et al. 2018). However, direct application of the PRISMA guideline to research on banking and finance research needs significant modifications, given that, unlike medicine, where randomized controlled trials can guide the hierarchy of evidence, the evidence on banking and finance tends to be heterogeneous, consisting of econometric studies, conceptual work, reports from international bodies, regulatory reports, and various working papers with different methodological quality. Strictly applying empirical filters of the PRISMA guideline to research on banking and finance literature could lead to the exclusion of vital theoretical and regulatory studies that are indispensable for explaining credit risk issues in a developing country like Pakistan, where institutions, regulatory forbearance, and political economy profoundly influence the banking sector.

Accordingly, the methodological approach here adopts a two-pronged framework: (i) a systematic, PRISMA-based, review which produced 247 studies out of 1350 identified after systematic deduplication, screening and selection procedures and (ii) an integrative narrative review as an update, carried out in 2025, which produced an additional 34 studies in the form of new articles, reports from State Bank of Pakistan (SBP), international institutions, and theoretical literature excluded on the grounds of empirical filters. The total evidence used for the analysis came to 281 studies. Reference management software Zotero was used to manage deduplication, themes, and tagging.

This methodological undertaking will have practical value beyond that needed for the purpose of a doctoral thesis. The vast body of research within finance today and growing at an ever-increasing rate means that systematic, replicable, and critically aware review procedures are ever more necessary (Petticrew & Roberts, 2006). This paper aims to achieve this by explicating each part of the review process in full detail: from how the Boolean search strings were formed, to how the quality assessment guide was devised, from how the publication and geographic bias were accounted for, to the thematic coding that was used to formulate the synthesis. It should provide a procedural template that other studies of credit risk, banking stability, and financial performance in emerging economies could use and expand upon.

The rest of the paper continues as follows: the following section explains why the two-stage systematic-integrative approach was used, explaining both its philosophical and methodological basis. Section 3 details how the review was searched, including time, database sources, and keyword composition. Section 4 defines the inclusion and exclusion criteria. Section 5 provides an outline of the PRISMA screening and flow diagram. Section 6 outlines the quality assessment framework used in the study. Section 7 details the extraction and coding of data. Section 8 describes the bias mitigation steps that were taken. Section 9 illustrates how the literature provided the basis for the empirical research framework. Section 10 reflects upon trends in methodological issues in the literature review process in finance research.

This paper has three main contributions. The first is the modification of PRISMA 2020 to the literature on banking and finance by considering the differences in sources of evidence. The second is the development of a systematic-integrative dual review method that allows for the rigor of systematic reviewing combined with the inclusivity of context. Finally, it presents an in-depth and context-specific synthesis of the literature on credit risk and bank performance in a way that is particularly relevant for Pakistan and developing economies, while explicitly establishing a link between the literature base and the empirical research design.

2. Philosophical and Methodological Underpinning

Methodologically, the choice of literature review approach is no less crucial than that of the econometric methods applied in subsequent empirical chapters. A doctoral thesis, in fact, is judged as much on its theoretical underpinnings as on the originality of its empirical contributions (Hart, 2018). To pursue a literature review that is selective or eclectic (however long) is to open it up to accusations of bias; systematic conducting provides an audit trail by documenting the selection processes at each step (Cochrane, 1972). Such transparency was a key consideration when deciding how to conduct this review.

One of the driving forces behind methodological innovation in social science literature reviews has been advancements in evidence-based medicine, beginning with Cochrane's (1972) foundational argument for systematic reviews of medical treatments. The subsequent decades brought a series of methodological guidelines that grew increasingly sophisticated in their approach to locating, appraising, and synthesizing literature. PRISMA is the most widely used guideline and, first introduced in 2009 and updated in 2020, comprises both a checklist for reporting systematic reviews and a corresponding flow diagram that depicts their conduct (Page et al., 2021; Moher et al., 2009). The principles it embodies were increasingly being applied in non-medical fields, including business, education, information systems, and finance (Tranfield et al., 2003; Okoli, 2015) by the year 2020.

Yet, as much as PRISMA has become the international standard in its native medical context, direct translation into finance faces structural obstacles. Randomized controlled trials constitute the highest level of evidence in medical research, and the distinction between primary research and secondary syntheses of research is more readily apparent. Finance and economics are a more eclectic realm, with research on credit risk and banking spanning the entire range from strictly specified econometric models, employing panel data analysis and large-scale datasets, to broad analyses of banking regulation and institutional change drawing upon official policy documents and reports. Furthermore, while Merton's (1974) structural theory of the pricing of corporate debt certainly comprises relevant prior research on the mechanics of corporate financial distress, much of the relevant literature is on the prevalence of non-performing loans in developing countries (as the section on Pakistan shows) and institutional context shapes the research agenda: the Basel Accords, for example, directly influence the way that financial institutions approach credit risk management and thus the research questions that scholars are motivated to explore. An assessment protocol, created with medical trials as its primary referent, can hardly cope with this variety of data types without significant adjustment.

Hence, the guiding principle when designing the literature review component of this study was to adapt PRISMA's values of transparency, replicability, and explicitly stated inclusion criteria to the evidentiary context of financial research (Denyer & Tranfield, 2009). This was addressed in two ways. First, PRISMA was utilized as the protocol for the initial screening and inclusion of sources; an auditable, repeatable mechanism was thus adopted at the first hurdle. Second, an integrative-narrative component was added as a second mechanism of selection to incorporate sources that PRISMA, with its stringent requirements for empirical structure, would fail to capture but are nevertheless important in interpreting context. These included policy reports and governmental documents, relevant theoretical discussions, accounts of the evolution of financial institution characteristics, and analyses of financial policy that lack formal, empirical backing but are essential to understanding credit risk in Pakistan.

The combination of these two mechanisms reflects Grant and Booth's (2009) framework in which they delineate 14 varieties of literature review, ranging from narrative reviews to systematic reviews and critical reviews; this study aligns with the former category and that of a 'systematic review', and elements from an 'integrative review' in the sense described by Grant and Booth. It is a holistic and coherent approach designed not only to avoid the pitfalls of ad-hoc research but also to include methodological diversity when appropriate. This approach, I believe, is particularly important in the study of credit risk in Pakistan, given the protracted nature of financial innovation, varied forms of relevant documentation, and disparity in data quality across several decades.

The central management of references in Zotero also contributed significantly to this effort. Records from each of the databases employed were readily imported into Zotero, and duplicates were quickly removed. The system of tagging allowed for thematic organization of included articles, and all stages of the review were meticulously logged. Every inclusion/exclusion decision was documented within the Zotero file, and two members of a doctoral group cross-checked the coding to assure a consistent, reliable application of the screening process, in both the PRISMA-driven portion of the review and its integrative narrative portion, where such rigor would have been less readily achievable on its own without additional mechanisms to prevent bias in selecting these broader sources.

3. Search Strategy

The structure of the search strategy was developed to capture the research within the relevant thematic and coverage scope. Due to the cross-disciplinary nature of credit risk literature (spanning economics, finance, banking regulation, and Islamic finance), a multi-source literature search strategy was vital. No sole source may be able to capture the literature comprehensively when institutional reports and policy papers are also considered. Hence, both academic formal literature databases and institutional sources were planned to cover the evidence.

Academic database literature search strategy used Publish or Perish (PoP) software to systematically run search strings on two different academic literature search engines. The selected academic sources were Google Scholar, as it includes academic working papers, conference papers, and institutional documents (not available in databases like Web of Science or Scopus), and the CrossRef for accessing peer-reviewed journal articles with standardized metadata. Together, these two sources could capture the broad range of academic work on credit risk and bank performance literature.

The formal academic sources search was supplemented with a specific search for institutional literature. Selected institution literature sources included the State Bank of Pakistan (SBP), International Monetary Fund (IMF), World Bank, Bank for International Settlements (BIS), and Basel Committee on Banking Supervision (BCBS). These institutions produce a vast number of policy reports, technical working papers, and regulatory assessments, which may not be indexed in standard academic sources yet contain relevant empirical information and have considerable implications for policy. For example, the Financial System Stability Assessment (FSSA) of Pakistan produced by the IMF, and the periodic Financial Stability Review published by the State Bank of Pakistan (SBP) have been useful sources of empirical data and institutional information specific to Pakistan.
The period of the literature search was established from the late 1960s to 2024 to cover all aspects of credit risk literature. The beginning year of the literature search from the late 1960s was based on the inclusion of path-breaking papers like the Z-score model of Altman (1968), as it remains the most cited paper on bankruptcy prediction, and Merton's (1974) structural model of corporate debt pricing. It covers the literature on non-performing loans development during the Latin American debt crisis in the 1980s and the Asian financial crisis of the late 1990s, and so on, until 2024 to include the latest papers on credit risk and bank performance, e.g., the recent paper of Leo et al. (2019) on machine learning application for credit risk evaluation and studies on the impact of the COVID-19 pandemic.

Search strings of keywords were refined to include a balance between scope and specificity. A broad keyword set consisting of "credit risk" and "bank performance" was initially; however, this was narrowed down by the result of preliminary searches and further inclusion of new relevant keywords. The final keyword structure was organized by the five main categories: "credit risk measures," "bank performance indicators," "macroeconomic context," "Islamic banking," and "Pakistan and South Asia," where Boolean operators (AND/OR) were used to construct specific keyword combinations.

Table 1: Search Keywords and Boolean Combinations

Theme	Keywords	Boolean Combinations
Credit risk measures	"credit risk," "non-performing loans," "NPLs," "loan defaults," "GNPLR," "distance-to-default," "Z-score"	("credit risk" OR "non-performing loans" OR "NPLs" OR "loan defaults") AND ("Z-score" OR "distance-to-default")
Bank performance	"bank profitability," "ROA," "ROE," "capital adequacy ratio," "CAR," "efficiency"	("bank profitability" OR "ROA" OR "ROE" OR "capital adequacy" OR "CAR") AND ("credit risk")
Macroeconomic context	"GDP growth," "inflation," "interest rates," "exchange rate volatility," "macro-financial linkages"	("credit risk" OR "NPLs") AND ("GDP" OR "inflation" OR "interest rate" OR "exchange rate")
Islamic banking	"Islamic banks," "Shariah governance," "murabaha," "sukuk," "Islamic finance risk"	("Islamic banks" OR "Islamic finance" OR "Shariah governance" OR "murabaha" OR "sukuk") AND ("credit risk" OR "performance")
Pakistan & South Asia	"Pakistan banking sector," "South Asia credit risk," "emerging markets banking," "political lending Pakistan," "SBP regulation"	("Pakistan" OR "South Asia") AND ("credit risk" OR "NPLs" OR "bank performance")

Note. Keywords were iteratively refined through pilot searches. GNPLR = gross non-performing loan ratio; ROA = return on assets; ROE = return on equity; CAR = capital adequacy ratio. SBP = State Bank of Pakistan.

From the searches, we obtained 1350 records: 930 results from PoP-based Google Scholar searches, 173 results from searches within CrossRef, and 247 results from institutional databases and manual searching. The first list was then screened systematically, as outlined in the following sections. The combined use of academic database and institutional repository searches was crucial: a sizable portion of the papers that were most relevant to the Pakistan banking context – specifically policy analysis and regulatory evaluations – could only be located through institutional resources.

All results were exported to and managed within the central reference management system Zotero throughout the entire review process. Zotero's automatic deduplication and tagging capabilities, coupled with its ability to manage a disparate collection of records, were instrumental in keeping track of the large and disparate evidence base. Since Zotero was used for both components of the systematic and integrative review, the entire process could be managed centrally, and cross-referencing between themes was facilitated.

4. Inclusion and Exclusion Criteria

As with the search strategy, transparency of the inclusion/exclusion criteria is crucial for the validity of a systematic synthesis. The choice of inclusion and exclusion criteria determines which evidence base becomes the object of systematic analysis, acting as the principal mechanism to ensure that it is relevant and tractable (Higgins et al., 2019). The parameters used in this study seek to combine broad coverage and thematic depth, aiming to capture the wide array of literature on credit risk, but at the same time, focusing strictly on material directly addressing the study's research questions.

The four dimensions along which the inclusion criteria were drafted were: first, thematic (requiring an explicit treatment of the link between credit risk and bank performance, measured using standard parameters like the gross non-performing loan ratio (GNPLR), Z-score, distance-to-default, return on assets (ROA), and capital adequacy ratio (CAR), with other areas being addressed only when they touched on credit risk implications; e.g. Market and operational risk); second, geographic (with preference given to Pakistan, South Asia, and emerging markets, but including the global literature for purposes of comparison or as a theoretical underpinning); third, method (relying on empirical and analytical methods like econometrics, case study and policy assessment, and discarding descriptive-only studies); and fourth, quality (prioritizing peer-reviewed articles and relevant policy reports from established international and domestic institutions and research networks, with an acceptance of top quality working papers).

The criteria for exclusion were also made explicit: excluding from consideration studies of the non-banking sector and household finance without links to banks, solely developed-world countries that were not demonstrably applicable to developing country issues, any materials not in the English language that could not be reliably translated for reasons of practicality, and blogs, news and opinion articles lacking analytical or methodological rigor, irrespective of any potential interest).

Table 2: Inclusion and Exclusion Criteria

Criterion	Inclusion	Exclusion
Focus	Credit risk and bank performance (GNPLR, Z-score, Distance-to-Default, CAR, ROA, macro drivers)	Non-banking sectors; household credit only
Geography	Global, emerging markets, South Asia, Pakistan	Advanced economies only (without transferable insights)
Source type	Peer-reviewed journals, IMF/WB/SBP/Basel reports, books, and high-quality working papers	Blogs, news articles, opinion pieces
Language	English	Non-English without a credible translation
Relevance	Linked to credit risk, bank performance, or Islamic banking	Topics outside scope (e.g., marketing, HR, management)

Note. GNPLR = gross non-performing loan ratio; CAR = capital adequacy ratio; ROA = return on assets; IMF = International Monetary Fund; WB = World Bank; SBP = State Bank of Pakistan.

Using these criteria with the first 1,350 search records yielded 247 studies with PRISMA inclusion. In addition to this systematic search process, a subsequent narrative-integrative review added a variety of regulatory, conceptual, and institutional literature from Pakistani history and the Pakistan context (see Section 7 for further discussion). These additions brought the evidence base to 281 documents. Zotero's tagging function proved useful throughout this double-track search, classifying documents thematically, by geographic scope, and by methodological quality systematically.

5. Screening Process and PRISMA Flow Diagram

PRISMA 2020 guidelines were followed to facilitate transparency and replicability throughout the four stages prescribed by the protocol, viz identification, screening, eligibility, and inclusion (Page et al. 2021). All four stages are described in more detail below, and the entire process is illustrated by the PRISMA flow diagram in Figure 1. The identification stage captured a total of 1,350 records through the three primary search channels, namely: 930 from Google Scholar via PoP, 173 from CrossRef, and 247 from institutional and manual searches such as SBP, IMF, World Bank, Basel Committee, and BIS. The high proportion captured by institutional sources highlights the prominence of policy and regulation-oriented literature in the domain of credit risk, especially in the Pakistani context, where official SBP reports and assessments by international bodies such as the IMF contribute to a large part of the available evidence.

The deduplication stage was conducted using Zotero's automated deduplication tool, followed by manual checking. A total of 72 records were eliminated in the deduplication stage as they were duplicates, leaving 1,278 unique records for screening. The low duplication rate (5.3%) shows the complementarity of the three search channels, which tended to capture separate segments of the literature.

In the screening stage, the titles and abstracts of all 1,278 articles were screened to identify clearly irrelevant papers. The total number of excluded papers was 1,031 and the papers were categorized into five groups of exclusion: 412 unrelated to banking or credit risk (corporate finance, household finance and insurance related), 238 not applicable to the geographical or contextual situation (mostly studies related to advanced country banking and transferable to Pakistan is limited), 201 not appropriate due to methodology (purely descriptive or method incompatible), 96 incomplete or unavailable papers (including unpublished research), and 84 of low relevance or quality. The distribution of excluded reasons is provided in Figure 2.

In the eligibility stage, 247 papers were downloaded and assessed on the criteria of methodology and contextual appropriateness. As compared to the screening stage (title and abstracts review), in this stage, all papers were read thoroughly with focus on methodology, data, and results, and all 247 articles reaching this stage were included for PRISMA synthesis, owing to sufficiently strict selection in the previous stages.

After the initial PRISMA synthesis, the review was updated in 2025 to include new research and literature identified in the integrative-narrative branch. The update led to the addition of 34 articles: the latest empirical research available, new regulations from the SBP and Basel Committee, and theoretical development pertinent to Pakistan. Overall, the evidence base expanded from 247 to 281 articles and was up-to-date and met methodological robustness criteria.

Advance Journal of Econometrics and Finance
Vol-4, Issue-2, 2026

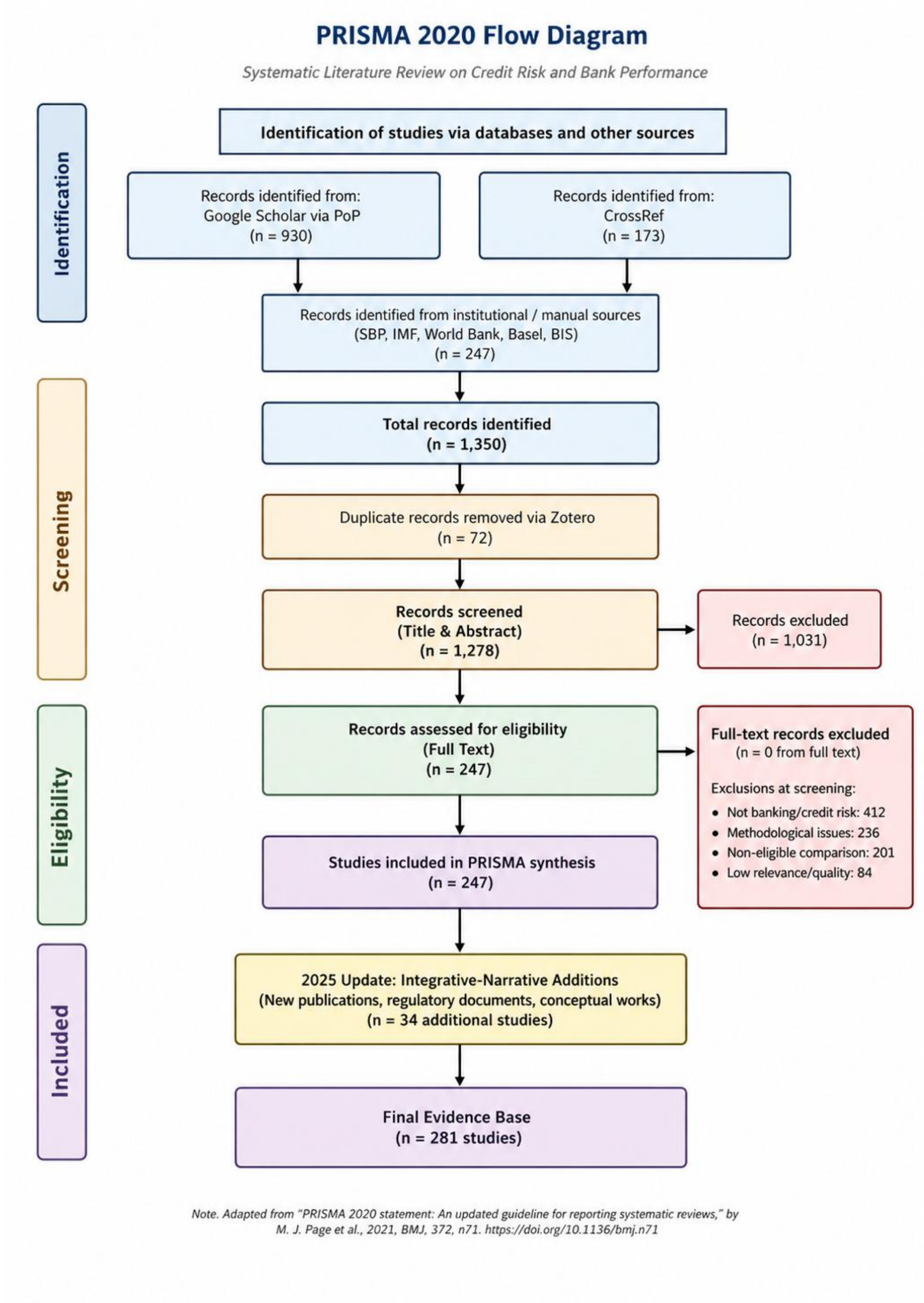

Figure 1: PRISMA 2020 flow diagram for the systematic literature review process, showing identification, deduplication, screening, eligibility, inclusion, and the 2025 update stages. The conclusive evidence base comprises 281 studies.

As shown in Figure 1, the sequential screening approach after the PRISMA 2020 protocol allowed for progressively decreasing the evidence pool from the 1350 records initial pool, 247 studies included after the systematic process, and 281 studies after the 2025 integrative update. The high number of eliminated studies during the screening phase is explained by the wide search strategy at the beginning and by the stringent eligibility criteria applied to select the articles and guarantee their accuracy and methodological rigor. The small number of duplicate papers indicated that the multi-source search strategy (databases and institutional repositories) effectively obtained different bodies of work. The inclusion of 34 papers in the integrative process suggests that fully systematic protocols are not suitable for finance literature and emphasizes the role of contextual and policy sources. The figure clearly shows the reproducibility and transparency of the review process. The figure also shows the required flexibility in the methodology to deal with this diverse field of research.

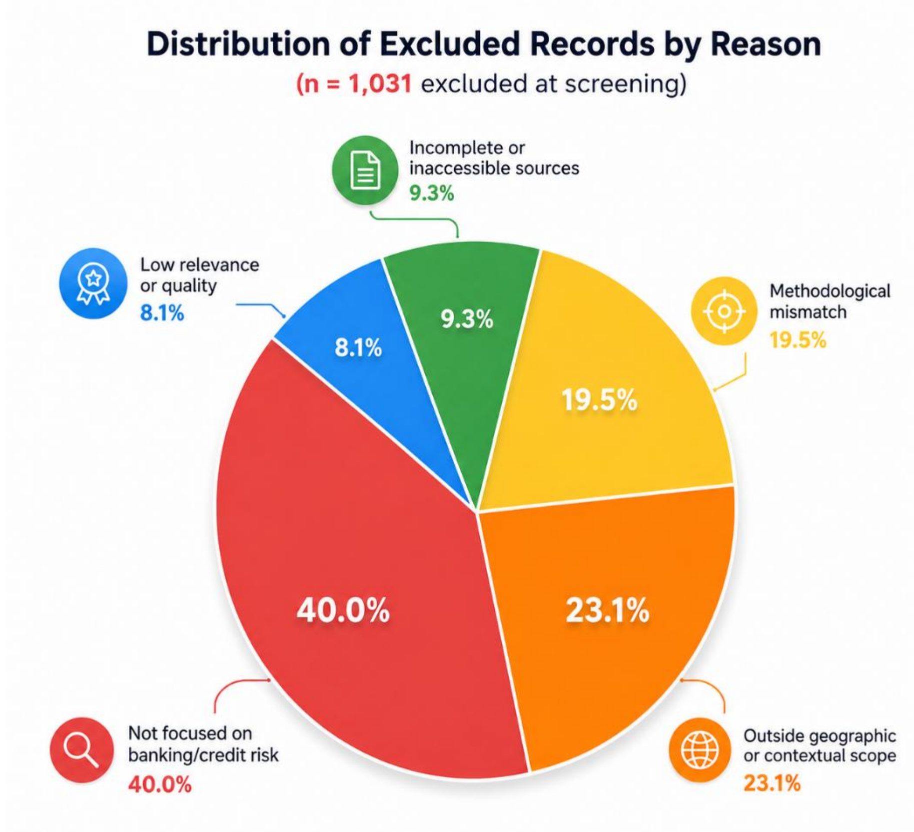

Figure 2: Distribution of excluded records by reason at the screening stage (n = 1,031). The largest category was studies not focused on banking or credit risk (40.0%), followed by those outside the geographic or contextual scope (23.1%).

Figure 2 outlines excluding papers by screening level and illustrates both what type of papers are captured at the initial search and the reasons why. Over half of the excluded papers are from a domain outside of banking and credit risk; this demonstrates that while the necessary broadness of keyword searches allows inclusion of literature on the topic, it simultaneously captures literature on closely related but irrelevant subject matter. The exclusion based on country, which also accounts for a significant fraction, reiterates the argument about the geographic dominance of the literature in the form of advanced country research. Finally, excluded papers were primarily descriptive and/or imprecisely specified in terms of methodology. Such figures illustrate why a careful selection process is important for ensuring a sound and well-developed literature search, and that the sample we end up with is neither arbitrarily selected nor unintentionally constructed.

The two-stage approach—initial PRISMA synthesis followed by an integrative update—demonstrates that the review was the product of a transparent, replicable, and defensible process. The PRISMA protocol ensured that the initial evidence base was constructed through systematic procedures, while the 2025 update captured recent developments and contextual sources that enhanced the review's relevance and comprehensiveness.

6. Quality Assessment Framework

Despite the careful selection by the PRISMA process, the quality of every single remaining study could not be assumed to be high enough for use in the literature synthesis. Therefore, the subsequent stage of the analysis involved assessing the quality of every retained study on the three main dimensions (methodology, context, and influence), using frameworks from well- established systematic review protocols in management and economics (Tranfield et al., 2003; Denyer & Tranfield, 2009). This quality appraisal was conducted for all 281 papers retained by the literature search process (247 from the PRISMA stage and 34 added by the integrative-narrative strand).

The methodological rigor of each paper was assessed according to the soundness of the empirical framework, the appropriateness of model specification, the econometric estimation technique, and the transparency of data sources and models. Panel data analysis (fixed effects, random effects, generalized method of moments, etc.), VAR methods, structural models, and time- series estimation procedures received more positive ratings than the use of plain cross-sectional regressions or simple descriptions. Special attention was also paid to the handling of endogeneity problems, given that credit risk, financial profitability, and capital adequacy are jointly determined in banks' balance sheets (Berger and Bouwman, 2013); empirical studies could rarely be exempt from it.

The contextual relevance of a paper was evaluated based on the degree of its transferability to developing economies' banking sectors and, more specifically, to the developing country banking sector of Pakistan and South Asia. Studies on Pakistan's banking sector scored highest for contextual relevance, whereas for advanced countries, the scoring was based on the transferability potential of their findings. For example, studies on determinants of NPL in the Eurozone could be regarded as more relevant for developing economies' banking sectors than household credit scoring studies in the United States, as there were close similarities between banking structures and regulatory frameworks between the former and South Asian banking systems (Espinoza & Prasad, 2010).

Finally, scholars' influence was estimated based on multiple proxy variables: the number of citations in academic literature (Google Scholar and Scopus) and academic journals' ranking (Scopus quartile rank and ABS/ABDC ranking), or, where applicable, for reports of institutional bodies, the ranking of issuing institution (e.g., IMF, World Bank, BIS).

Top-quartile journal articles and institutional reports received the highest citation score. Scholarly influence was used to augment (not substitute for) methodological rigor and contextual relevance to allow for inclusion of research with a stronger influence impact but more methodological limitations or limited context relevance.

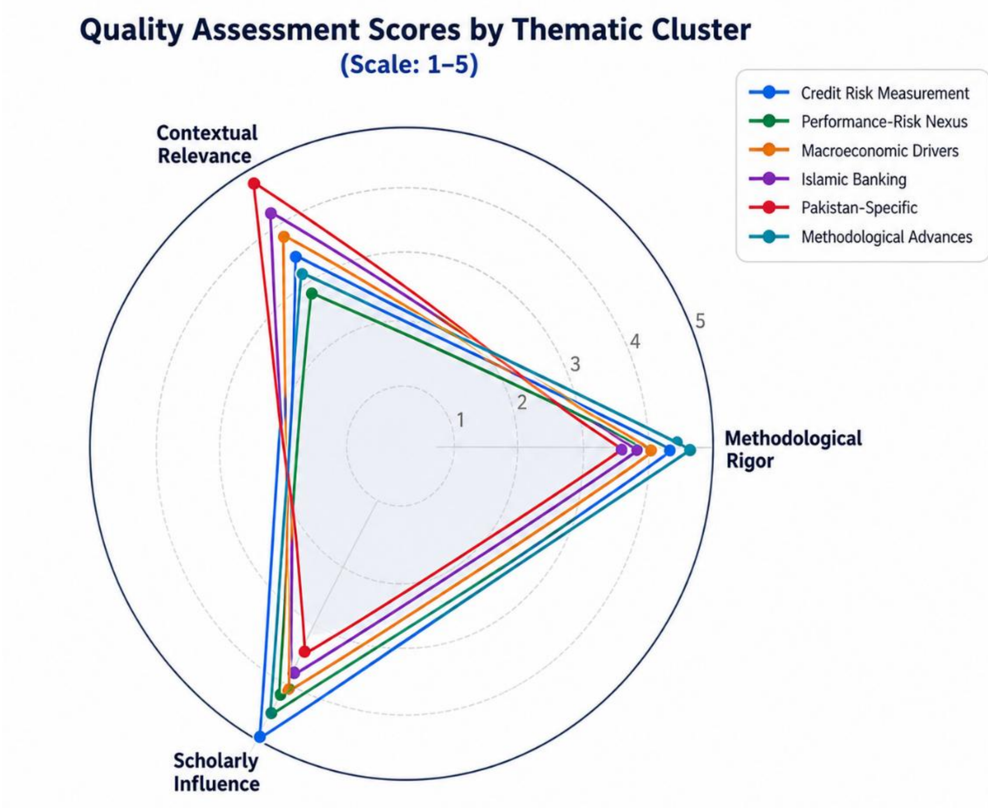

Figure 3: Quality assessment scores by thematic cluster across three dimensions: methodological rigor, contextual relevance, and scholarly influence (scale: 1–5). Credit risk measurement scored highest on methodological rigor and scholarly influence, while Pakistan-specific studies achieved the highest contextual relevance.

Figure 3 presents the three quality measures across the different thematic clusters. We see substantial variation in terms of rigorousness, contextual relevance, and academic impact of the included literature. The thematic clusters on credit risk measurement and methodological advancement possess the highest methodological rigor and influence/citation, due to the mature quantitative nature of the approaches and highly influential theoretical literature on these topics. Conversely, the thematic cluster on Pakistan-specific studies has the highest relevance, reflecting the need for local understanding and policy formulation, even though its influence is low. This dichotomy between influence and contextual relevance points to a structural issue in literature, where highly cited international literature may not necessarily be suitable in emerging market contexts, reiterating the importance of including region-specific papers to balance the literature from highly influential international literature.

7. Data Extraction and Thematic Coding

After defining the 247 studies to be included from PRISMA, we performed a rigorous extraction of data. A standard extraction template was developed to systematically collect the necessary data from each study, which included bibliographic details, study objectives, research methods, the variables used, geographical coverage of the study, main results, and reported limitations of the studies. This template facilitated consistency throughout the extraction and was useful in comparing study designs and results afterwards. An additional 34 studies that we included as part of the 2025 integrative-narrative update were also extracted by following the same template for comparable data and methodologically rigorous manner of analysis on a total of 281 studies.

After extraction, a thematic analysis was performed in which studies were categorized not by chronology but by conceptual and methodologically similar themes. This was grounded research in a thematic synthesis tradition (Thomas & Harden, 2008) so that the literature review can lay an analytical ground to facilitate a cross-cutting analysis across method, location, and institution. The categorization of themes was done using Zotero's tags facility, which was beneficial in assigning several thematic and other tags for each study, like theme, region, and type of method used.

Six major thematic clusters were developed. These are: credit risk measurement; performance-risk nexus; macroeconomic factors affecting credit risk; Islamic banking; Pakistan-specific studies; methodological advances. The 'credit risk measurement' is the most densely populated cluster with 93 studies examining the measurement of credit risk, which includes retrospective or backwards-looking indices, such as GNPLR, composite indices such as Z-score, and forward-looking, market-based approaches, such as distance-to-default. The set includes 93 papers. This cluster includes the broadest set of studies as it ranges from Altman's (1968) landmark multivariate discriminant analysis paper to recent applications of machine learning methods for default prediction (Leo et al., 2019). The most widely used credit risk proxy across these papers is GNPLR (72% of papers in this cluster), due to its data availability and SBP's mandate to disclose in Pakistan.

The second theme, the 'performance-risk nexus,' captures the bidirectionality between bank profitability and credit risk, that is, whether higher performance makes the bank riskier or if it enables it to build higher capital buffers to absorb unexpected losses. Some authors find a negative relationship between profitability and NPLs (Boahene et al., 2012), while others find a positive correlation for certain periods, when the banks are trying to grow aggressively (Klein, 2013). The other topics that feature in this cluster are the impact of capital on the credit risk and how the levels of capital affect the lending behavior. This cluster has 58 papers. The relationship of a bank's capital with credit risk is studied as both a dependent and an independent variable, and it is suggested that banks with higher capital levels take on less risk (Berger et al., 1995).

The third theme, 'macroeconomic factors', covers 41 studies investigating the relationship between macroeconomic aggregates and banks' credit risk/performance. The common macroeconomic variables include GDP growth rate, inflation rate, interest rates, and exchange rate volatility. The findings on the relationship between macroeconomic trends and NPLs are well-established: credit risk rises in an economic downturn and declines in an upturn (Espinoza & Prasad, 2010; Nkusu, 2011). Inflation exerts a negative impact on the loan portfolio due to a reduction in the real value of collateral and a positive impact on the payment capacity of debtors by eroding their real income. However, its net effect is still unclear in the literature. The exchange rate has a significant impact, especially in the Pakistani context, due to the heavy reliance of the economy on imports. It causes an increase in credit risk through its impact on the firm's balance sheet for those that have huge foreign exchange liabilities (Javed & Anwar, 2018).

The fourth theme is the 'Islamic Banking,' consisting of 27 studies examining the unique characteristics of risk in Islamic finance. The core question explored here is whether Islamic banks, on account of their specific structure, including asset-based financing and profit and loss sharing mechanisms, are resilient to credit shocks compared to conventional banks. Existing evidence on this question remains inconclusive but shows that some authors attribute lower NPLs for Islamic banks during financial crises to their unique structure (Beck et al., 2013; Cihak & Hesse, 2010), whereas others find no statistical difference after controlling for size and capital structures, etc. (Hasan & Dridi, 2011). The aspect of governance in Islamic banks is unique as Shariah governance is required for Islamic banks, which plays a role in the management of risk (Chachi, 2008).

The fifth theme of 'Pakistan-specific studies' comprises 37 papers analyzing the credit risk within the specific context of Pakistan. The main issues covered are legacy of politically directed lending to private and state owned enterprises leading to buildup of non-performing loans in 1990s (Burki & Niazi, 2006; Zaidi, 1999); reforms by State Bank of Pakistan including privatization of major banks, adoption of Basel Standards; and the nature of risks associated with Pakistan's economy characterized by recurrent episodes of macroeconomic volatility including high inflation, current account and balance of payments crises and depreciation of national currency (IMF, 2016). The World Bank and IMF (2016) analyses confirm the improvements in the asset quality over the last decade, but structural vulnerabilities remain, especially in the State-Owned Banks as well as in the SME credit segment.

The last theme, 'methodological advances', covers 25 papers detailing and advancing the methodological tools in credit risk analysis, starting from early cross-section regressions up to panel data, PVAR, and stress testing methods. In the latest literature, this theme includes the application of machine learning methods like random forest, SVM, and neural networks, and other data-driven approaches in modeling credit risk (Leo et al., 2019). Though promising to be better predictors of credit risk, the use of these techniques in Pakistan remains limited by the availability of data and interpretation issues.

Table 3: Thematic Coding of Literature (n = 281)

Theme	Studies	Example References	Key Characteristics
Credit risk measurement	93	Altman (1968); Klein (2013); Louzis et al. (2012); Merton (1974)	GNPLR dominant; Z-score/DD increasingly applied
Performance-risk nexus	58	Boahene et al. (2012); Hanif et al. (2012); Berger et al. (1995)	ROA and CAR central; mixed evidence on directionality
Macroeconomic drivers	41	Afzal & Mirza (2012); Espinoza & Prasad (2010); Nkusu (2011)	GDP, inflation, and FX volatility are consistently examined
Islamic banking	27	Beck et al. (2013); Cihak & Hesse (2010); Hasan & Dridi (2011)	Focus on resilience and Shariah governance.
Pakistan-specific studies	37	Burki & Niazi (2006); IMF (2016); Zaidi (1999)	Political lending and regulatory forbearance
Methodological advances	25	Arellano & Bond (1991); Leo et al. (2019); Love & Zicchino (2006)	PVAR, stress testing, ML techniques

Note. DD = distance-to-default; GNPLR = gross non-performing loan ratio; ROA = return on assets; CAR = capital adequacy ratio; FX = foreign exchange; PVAR = panel vector autoregression; ML = machine learning. The total of 281 exceeds the sum across themes (281) because some studies were coded under multiple themes.

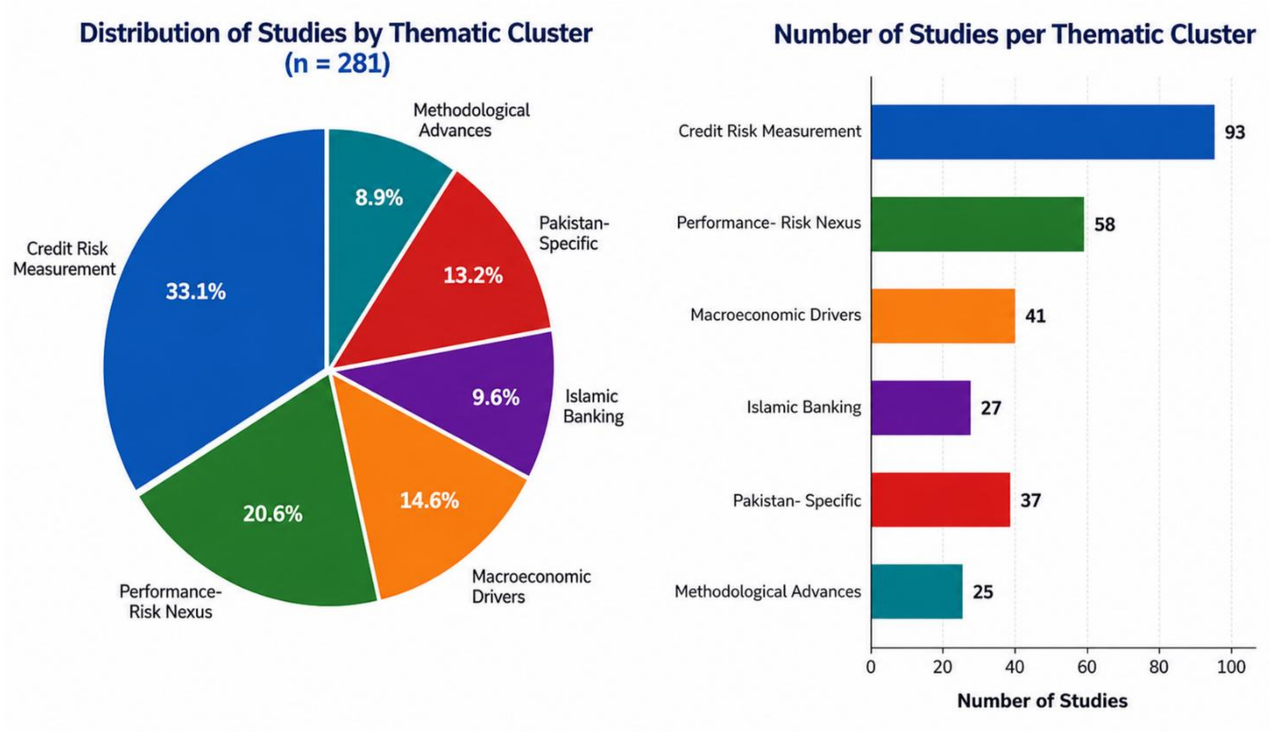

Figure 4: Distribution of studies across thematic clusters. Credit risk measurement is the largest cluster (33.1%), followed by the performance-risk nexus (20.6%) and macroeconomic drivers (14.6%).

The distribution of papers across the six themes is presented in Figure 4, and the finding that clearly dominates is the measurement of credit risk. The preeminence of the measurement theme reflects a tradition in literature of exploring and refining quantitative proxy measures of credit risk, influenced largely by the need for data and regulatory reporting. An indicator of under-research in the measurement context, despite their perceived importance, is the small number of papers relating to both Islamic banking and innovation in measurement. These latter findings can also imply a potential skew in theory where measurement aspects dominate compared to issues of institutional, behavioral, and structure-related aspects driving credit risk. This indicates that there are avenues in future work to develop research beyond these simple measures.

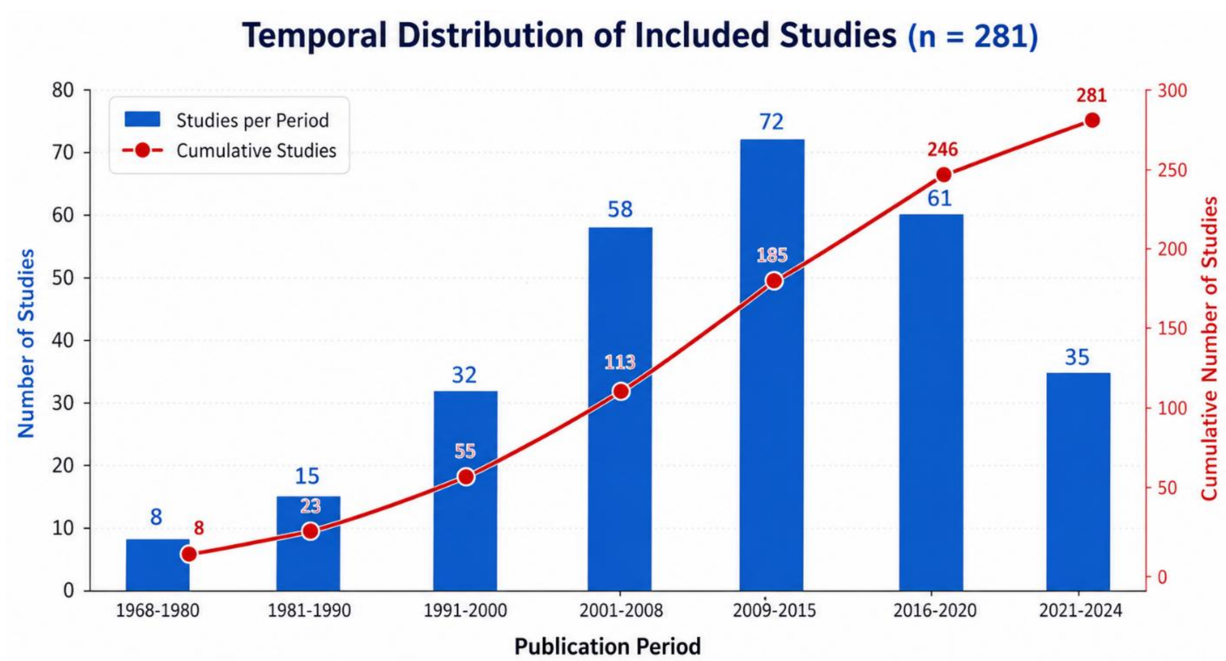

Figure 5: Temporal distribution of included studies from 1968 to 2024. The largest number of studies was published in the 2009–2015 period following the global fiscal crisis, reflecting the surge in credit risk research during this time.

Figure 5 displays a clear trend in how this literature evolved throughout the years, from 1968 to 2024. We observe a remarkable growing trend in research output from the fiscal crisis (2007-2008) onwards, which is attributable to an increasing research and policy interest in credit risk and financial stability following systemic financial events. The tendency to study credit risk following crises suggests that most of the current literature is affected by crisis-induced research, which leads to findings that may be skewed towards stress periods rather than stable economic states. Previously, there is relative sparse evidence of studies in this field with emphasis on classical theoretical work such as bankruptcy prediction and structural credit risk models. We observe that in recent years, there has been an increasing tendency for research employing state-of-the-art methodologies, such as machine learning and stress-testing models. Overall, the figure demonstrates the cyclical interest in research topics in the field of finance, with significant global events being triggers for research outputs, and a demand for research that may be more continuous and with forward-looking properties.

Structuring extraction and thematic coding accomplished two important things: First, it converted the large and diverse pool of literature into a structured analytical framework and allowed for synthesis across different methodologies, geographies, and institutional settings; second, it enabled the uniform treatment of 247 PRISMA retained and 34 other papers, avoiding methodological discontinuity in the integrative thread. Avoiding chronological ordering, thematic coding helped detect issues across different approaches, countries, and institutions: A sustained tension between backward- and forward-looking credit risk measures; that macroeconomic instability increases bank risk in emerging markets, etc. Which would be difficult to uncover by simply describing papers chronologically.

8. Strategies for Addressing Bias

Bias is a chronic, though valid, issue in literature reviews, especially when there is a thematic or geographic imbalance in the coverage (Egger et al., 2008). The five sources of bias identified, and the strategies employed to counter them, show the elevated level of commitment that has been made beyond PRISMA's minimum transparency requirements: Publication bias-the systematic tendency for articles to be published based on their statistical significance and novelty, and negative or confirmatory results to go unpublished-is an ever-present threat to the integrity of any literature synthesis (Rothstein et al., 2005). For credit risk literature, the potential bias may result in overstating relationships between variables, as those studies find statistically significant variables that explain non-performing loans or bank profitability are more likely to find their way into published work. Publication bias was mitigated by the inclusion of high-quality working papers and institutional reports such as SBP Financial Stability Reviews, IMF country assessments, and World Bank policy research working papers, irrespective of the statistical significance of the results. These sources are less subject to publication bias than peer-reviewed journals, as the primary purpose is often to share information rather than contribute novel findings. The use of 34 additional studies for the integrative-narrative strand also increased the available evidence beyond what would be identified with a journal search alone.

Geographic bias-the over-representation of literature emanating from developed economies- is a recognized issue within the finance literature (Berkowitz et al., 2022). The majority of credit risk literature originates from the USA, Europe, and East Asia, while South Asia and Sub-Saharan Africa remain severely under-researched. Geographically targeted keywords and the inclusion of regional institutional research provided for over-sampling for Pakistan, South Asia, and emerging economies. The resultant evidence base comprises a considerable proportion of Pakistan-specific studies (n=37) but also includes comparison research from India, Bangladesh, and Sri Lanka, and emerging markets.

Conceptual bias-the potential for over-reliance upon one specific theory or way of measuring a concept-was prevented by ensuring the literature covers various methods for measuring credit risk. GNPLR remained the dominant variable due to its importance as a regulatory measure and ease of access to data; however, it was used alongside variables measuring Z-score, distance to default, expected loss models, and market indicators for default risk. Researcher bias-the potential for an investigator's judgment to influence study selection, coding, and analysis-was guarded against by documenting all stages of the search, screening, and coding in Zotero to produce an auditable record of choices made, validation by a doctoral colleague who double-checked all coded data, and discussion of conflicting coding against the specified inclusion criteria. Whilst researcher bias can never be eliminated in the research process, researcher bias can be significantly minimized with these procedural mechanisms. Finally, integrative bias-the risk that the PRISMA protocol alone will omit important conceptual and regulatory works-was addressed by the addition of the integrative-narrative strand. The methodology of PRISMA itself can privilege studies with formal research designs at the expense of regulatory analysis or theoretical work, which lack the rigor of typical empirical research but contribute valuable insights; the integrative-narrative strand included works that were fundamentally designed to supplement and incorporate such items.

The explicit recognition and mitigation of these five sources of bias highlight the commitment to methodology that is integral to credible literature review work. Addressing the biases up front guarantees that the findings are based on a robust and representative selection of literature incorporating the rigor of PRISMA and the contextual specificity that the banking sector of Pakistan demands.

9. Integration with the Empirical Research Framework

The literature review was not an end but had a direct relation with the empirical design of the overall research project. Based on the screened and quality-checked literature, the data collected were thematically coded to identify the set of variables that were subsequently operationalized within the Panel VAR framework in the empirical chapters of the thesis. The crucial inductive aspect of the relation between literature and the empirical model lies in illustrating how evidence-driven (rather than imposed a priori) the variables chosen were.

The confluence of evidence across the 281 studies singled out a set of variables as core to credit risk and bank performance in Pakistan. GNPLR stood out as the most used proxy variable for credit risk, being used in most of both international and Pakistan-specific papers, owing to its widespread availability and regulatory relevance (Ali & Ahmed, 2024; Klein, 2013; Louzis et al., 2012). Z-score and distance-to-default, forward-looking measures of bank stability, though infrequently used in Pakistan due to data limitations, are prominently discussed in international literature, representing the other side of bank fragility (Beck et al., 2013; Bharath & Shumway, 2008). ROA and CAR are consistently reported to both cause and be consequences of credit risk owing to the bi-directional link between profitability, capital buffers, and asset quality (Hanif et al., 2012; Berger et al., 1995). Macroeconomic variables, including GDP growth, inflation, interest rates, and exchange rate volatility, all recur in the literature as strongly linked to credit risk outcomes across all sorts of institutional contexts and were thus seen as a sine qua non for any empirically based bank risk model (Espinoza & Prasad, 2010; Nkusu, 2011).

Advance Journal of Econometrics and Finance
Vol-4, Issue-2, 2026

Table 4: Mapping of Empirical Variables to Literature Sources

Variable	Literature Support	Justification
GNPLR	Ali & Ahmed (2024); Klein (2013); Louzis et al. (2012)	Dominant credit risk proxy; data availability for Pakistan
Z-score	Beck et al. (2013); Cihak & Hesse (2010)	Stability measures integrating profitability and volatility.
Distance-to-Default	Bharath & Shumway (2008); Merton (1974); Duffie et al. (2007)	Market-based default probability
ROA	Hanif et al. (2012); Shah & Ahmed (2020)	Profitability indicator linked to credit quality.
CAR	Altman et al. (2017); Berger et al. (1995)	Basel standard capital buffer: key for resilience
GDP growth	Espinoza & Prasad (2010); Nkusu (2011)	Expansions reduce defaults; contractions raise them
Inflation	Ahmed & Malik (2020); Klein (2013)	Ambiguous effect: reducing debt burden but weakens repayment
Interest rates	Espinoza & Prasad (2010)	Convincing evidence for higher risk under rising rates
Exchange rate	Espinoza & Prasad (2010); Javed & Anwar (2018)	Volatility critical in Pakistan's import-driven economy

Note. GNPLR = gross non-performing loan ratio; ROA = return on assets; CAR = capital adequacy ratio; GDP = gross domestic product.

Table 4 maps each empirical variable to its supporting literature, demonstrating the inductive pathway from evidence to model specification. This mapping serves a dual purpose: it establishes the empirical legitimacy of each variable, and it provides a transparent audit trail linking the literature review to the research design. The Panel VAR framework, which allows for the simultaneous modeling of multiple endogenous variables, was specifically chosen to accommodate the bidirectional relationships identified in the literature—such as the mutual dependence of credit risk, profitability, and capital adequacy—that simpler regression frameworks cannot adequately capture.

10. Reflections on Methodological Trends in Finance Literature Reviews

Over the last three decades, the methodological framework for conducting literature reviews in finance has radically changed, as part of the wider trend toward transparent and evidence-based research that has permeated the social sciences. Literature reviews in finance and economics were initially essentially narrative in nature. While they did provide a valuable source of intellectual context, they lacked transparency in how the literature to be included was selected, weighted, and included in each review (Webster & Watson, 2002). This meant that the conclusions from such literature reviews were based on a systematic search, selection, and integration of studies without explicitly documented search strategies, inclusion criteria, or systematic selection processes and could therefore not easily be replicated or updated by other researchers.

In finance, systematic review methodologies grew rapidly from around the mid-2000s due to the combination of a few factors. Firstly was the widespread popularity and success of evidence-based research paradigms in medicine and public policy (Petticrew & Roberts, 2006), showing that transparency of method was essential if a literature review was to have value beyond that of an intellectual review; secondly was the explosion in both the quantity and diversity of financial literature, making informal, ad-hoc reviewing methods increasingly unreliable; and thirdly was the development of technology which allowed for a more systematic way to identify and manage literature (Okoli, 2015), including bibliometric databases, reference management software and text-mining algorithms.

The arrival of PRISMA as a reporting standard for systematic reviews was a particularly important event. Originally designed for health studies, PRISMA has progressively been extended for use in the social sciences (Page et al., 2021) and has given rise to a standard vocabulary and visual language – the PRISMA flow chart – for documentation of the systematic review process. PRISMA has clearly had a positive impact on increasing the transparency of the review process although its protocols are yet to be fully adapted to incorporate the highly diverse sources of evidence prevalent in financial literature; the purpose of this paper is to add to the growing literature on PRISMA adaptations by showing how PRISMA can be incorporated with an integrative-narrative strand to draw upon the full spectrum of relevant evidence, including policy documents and theoretical pieces which PRISMA primarily aims to test empirically.

A related trend is the growing popularity of quantitative synthesis using meta-analysis; although already widely used in psychology, education and management, meta-analysis, which involves combining effect sizes from various studies into a summary estimate, is relatively underdeveloped in finance as a consequence of the huge heterogeneity between different studies; dataset sources, the econometric models used, what is being tested and what specific sector or market is being examined, among other factors, make such an analysis difficult. This applies equally to the dependent variable itself, with the measures of credit risk being extremely diverse, from GNPLR to Z-score to distance-to-default, making pooling across different studies difficult in any quantitative form. Despite improvements in meta-analytic methods for heterogeneous data, they remain limited for use in the majority of credit risk applications and, for this reason, rather than conduct a quantitative meta-analysis, the study used a narrative-analytical synthesis approach with the intention of drawing together diverse sources of evidence that would likely limit the interpretation of any pooled effect-size estimate.

A fourth development in literature reviews is the increasing acceptance that grey literature such as working papers, reports, policy briefs and technical studies produced by institutional actors can be of legitimate interest for review; the exclusion of grey literature from a literature review of credit risk research would mean excluding a substantial portion of the most up-to-date and context-relevant research because so many institutional actors and central banks undertake in-depth studies in this area. In this paper, Grey literature was intentionally sought out and included, with contributions drawn from: SBP, IMF, World Bank, Basel Committee, and the BIS in recognition that this is often where primary data or conceptual information can be obtained.

Fifthly, a trend towards greater contextual inclusivity has been developed over the last decade. There is a widely recognized view that reviews based on advanced country frameworks and evidence should not be generalizable to developing country contexts (Greenhalgh et al., 2018), and this is especially true for the sector of banking and finance, with, for example, institutionality, regulatory environment, and the economic context being vastly different between emerging and developed economies. The present review addressed this trend through intentionally oversampling studies in Pakistan and South Asia to attempt to develop findings relevant to the context that the analysis intends to apply to.

11. Policy Implications

The systematic synthesis of 281 studies on credit risk and bank performance, spanning six decades of scholarship and encompassing evidence from Pakistan, South Asia, and global emerging markets, yields actionable insights for three principal stakeholders: banking regulators, commercial banks, and the academic research community. These implications are grounded in the convergent evidence identified across the six thematic clusters of the review and are informed by the methodological rigor of the PRISMA-based synthesis. Each set of implications is presented below with reference to the weight of evidence supporting it.

11.1 Implications for Regulators

Strengthening Credit Risk Measurement Frameworks

One clear finding of the systematic review was the strong prevalence of the gross non-performing loan ratio (GNPLR) as the primary credit risk indicator both in regulatory practice and in academia, accounting for about 72% of papers in the credit risk measurement cluster. While GNPLR provides the benefit of familiarity to regulators, standard reporting, and availability of time-series data, its backward-looking nature makes it unable to be used as an early warning indicator. It can be demonstrated from the literature that forward-looking indicators, especially Z-score and distance-to-default (Merton, 1974; Beck et al., 2013), reflect bank fragility dynamics that can only be detected by GNPLR when credit deterioration has already taken place and is reflected on the income statement. For the regulators in Pakistan and other emerging economies, the policy prescription is to complement the current regulatory use of GNPLR-based surveillance by adopting and using forward-looking indicators of bank stability, which would enhance the early warning architecture of the financial regulatory system. The State Bank of Pakistan ought to mandate calculation and regularly report Z-score indicators for all scheduled banks. The specific Z-score indicators would need to be customized to the specific capital structure and volatility characteristics of the Pakistani banking system and be used in tandem with existing GNPLR reporting mandated under Basel III.

Macroeconomic Sensitivity and Countercyclical Buffers

The group of macroeconomic driving factors in the literature review (41 papers) demonstrated a stable and significant relationship between macroeconomic conditions and bank credit risk outcomes. Reductions in GDP, increases in inflation, rising interest rates, and falling exchange rates were all observed to correlate with higher NPL ratios across different institutional settings (Espinoza & Prasad, 2010; Nkusu, 2011; Klein, 2013). For regulators, this suggests that static, point-in-time capital adequacy assessments fail to account for the cyclical vulnerability of the bank loan portfolio. Therefore, authorities should implement or enhance countercyclical buffer mechanisms that allow for automatically adjusting capital ratios according to macroeconomic indicators. For Pakistan, where exchange rate movements and inflation have both traditionally exhibited a greater magnitude relative to its peers among developing countries, greater weight on exchange rate and inflation as conditioning variables should be placed in determining such buffer levels, as they were consistently found to have a significant influence on credit risk in an import-oriented economy (Javed & Anwar, 2018).

Islamic Banking Regulatory Architecture

For the Islamic banks cluster (27 studies), there was some mixed but suggestive evidence that Islamic banks demonstrate different risk-taking patterns in times of stress, and several of them reported a better performance of NPL ratios in crisis periods (Beck et al., 2013; Cihak & Hesse, 2010). This implies directly for Pakistan-being one of the world's largest dual banking systems in terms of percentage of Islamic assets-that the existing regulatory framework that applies the capital adequacy requirements under Basel III to both conventional and Islamic banks in largely an homogenous fashion seems not to address sufficiently the risk-taking behaviors that arise from the structure of asset-backed financing, profit and loss sharing terms, and the layer of Shariah governance. Regulators may need to contemplate a separate supervisory framework for Islamic banking that accommodates the differences, such as designing Shariah-specific stress tests and incorporating the Shariah governance quality into the supervisory evaluation framework. Creating an Islamic Banking Stability Monitoring Dashboard within the SBP is another feasible idea to implement this initiative.

11.2 Implications for Commercial Banks

Integrating Multi-Metric Credit Risk Assessment

From a thematic coding perspective, it became apparent that the measure of credit risk has dramatically progressed from a traditional single indicator approach to the accepted multi-metric model, which takes into consideration backward-looking variables (GNPLR), composite stability indicators (Z-score), and forward-looking market-based measures (distance-to-default) to capture bank vulnerability more holistically. Therefore, for commercial banks' credit risk management, reliance solely on the GNPLR constitutes a deadly blind spot, and banks must allocate resources to develop in-house ability to calculate and track Z-score and distance-to-default metrics by using market-provided data and public company filings. This multi-metric approach should be integrated into the bank's internal ICAAP and the bank's risk appetite framework to allow for earlier detection of credit quality deterioration and better and faster re-balancing of credit portfolios. From the performance-risk nexus cluster (58 studies), we also saw that it is a two-way relationship between profitability and credit risk, implying that credit risk is not an inhibitor to performance but rather an integral part of value generation.

Macroeconomic Risk Integration in Credit Decisioning

The persistent result across the 41 studies that macroeconomic variables do impact loan portfolio quality has direct implications for bank operations. Pakistani (and developing country) banks should explicitly include macro-scenario analysis into their loan origination and portfolio management practices, going beyond simply qualitatively assessing macro impacts to using quantitative stress testing that explicitly maps the sensitivity of portfolio quality to shocks like recessions, currency depreciations, and interest rate hikes, based on the PVAR evidence presented in the cluster (25 studies). Banks should build resolute macro-financial risk units that monitor early macro indicators, convert these into forward loan portfolio risks, and recommend a credit committee portfolio orientation to manage these identified risks.

Capital Strategy and Risk-Taking Discipline

All the evidence collected from the performance-risk nexus and macro determinants supports that capital adequacy has the twin functions of a buffer against unforeseen losses and an imposed deterrent against excessive risk-taking. Banks should hold capital buffers more than the minimum required by regulators. The desired level should be adapted to the individual risk profile of each bank and the conditions prevailing in the environment in which it operates. Although, evidence suggests that higher bank capital levels are associated with increased bank resilience (Berger et al., 1995; Berger & Bouwman, 2013), there is no linear relationship since the marginal impact of capital above a certain point diminishes, therefore bank boards and risk committees must determine an optimum level based on the evidence from using the Z-score concept and considering a balance between the cost of equity and the resilience. Evidence that aggressive loan expansion is likely to be followed by increased levels of non-performing loans (Klein, 2013) is evidence for controlled credit growth rates, and targeting market share should be abandoned in favor of risk-adjusted return objectives.

11.3 Implications for Researchers

Adopting Replicable Systematic Review Methodologies

We show in this paper that the explicit use of PRISMA 2020 in conjunction with an integrative-narrative strand offers a transparent and replicable approach to the synthesis of literature in banking and finance. It was found in the methodological reflection cluster of the review that finance literature reviews have shown an important trend over the past 30 years toward more systematic protocols, rather than implicit narrative frameworks (Tranfield et al. 2003; Page et al. 2021). For researchers conducting literature reviews in finance, whether for dissertations, articles, or policy evaluations, it may follow that systematic review methods are becoming more of a requirement, rather than a choice (i.e., Journal editors and supervisors may demand them). The approach illustrated in this paper-the dual-track review methodology incorporating PRISMA with a component of integrative-narrative-may provide useful guidance for research in this area, where the literature base contains a high percentage of regulatory documents, institutional reports, and theoretical contributions that are excluded in solely systematic protocols. Review methods should be reported with the same level of detail as required for the reporting of empirical research design (e.g., Search strategy, inclusion and exclusion criteria, quality assessment, bias control).

Advancing Forward-Looking Credit Risk Research in Pakistan

From both quality assessment and thematic coding, a critical research gap identified was the limited use of forward-looking credit risk measures in the Pakistani banking environment. Even though Z-score and distance to default are a mature literature (Merton, 1974; Bharath and Shumway, 2008), their adoption in Pakistan is minimal, primarily limited by data availability and predominance of GNPLR as a proxy of default risk, an avenue of research worth considerable effort. It would be beneficial to see Pakistan-specific calibration of forward-looking credit risk measures; the methodology applied by research in developing markets with structural similarities should be adopted. The availability of bank-level data through the SBP bulletins, stock exchange, and credit bureaus has made this an achievable goal. Another trend identified by the methodological advances cluster was the emerging and exciting trend of application of machine learning methods in predicting credit risk (Leo et al., 2019), an area with no existing Pakistan-focused study and worth of a dedicated research effort focusing on the application of various techniques like random forest, gradient boosting, and neural networks, and not least, regulatory concerns regarding their use.

Addressing Geographical and Conceptual Bias

In this review, the bias analysis confirmed a substantial problem of geographical bias in the literature on credit risk, as all the studies are in developed countries. This problem hinders the generalizability of the results to emerging markets and gives rise to knowledge asymmetry, which places regulatory bodies in emerging markets at a disadvantage. Researchers must find ways to mitigate this bias by conducting robust and contextualized research in underexplored markets (specifically Pakistan, South Asia, Sub-Saharan Africa, and Central Asia) and disseminating their findings through high-impact internationally indexed journals. Moreover, the cognitive bias towards using GNPLR as the default proxy for risk must be confronted by studies that conduct robust comparisons of other credit risk measures within the same institutional environment and provide evidence of the marginal informational gain from using forward-looking and market-based indicators versus traditional GNPLR. Such analyses would be relevant to both academics and regulatory bodies, thus informing the broader discussion about appropriate measurement frameworks for credit risk in emerging market banking systems.

Filling the Islamic Banking Evidence Gap

Only 27 (out of 281) studies belonged to the Islamic banking cluster. This is not a substantial proportion, given the growing importance of Islamic finance and Pakistan's position as a leading Islamic finance hub. The evidence on whether Islamic banks are indeed differently characterized by risk as compared to conventional banks, though suggestive, is not conclusive. In part, this is due to the relative scarcity of rigorous studies on this topic and due to methodological difficulties in isolating the impact of the Islamic contract structures from other potentially confounding variables, such as the size, regulatory, and market concentration of the banks. Future research endeavors would do well to focus on studies based on matched samples, natural experiments, etc., to shed further light on the causal effect of the Islamic banking model on the credit risk implications. The distinctive characteristics of Islamic banking-the prohibition of interest and provision of loans only on profit-sharing bases, the insistence on the financing to be backed by physical assets, and the role of Shariah governance mechanisms-provide fertile ground for the testing of a set of hypotheses on how institutions' structures affect financial stability, with a far wider relevance outside the borders of Pakistan.

12 Conclusion

In this paper, the detailed methodological process for a systematic literature review of the relationship between credit risk and bank performance in the context of the dual banking system in Pakistan is presented. Across six decades of research and 281 empirical papers from Altman's (1968) pioneering Z-score model to current applications of machine learning in credit risk analysis, a methodological framework has been developed that combines the systematic transparency of the PRISMA 2020 protocol with the context sensitivity of an integrative-narrative review, resulting in a literature evidence base that is both replicable and context-relevant.

Following PRISMA 2020 stages (identification, screening, eligibility, inclusion), four stages resulted in the initial systematic literature review synthesis (n = 247) of the literature drawn from 1,350 records. The systematic review was supplemented with an integrative-narrative search update in 2025, which identified an additional 34 papers to create a final literature evidence base of n = 281 studies classified into six thematic clusters. The evaluation of research articles was based on three aspects: methodological quality, context relevance, and scholastic importance. The five sources of bias considered were publication, geographical, conceptual, researcher, and integrative, with specified means of avoiding them.

The contributions of this study to the field's methodological knowledge are triple. First, the suitability and the need to modify PRISMA 2020 for financial research are illustrated, considering the diverse literature found in banking and finance. Second, the combined systematic-integrative approach provides a framework that balances PRISMA's reproducibility with the context inclusivity needed for research in emerging markets. Third, the study exhibits a transparent, auditable linkage between literature, the evidence base, and the empirical research framework, as all variable selection was grounded in the findings rather than imposed on them.

As academic research in finance continues to expand, methods of reviewing it must also continue to evolve. Systematic protocols borrowed from health sciences serve as the first base for this. Financial evidence, however, has different methodological characteristics; source diversity, institutional context, and the long shelf-life of established theories, compared to those of medical studies, not to mention geographical concentration of literature, primarily in developed economies. This paper provides a methodology that could offer a template for dealing with such literature-based research in banking and finance.

References

Afzal, A., & Mirza, H. (2012). Determinants of non-performing loans in Pakistan. Pakistan Economic and Social Review, 50(2), 163 to 180.
Ahmed, A., & Malik, M. (2020). Macroeconomic determinants of non-performing loans. Journal of Asian Economics, 67, 101207. https://doi.org/10.1016/j.asieco.2020.101207
Ali, R., & Ahmed, J. (2024). Venturing into risk management. Examining credit risk and performance in Pakistani commercial banks through panel VAR analysis from a business perspective. Journal of Entrepreneurship and Business Venturing, 4(1). https://doi.org/10.56536/jebv.v4i1.101
Altman, E. I. (1968). Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. Journal of Finance, 23(4), 589 to 609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x
Altman, E. I., Brooks, B., Resti, A., & Sironi, A. (2005). The link between default and recovery rates. Theory, empirical evidence, and implications. Journal of Business, 78(6), 2203 to 2228. https://doi.org/10.1086/497044
Arellano, M., & Bond, S. (1991). Some tests of specification for panel data. Monte Carlo evidence and an application to employment equations. Review of Economic Studies, 58(2), 277 to 297. https://doi.org/10.2307/2297968
Beck, T., Demirgüç-Kunt, A., & Merrouche, O. (2013). Islamic vs. conventional banking. Business model, efficiency, and stability. Journal of Banking and Finance, 37(2), 433 to 447. https://doi.org/10.1016/j.jbankfin.2012.09.016
Berger, A. N., & Bouwman, C. H. (2013). How does capital affect bank performance during financial crises. Journal of Financial Economics, 109(2), 287 to 307. https://doi.org/10.1016/j.jfineco.2013.02.008
Berger, A. N., Herring, R. J., & Szegö, G. P. (1995). The role of capital in financial institutions. Journal of Banking and Finance, 19(3 to 4), 393 to 430. https://doi.org/10.1016/0378-4266(94)00101-5
Berkowitz, M. K., Qiu, J., & Yeo, J. (2022). Geographical bias in academic finance research. Journal of Financial Economics, 144(2), 598 to 619. https://doi.org/10.1016/j.jfineco.2021.12.002
Boahene, S. K., Dasah, J., & Agyei, S. K. (2012). Credit risk and profitability of selected banks in Ghana. Research Journal of Finance and Accounting, 3(7), 48 to 57.
Burki, A. A., & Niazi, G. S. K. (2006). Impact of financial reforms on efficiency of banks. A study of the banking sector of Pakistan. Lahore Journal of Economics, 11(1), 89 to 108.
Cihák, M., & Hesse, H. (2010). Islamic banks and financial stability. An empirical analysis. Journal of Financial Services Research, 38(2 to 3), 95 to 113. https://doi.org/10.1007/s10693-010-0089-0
Cochrane, A. L. (1972). Effectiveness and efficiency. Random reflections on health services. Nuffield Provincial Hospitals Trust.
Denyer, D., & Tranfield, D. (2009). Producing a systematic review. In D. Buchanan & A. Bryman (Eds.), The Sage handbook of organizational research methods (pp. 671 to 689). Sage.

Duffie, D., Saita, L., & Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. Journal of Financial Economics, 83(3), 635 to 665. https://doi.org/10.1016/j.jfineco.2005.10.011

Egger, M., Smith, G. D., & Altman, D. G. (Eds.). (2008). Systematic reviews in health care. Meta-analysis in context (2nd ed.). BMJ Publishing Group.

Espinoza, R., & Prasad, A. (2010). Nonperforming loans in the GCC banking system and their macroeconomic implications (IMF Working Paper No. WP/10/224). International Monetary Fund.

Grant, M. J., & Booth, A. (2009). A typology of reviews. An analysis of 14 review types and associated methodologies. Health Information and Libraries Journal, 26(2), 91 to 108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic reviews over qualitative research. BMJ, 364, k4215. https://doi.org/10.1136/bmj.k4215

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020. An R package and Shiny app for producing PRISMA 2020 flow diagrams. Systematic Reviews, 11, 89. https://doi.org/10.1186/s13643-022-01918-0

Hart, C. (2018). Doing a literature review. Releasing the research imagination (2nd ed.). Sage.

Hasan, M., & Dridi, J. (2011). The effects of the global crisis on Islamic and conventional banks. A comparative study (IMF Working Paper No. WP/11/201). International Monetary Fund.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). Cochrane handbook for systematic reviews of interventions (Version 6.0). Cochrane.

International Monetary Fund. (2016). Pakistan. Financial system stability assessment (IMF Country Report No. 16/62). International Monetary Fund.

Javed, A., & Anwar, S. (2018). Impact of macroeconomic factors on non-performing loans. Evidence from Pakistan. Pakistan Development Review, 57(4), 385 to 401.

Klein, N. (2013). Non-performing loans in CESEE. Determinants and impact on macroeconomic performance (IMF Working Paper No. WP/13/72). International Monetary Fund.

Leo, M., di Graziano, M., & Lazzari, G. (2019). Machine learning in credit risk. A review. Expert Systems with Applications, 137, 367 to 385. https://doi.org/10.1016/j.eswa.2019.07.001

Louzis, D. P., Vouldis, A. T., & Metaxas, V. L. (2012). Macroeconomic and bank-specific determinants of non-performing loans in Greece. Journal of Banking and Finance, 36(4), 1012 to 1027. https://doi.org/10.1016/j.jbankfin.2011.10.012

Love, I., & Zicchino, L. (2006). Financial development and dynamic investment behavior. Evidence from panel VAR. Quarterly Review of Economics and Finance, 46(2), 190 to 210. https://doi.org/10.1016/j.qref.2005.11.007

Merton, R. C. (1974). On the pricing of corporate debt. The risk structure of interest rates. Journal of Finance, 29(2), 449 to 470. https://doi.org/10.1111/j.1540-6261.1974.tb03058.x

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses. The PRISMA statement. PLOS Medicine, 6(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Nkusu, M. (2011). Nonperforming loans and macrofinancial vulnerabilities in advanced economies (IMF Working Paper No. WP/11/161). International Monetary Fund.

Okoli, C. (2015). A guide to conducting a standalone systematic literature review. Communications of the Association for Information Systems, 37, Article 43. https://doi.org/10.17705/1CAIS.03743

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement. An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Petticrew, M., & Roberts, H. (2006). Systematic reviews in the social sciences. A practical guide. Blackwell Publishing.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.). (2005). Publication bias in meta-analysis. Prevention, assessment, and adjustments. John Wiley and Sons.

Shah, M., & Ahmed, S. (2020). Bank performance and credit risk in Pakistan. An empirical analysis. Journal of Risk and Financial Management, 13(12), 297. https://doi.org/10.3390/jrfm13120297

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8, Article 45. https://doi.org/10.1186/1471-2288-8-45

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British Journal of Management, 14(3), 207 to 222. https://doi.org/10.1111/1467-8551.00375

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future. Writing a literature review. MIS Quarterly, 26(2), xiii to xxiii. https://doi.org/10.2307/4132319

Zaidi, S. A. (1999). Issues in Pakistan economy. A political economy perspective. Oxford University Press.