# Birds of a Feather Flock Together: A Deep Learning Bank Co-Lending Network Risk Measure

Junbin Gao    Yunying Huang *    Ruikun Li    Buhui Qiu    Eliza Wu

This version: November 2024

**Abstract**

We employ an advanced deep learning technique, the Graph Neural Network, to develop a Co-Lending Graph Neural Network (CoLGNN) model to address the complexity of financial interconnectedness. This model maps the risk spillovers among financial institutions in the syndicated lending market. By leveraging comprehensive data on bank characteristics, lending attributes, and the network's topological structure, the model generates a co-lending network risk measure (CLN score) for each bank. We show that this measure effectively captures risk spillovers in the co-lending network, serving as an early-warning indicator of each bank's embedded network risk. Our analysis demonstrates that the CLN score robustly predicts future bank risks and profitability up to two years ahead, applicable to both public and private banks. Furthermore, we find that the measure's predictive power is stronger for banks identified as vulnerable due to factors such as small bank size, poorer performance, lower capital adequacy, and higher complexity. Finally, we validate the risk measure through a quasi-natural experiment on bank credit-rating downgrades, demonstrating that such events propagate risk from focal banks to neighboring co-lending banks.

**Keywords:** deep learning; graph neural networks; bank risks; syndicated lending

**JEL Codes:** G17, G21

---

Junbin Gao (Email: junbin.gao@sydney.edu.au), Yunying Huang (Email: yunying.huang@sydney.edu.au), Ruikun Li (Email: ruikun.li@sydney.edu.au), Buhui Qiu (Email: buhui.qiu@sydney.edu.au) and Eliza Wu (Email: eliza.wu@sydney.edu.au) are with University of Sydney Business School, University of Sydney, NSW 2006, Australia. For helpful comments, we thank Mingze Gao, Iftekar Hasan, Yingdong Mao, Zheyao Pan, Hanlin Shang, Gary Tian, Shang-Jin Wei, Xuebin Zheng, seminar participants at Curtin University and Macquarie University, and participants at the 2023 Machine Learning Methods for Business Workshop and the 2024 Finance PhD Colloquium at the University of Sydney Business School. This paper received the Best Paper Award at the 2024 Finance PhD Colloquium at the University of Sydney Business School. We gratefully acknowledge the research support from the Australian Research Council Discovery Project Grant (DP210102611). All errors are our own.

*Corresponding author.

*"Complex links among financial market participants and institutions are a hallmark of the modern global financial system. Across geographic and market boundaries, agents within the financial system engage in a diverse array of transactions and relationships that connect them to other participants." (Yellen, 2013)*

# 1 Introduction

This study employs an advanced machine learning technique to develop a Co-Lending Graph Neural Network (CoLGNN) model, which maps the risk spillovers among financial institutions in the syndicated lending market. In the rapidly evolving global economies, the complex interdependencies among financial institutions are increasingly recognized as crucial in shaping systemic risk dynamics. Many studies highlight the significant role of financial system interconnectedness in the cascading of failures and propagation of risks (e.g. Allen & Gale, 2000; Elliott et al., 2014; Acemoglu et al., 2015; Cabrales et al., 2017). The global financial crisis (GFC) provides important lessons on how risks spread throughout the financial system (Yellen, 2013). Recent examples include the risk spillover to the banking industry from the Dexia bank bailout in 2011, the Cyprus banking crisis in 2013, the failure of Banco Espírito Santo in 2014, and the bankruptcy of Silicon Valley Bank (SVB) in 2023. Current regulations, such as the Basel capital requirements, are designed to limit banks' risks when viewed in isolation (Acharya et al., 2017). However, the interconnected nature of the banking industry highlights the pressing need for more nuanced risk measures to refine individual banks' risk management practices and for regulators to gain a deeper understanding of risk transmission mechanisms to enhance financial stability.

While mapping risk transmission across various segments of the financial landscape has received extensive attention, one area that remains relatively underexplored is the cross-sectional interdependencies in the syndicated loan market. This market, representing one of the largest sources of corporate financing, facilitates large-scale loans to corporations by syndicated lenders, with a total value of $4.7 trillion US dollars globally in 2022, according to Refinitiv Dealscan. The syndicated loan market's inherently asymmetric structure propagates risks from lead arrangers to participants, introducing unique challenges and complexities in modeling risk dynamics. Furthermore, although stock returns function as a forward-looking metric that encapsulates the efficient market's integration of information and reflects market expectations (e.g., Fama, 1970, 1991), the reliance on public stock market data for risk metrics (e.g., Merton, 1974; Tobias & Brunnermeier, 2016; Nagel & Purnanandam, 2020) poses additional challenges for regulators and banks. Private banks, integral to the banking sector's operational fabric and risk spillover network, lack market-based risk metrics due to their absence from public equity markets. Ignoring private banks in market-based risk assessment overlooks critical elements in risk transmission mechanisms. Addressing these gaps with more comprehensive risk metrics can lead to the development of more robust financial regulatory frameworks.

We construct a directed co-lending network reflecting the asymmetrical relationships inherent in the syndicated loan market and develop a customized deep learning framework—the Co-Lending Graph Neural Network Model (CoLGNN)—to analyze risk spillover within this network. In the syndicated lending market, banks diversify their risk exposure via co-lending (loan syndication) with other lenders, which is functioning crucially on the perceived quality and reputation of the lead bank (lead arranger) (e.g., Pichler & Wilhelm, 2001; Ivashina, 2009; Gopalan et al., 2011). This market is characterized by information asymmetry between lead arrangers and participants, with the latter having limited oversight in the screening and monitoring of loans (Ivashina, 2009; Ivashina & Scharfstein, 2010b). The asymmetry motivates us to model connections as directed edges from lead arrangers to participants within syndicates. Given the dual roles banks may play, acting as lead arrangers in some deals and participants in others, we consider bank holding companies as nodes and co-lending relationships (loans) as edges in the network. The directed edges serve as a risk transmission channel to propagate potential shocks and influences from lead banks.

However, the complex topological nature of financial networks imposes methodological challenges. Traditional empirical approaches such as centrality metrics, offer a limited perspective on the multidimensional relationships and risk spillover mechanisms by focusing primarily on the relative importance of individual nodes. Our study differentiates from traditional approaches by applying the CoLGNN model to dissect bank-level risk within the co-lending network. As a cutting-edge and trending deep learning framework specialized in a non-Euclidean domain,[1] graph neural networks (GNNs) have become the preferred approach for prediction tasks across various network-related domains, including social networks, supply chains, and computer vision (Xu et al., 2019; Wu et al., 2020).[2] GNN transcends the limitations of traditional empirical methods by integrating comprehensive bank features with the topological structure of banks' lending portfolios and the high-dimensional characteristics of syndicated loans. Specifically, the developed CoLGNN model utilizes the "message-passing paradigm" to capture the risk spillover within the co-lending network (Chami et al., 2022). We design a graph diffusion convolution operator for CoLGNN to aggregate the neighboring node information and high-dimensional edge attributes. The diffusion convolution modules are specially designed to capture the directional spillover within the co-lending network, combining the characteristics of bank characteristics and topological information, as well as the high-dimensional syndicated loan features.

Empirically, using a comprehensive sample of U.S. syndicated loans from Refinitive DealScan and bank holding company characteristics from the FR Y9-C database, we utilize the developed CoLGNN model to generate a bank-level co-lending network risk score (CLN score) for each bank (both public and private banks) at the year-quarter level. We construct a series of rolling-window co-lending networks at each year-quarter end using all syndicated loans originated in the past five years. Each co-lending network is an attributed graph with bank characteristics as node features and loan characteristics as edge features. Adopting a

[1]The non-Euclidean data imply that there are no such properties as global parameterization, a common system of coordinates, vector space structure, or shift invariance (Bronstein et al., 2017). Meanwhile, the Euclidean data (the grid-like structure) are organized in fixed dimensions, such as panel data. For example, the supply chain and social networks are non-Euclidean structures due to their inherent complex topological structure. The data becomes Euclidean if we measure the centrality of each node $i$ at time $t$ and organize it as panel data.

[2]For example, GNN-related frameworks are leading in many graph-related benchmark datasets in deep learning academic papers. See https://paperswithcode.com/task/node-classification

semi-supervised learning scheme, we train the CoLGNN model to estimate the CLN score across the network using a subset of the public banks for which risk labels—derived from stock market performance—are known. Specifically, motivated by the market-based risk measures relying on the informativeness of stock prices, we label the public banks exhibiting the top quartile of stock performance during a quarter as "safe" and public banks in the bottom quartile as "risky" for training purposes.[3] By mapping the risk transmission via "message passing" and leveraging the topological structure of the co-lending networks via the convolution module, our trained CoLGNN model estimates a risk score for each node (i.e., each bank) at the prediction stage.

Using a comprehensive panel dataset spanning 1996 to 2020, we examine the predictive power of the CoLGNN-derived bank CLN score in forecasting future bank risks. Our analysis shows that the bank CLN score significantly predicts future bank loan loss provisions. Importantly, this predictive ability extends to both unlabeled public and private banks in out-of-sample predictive regressions, highlighting the CLN score as an innovative early-warning measure particularly useful for banks that lack market-based risk metrics. The inclusion of the CLN score significantly improves the predictive regressions, as evidenced by increased adjusted R-squared values and decreased root-mean-square errors. This advancement is primarily due to the model's incorporation of individual bank characteristics, loan-specific attributes, and network topology, capturing the dynamics of risk spillover within the co-lending network. Our results remain robust after controlling for extensive bank characteristics, lending specializations, year-quarter fixed effects, and bank fixed effects.

Furthermore, as the CLN score captures risk spillovers from neighboring co-lending banks within the network, we posit that banks with certain vulnerabilities and fragilities—specifically, smaller size, weaker earnings performance, higher return volatility, and

[3] We use raw stock returns rather than factor-adjusted or idiosyncratic returns for our labeling strategy as both systematic and idiosyncratic risks are relevant for bank stability. Raw returns provide a comprehensive measure that captures both market-wide and bank-specific risks, consistent with the efficient market's incorporation of all available information (Fama, 1970). Moreover, as shown by Campbell et al. (2008), both systematic and idiosyncratic components of stock returns contain important information about bank risk.

lower capital adequacy—would be more sensitive to these spillovers. Consistent with this hypothesis, our empirical results demonstrate that the predictive power of the bank CLN score is indeed significantly stronger for such banks. This suggests that the future risk profiles of these banks are more influenced by the risk dynamics of their adjacent banks within the co-lending network. Additionally, we expect complex or opaque banks to face more challenges in effective risk management, leading to a greater sensitivity to the network's risk transmission dynamics. Our findings also indicate that the predictive power of the bank CLN score is stronger for banks with greater complexity or opacity.

Moreover, We find that bank CLN score exhibits similar predictive power for other risk metrics beyond bank loan loss provisions, including non-performing loans, default probability (Merton, 1974), bank default probability (Nagel & Purnanandam, 2020), and future stock return idiosyncratic volatility. These findings collectively affirm that the bank CLN score, estimated by the CoLGNN framework, provides valuable and insightful forecasts of future bank risks. Consistent with its predictive ability for future bank risks, we also find that a higher bank CLN score significantly predicts lower future bank profitability. Importantly, we find that controlling for bank stock returns does not diminish the predictive power of the bank CLN score, which even significantly outperforms stock returns in predicting risk and profitability for public banks whose stock performance lies in the interquartile range.[4] In terms of the time-varying predictability of the bank CLN score, we find persistent predictive power across both halves of the sample period. Additionally, we observe an increase in the predictive power of the bank CLN score during the global financial crisis when risk transmission intensified.

Finally, we validate the early-warning capabilities of the bank CLN score for future bank risks and its role in the risk transmission mechanism within the co-lending network using

[4] If the predictive power of the CLN score arose from banks with high CLN values being subject to common negative shocks with the labeled "risky" banks, then controlling for bank stock returns should substantially diminish the CLN score's effectiveness in predicting future bank risks and profitability. However, our findings suggest otherwise. Specifically, we find that the bank CLN score significantly outperforms quarterly stock returns in predicting future non-performing loans and ROA within the subsample of unlabeled public banks, supporting its role in capturing spillover risk through the co-lending network.

a quasi-natural experiment. We employ S&P credit-rating downgrade events to investigate whether the negative shocks on focal banks from long-term credit downgrades propagate through their co-lending networks and adversely affect connected neighboring banks. Using propensity score matching, we select control banks—those without direct ties to the focal bank yet exhibiting similar characteristics to the treated banks. Through stacked-cohort difference-in-differences (DiD) and dynamic DiD estimations, we find that, following a downgrade event, banks directly connected to the downgraded bank experience significantly higher increases in their co-lending network risk score (i.e., the CLN score) compared to control banks. These results confirm the effectiveness of our CoLGNN model and the derived bank CLN score in capturing and quantifying risk transmission in the co-lending network.

Our study contributes to the literature in several ways. First, we pioneer the application of GNN on the universe of syndicated loans to develop a Co-Lending Graph Neural Network. Our design is motivated by extant literature on information asymmetry and risk sharing within syndicated loans (e.g. Ivashina, 2009; Blickle et al., 2020). Prior studies have shown that loan outcomes and bank risks could be influenced by the syndicate structure (e.g. Lim et al., 2014; Gao et al., 2023) and lending relationships (e.g. Bharath et al., 2011). We contribute to the literature by developing a novel directed network design that captures lending relationships, lending structures, and directional risk spillovers within the syndicated loan market. Our study enriches the syndicated lending literature by employing deep learning to quantify the implications of risk spillovers for future bank risks.

Second, our research contributes to network studies in financial economics. Existing literature emphasizes the interconnectedness of financial institutions as a crucial channel for risk spillovers (e.g. Battiston et al., 2012; Golub & Jackson, 2012; Elliott et al., 2014; Acemoglu et al., 2015; Anderson et al., 2019). We extend the study of the network to the syndicated lending market, one of the most important sources of corporate financing. We exploit the topological structure of the co-lending network to estimate early-warning risk metrics for banks operating in the syndicated lending market.

Third, our study contributes to the literature on bank risk metrics. Existing literature often relies on stock market information to develop measures for banks' default risk (e.g. Merton, 1974; Nagel & Purnanandam, 2020) and systematic risk (e.g. Tobias & Brunnermeier, 2016). Our paper contributes to this literature by employing a semi-supervised deep learning design on the co-lending network to extend risk evaluation to banks that lack market-based risk metrics. This approach provides a more comprehensive risk metric that captures the risk spillovers in the co-lending network for the banking industry.

Finally, the study contributes to the burgeoning field of artificial intelligence (AI) in finance. As AI algorithms specialize in dealing with high-dimensional and non-structural data, recent studies identify innovative empirical models (Gu et al., 2020), new investment strategy (e.g. Cong et al., 2021), and alternative data source (e.g. Li et al., 2021; Cao et al., 2023; Jiang et al., 2023). Our study contributes to this trending area by employing a cutting-edge AI framework that specializes in dealing with non-Euclidean data and proposes an innovative deep learning model, CoLGNN, for risk management in the banking industry.

The rest of the paper proceeds as follows. Section 2 develops the main hypotheses. Section 3 discusses the design of the co-lending network and introduces the CoLGNN model. Section 4 discusses the bank-level CLN score, describes our data sources, and outlines key measurements. Section 5 presents the baseline risk-prediction results and explores the underlying risk-transmission mechanism. Section 6 validates the risk-transmission mechanism using a quasi-natural experiment on credit-rating downgrade events. Section 7 concludes. The Appendix provides detailed variable definitions and supplementary empirical results.

# 2 Hypothesis Development

The global financial system's high level of interconnectedness facilitates risk spillovers and contagion, particularly within tightly coupled financial networks (e.g., Elliott et al., 2014; Acemoglu et al., 2015; Anderson et al., 2019). The structure and topology of these networks are critical determinants of how risks are transmitted and amplified, influencing both the speed and magnitude of contagion (Haldane & May, 2011; Glasserman & Young, 2015). Direct and indirect connections within these networks serve as channels through which financial shocks propagate, exacerbating vulnerabilities across institutions (Battiston et al., 2012). This highlights the pressing need for robust risk assessment frameworks capable of capturing the complexities of financial interdependencies (Brunnermeier & Oehmke, 2013).

In this study, we focus on interconnected financial networks within the context of co-lending in the syndicated loan market. In a typical syndicated loan structure, multiple financial institutions collaborate to provide a loan to a specific borrower, with lead banks (or lead arrangers) playing a central role. Lead banks are responsible for building relationships with borrowers, gathering critical information, negotiating loan terms, originating the deal, and marketing the loan to participant banks and institutional investors (e.g., Holmstrom & Milgrom, 1987; Pichler & Wilhelm, 2001). Beyond facilitating loan origination, lead banks bear the primary responsibility for conducting ex-ante due diligence and ex-post monitoring of borrowers.

This syndicated lending structure inherently creates a web of financial interdependencies, where risks are not confined to a single institution but propagate through the network of lenders. Such propagation is especially pronounced when lead banks fail to accurately assess borrower risk or when unforeseen shocks compromise a borrower's repayment capacity. In these scenarios, cascading disruptions extend beyond the lead bank, adversely affecting participant lenders who often rely on the lead bank's due diligence and reputation when joining the loan (Ivashina & Scharfstein, 2010a; Gopalan et al., 2011). Thus, understanding and measuring risk propagation within the syndicated lending network is essential for designing effective risk management strategies.

To address this, we employ graph neural networks (GNNs) to model the intricate interplay of bank characteristics, loan terms, and network topology, leveraging bank stock performance to label a subset of public banks with good or poor stock performance as "safe" or "risky"

(Campbell et al., 2008).[5] Existing studies primarily examine banks' cross-sectional and time-series traits in isolation, often neglecting the topological structure of financial networks and the dynamics of syndicated lending relationships. Our GNN model fills this gap by using a message-passing paradigm to trace risk transmission pathways from lead banks to participant banks within the co-lending network. Using semi-supervised learning, we train the model on labeled public bank samples to estimate a co-lending network risk measure for all banks on the network, both public and private banks.

The interconnectedness of the syndicated lending network facilitates the transmission of risks among co-lending banks. When a lead bank faces significant risks—such as those arising from a borrower default—these risks can propagate to other banks that participate in the same loan syndicate. The close financial ties established through co-lending relationships serve as channels for risk transmission, exposing participant banks to vulnerabilities stemming from the lead bank's distress. This interconnected structure amplifies the impact of localized shocks, as the financial health of co-lending banks becomes increasingly interdependent. In this context, a higher level of co-lending network risk—arising from exposure to risky lead banks—places participant banks in a precarious position, increasing their likelihood of facing greater financial risks in the future. By quantifying this network-level co-lending risk and tracing its transmission pathways, we aim to investigate how risks embedded in co-lending networks influence individual banks' future bank risks. We develop the following hypothesis.

**Hypothesis 1**

*Higher co-lending network risk of a bank predicts greater bank risks in the future.*

Furthermore, market-based risk metrics, as demonstrated in seminal works by Merton (1974), Campbell et al. (2008), and Nagel and Purnanandam (2020), inherently limit risk

[5]Stock returns serves as a robust, forward-looking proxy for assessing the risks of public banks, anchored in the principles of market efficiency and informativeness (e.g., Fama, 1970, 1991; Fama & French, 2015). It reflects the market's perception of a bank's financial health. Stock prices are also a key input for market-based default risk measures (Merton, 1974; Nagel & Purnanandam, 2020).

evaluation to publicly traded banks. These approaches inadvertently exclude a significant portion of the banking sector—namely, private banks—due to their absence from public stock markets. By leveraging the capabilities of the semi-supervised GNN model, we extend the measurement of co-lending network risk to include private banks.

The message-passing paradigm uniquely enables us to estimate the risk profiles of private banks by integrating their inherent characteristics, lending activities, and the influences of their interconnectedness with public banks within the co-lending network. We posit that the observable attributes of labeled public banks, combined with the structural intricacies of the co-lending network, can be used to infer the risk profiles of neighboring private banks. Accordingly, we develop the following hypothesis.

**Hypothesis 2**

*The co-lending network risk measure exhibits robust out-of-sample performance in predicting the future risks of private banks.*

Moreover, we anticipate variations in the risk predictive power of our co-lending network risk measure based on differences in bank size, recent performance, financial stability, capital adequacy, and structural complexity. These attributes play a critical role in shaping a bank's fragility within the co-lending network and its sensitivity to risk spillovers from the network.

First, smaller banks are likely to exhibit greater vulnerability to spillovers from neighboring institutions within the network. With lending portfolios that are often less diversified, smaller banks may experience disproportionately severe impacts from adverse events involving closely connected banks. Supporting this, Giannetti and Saidi (2019) find that banks with smaller market shares are less capable of internalizing negative spillovers, emphasizing their heightened exposure during periods of shock propagation. Second, banks with poor performance or higher earnings volatility may be less resilient to risk spillovers from the network. Such banks have diminished financial cushions, are more reliant on external relationships, face greater market pressures, and are less capable of absorbing and mitigating the impact of network risks. These factors exacerbate their fragility in interconnected financial networks, leaving them more sensitive to spillovers.

Third, banks with lower levels of capital are less able to absorb shocks, making them more susceptible to the adverse effects of network-induced risks. Limited capital buffers exacerbate their fragility, amplifying the consequences of risk spillovers. Lastly, banks with greater structural complexity may find it more challenging to effectively manage risk and lending relationships. This complexity increases their sensitivity to network risk spillovers, as the intricate nature of their operations can magnify their exposure to spillover effects. We propose the following hypothesis.

**Hypothesis 3**

*The risk predictive power of the co-lending network risk measure is significantly stronger for banks with smaller size, negative earnings shocks, higher return volatility, lower capital ratios, and greater structural complexity.*

# 3 Modelling Bank Risk via Graph Neural Network

## 3.1 Co-Lending Network Design in the Syndicate Lending Market

The syndicated lending market features a unique structure in which lead arrangers play a crucial role in screening, initiating, and monitoring loans. This arrangement creates a network of financial interdependencies, motivating the development of a directed co-lending network to analyze network spillover risk. In this network, banks (nodes) are connected through co-lending relationships (edges), with a directed edge representing the flow of information and risk from a lead arranger to a participant bank in a syndicated loan (Ivashina, 2009; Benmelech et al., 2012).

Directed edges in the co-lending network illustrate how information and risk are transmitted, with their orientation reflecting the flow from lead banks to other participants within the same loan syndicate. Each directed connection thus symbolizes not only a co-lending relationship but also a potential risk-spillover channel. The network aggregates co-lending activities across all syndicated loans within a specified time frame, resulting in a dynamic series of co-lending networks that can be analyzed over rolling windows (e.g., five-year intervals). Mathematically, we define the co-lending network as an attributed graph as follows:

**Co-Lending Network** Each co-lending network is an attributed graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, \ldots, v_n\}$ is the set of banks and and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of edges, where a directed edge $e_{i,j} \in \mathcal{E}$ pointing from bank $i$ to bank $j$. $\mathcal{N}(v) = \{u : (u, v) \in \mathcal{E}\}$ denotes the neighbor set of bank $v$. $\mathbf{X} \in \mathbb{R}^{n \times f}$ and $\mathbf{E} \in \mathbb{R}^{m \times s}$ are the bank characteristics and loan characteristics, respectively.

## 3.2 Co-Lending Graph Neural Network (CoLGNN)

### 3.2.1 From Message-Passing to Mapping Risk Spillover

Conventional econometric methods often struggle to capture the rich topological information inherent in financial networks. Network centrality measures are frequently used to analyze the role of focal entities in networks (e.g., El-Khatib et al., 2015; Rossi et al., 2018; Richmond, 2019). However, these approaches may not be able to fully capture the multidimensional relationships and dynamic interactions that characterize complex financial systems. Similarly, traditional deep learning (DL) frameworks, while effective at handling high-dimensional datasets, are primarily designed for Euclidean data structures. For instance, convolutional neural networks (CNNs) use hidden layers to identify spatially localized features, but they are limited to fixed-dimension data, requiring feature engineering to incorporate network structures.[6] This adaptation can be challenging, as it demands embedding representations that approximate network structure.[7]

[6]Machine learning with networks requires feature embedding representations that encode node similarities, which can be difficult to capture in traditional frameworks.

[7]Common methods such as adjacency matrix might lead to inefficient computation problems and might not be permutation invariant. When the number of nodes and the number of edges per node are highly variable, this could lead to very sparse adjacency matrices which is space inefficient. Moreover, many adjacency matrices might encode the same graph structure, and there is no guarantee that different matrices would produce the same estimation in a deep neural network (not permutation invariant).

To address these challenges, we introduce the **Co-L**ending **G**raph **N**eural **N**etwork framework (CoLGNN), a customized deep learning framework built on Graph Neural Networks (GNNs), which have emerged as a powerful class of methods for handling graph-structured data. GNNs apply optimizable transformations across all graph components—nodes, edges, and global contexts—enabling the model to learn intricate patterns directly from the graph's structure. Recent advancements in GNN-related methods exhibit significant out-performance in a wide range of tasks such as node classification, link prediction, and clustering (Bronstein et al., 2017; Xu et al., 2019; Wu et al., 2020). GNN captures the locality of each node, aggregates the neighborhood information, and stacks multiple layers to estimate the coefficient and perform prediction tasks.

CoLGNN is specifically tailored to study risk dynamics in co-lending networks as defined in Section 3.1. It employs a message-passing mechanism, which iteratively aggregates information from each node's neighbors, allowing the model to learn and adapt to the topological characteristics unique to co-lending structures.

**Message-Passing Paradigm** The Message-Passing paradigm follows a multi-layer scheme of updating node representations based on neighborhood aggregation. Let $\mathbf{h}_i^{(l)}$ represent node $i$'s embedding at the $l$-th layer. The Message-Passing scheme is defined as:

$$\mathbf{h}_i^{(l)} = \mathrm{UP}\left(\mathbf{h}_i^{(l-1)}, \mathrm{AGGR}(\{\mathbf{h}_j^{(l-1)} : j \in \mathcal{N}(i)\})\right), \tag{1}$$

where the $\mathrm{UP}(\cdot)$ and $\mathrm{AGGR}(\cdot)$ are aggregation function and update function, respectively (Hamilton et al., 2017). Starting from initial node features $\mathbf{H}^{(0)} = \mathbf{X}$, CoLGNN learns refined representations $\mathbf{H}$ after multiple layers, effectively capturing how risk signals propagate through the network.[8]

The message-passing mechanism in GNNs is particularly well-suited for modeling risk spillover, as it mimics how risk spreads from lead arrangers to other banks within a co-lending

[8] $\mathbf{H}^{(l)} = \left[\mathbf{h}_i^{(l)} : i = 1, ..., n\right]$ for $l \in \{1, \dots, L\}$

network. In syndicated lending markets, lead arrangers—responsible for loan origination, due diligence, and monitoring—not only allocate lending shares but also potentially transmit risks to participant banks. Through message-passing, information (or "messages") about a lead bank's risk level is iteratively propagated along directed edges to other banks. Each message encapsulates data on risk characteristics, enabling the model to aggregate and refine risk signals across the network. This iterative process effectively captures the diffusion of risk throughout the syndicate, accounting for both direct and indirect transmission pathways.

Figure 1 presents a flowchart of the CoLGNN framework, illustrating the key stages involved in modeling risk within the co-lending network at a specific time point $t$. The process begins with constructing the network and establishing directed co-lending connections. Next, banks are labeled through a binary classification system, identifying them as "safe" or "risky" at time $t$. The CoLGNN framework then applies graph diffusion convolution to capture risk dynamics across the network. The final stage estimates CLN risk scores for each bank, reflecting the network spillover risk associated with each bank at time $t$.

**[Insert Figure 1 about here]**

#### 3.2.2 CoLGNN Framework

To model risk spillover in the co-lending network, we propose the **Co**-**L**ending **G**raph **N**eural **N**etwork framework (CoLGNN), which includes a directional diffusion convolution with edge consideration. To capture the diffusion process of risk across the co-lending network, we utilize the graph diffusion convolution operation first introduced by Li et al. (2017). The directional diffusion convolution over the $l$-th layer graph embedding $\mathbf{H}^{(l)}$ is defined as:

$$\mathbf{H}^{(l)} = \sum_{k=0}^{K-1} \left( \widetilde{\mathbf{A}}_{out}^{k} \mathbf{H}^{(l-1)} \mathbf{W}_{k,1} + \widetilde{\mathbf{A}}_{in}^{k} \mathbf{H}^{(l-1)} \mathbf{W}_{k,2} \right), \tag{2}$$

where $K$ represents the diffusion step, $\widetilde{\mathbf{A}}_{out} = \mathbf{D}_{out}^{-1}\mathbf{A}$ and $\widetilde{\mathbf{A}}_{in} = \mathbf{D}_{in}^{-1}\mathbf{A}^{\top}$, where $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{D}_{out/in} \in \mathbb{R}^{n\times n}$ are graph adjacency matrix and out/in degree matrix respectively, and

$\mathbf{W}_{k,out/in}$ are trainable weights for out/in flow components. In particular, the $\widetilde{\mathbf{A}}_{out/in}$ are dual state transition matrices, particularly in the context of a random walk on the graph. This matrix represents the probabilities of risk transmission moving out/in from one node to another in a one-step random walk.[9]

Edge (loan) attributes in co-lending networks carry significant implications for banks' risk dynamics.[10] Therefore, we incorporate both numerical and categorical loan features as edge attributions $\mathbf{e}_{i,t}$. Categorical attributes are represented through one-hot vector encoding. For multiple co-lending relationships between two nodes within a co-lending network, we employ loan amount weighted averages to consolidate edge attributes between nodes and modify the graph diffusion propagation scheme of Equation (2) by merged edge attributions.

Consider node $i$ on the graph, with $\mathcal{N}_{out}(i)$ and $\mathcal{N}_{in}(i)$ denoting the out and in direction neighbors of node $i$, respectively. From the node features $\mathbf{H}^{(l-1)}$ on the $(l-1)$-layer, we define the following node-wise edge attribution merging scheme:

$$\begin{aligned} \tilde{\mathbf{h}}_{out,i}^{(l-1)} &= \mathbf{h}_i^{(l-1)} + \sum_{j \in \mathcal{N}_{out}(i)} \mathrm{ReLU}(\mathbf{h}_j^{(l-1)} + \mathbf{e}_{i,j}\mathbf{W}); \\ \tilde{\mathbf{h}}_{in,i}^{(l-1)} &= \mathbf{h}_i^{(l-1)} + \sum_{r \in \mathcal{N}_{in}(i)} \mathrm{ReLU}(\mathbf{h}_r^{(l-1)} + \mathbf{e}_{r,i}\mathbf{W}). \end{aligned} \tag{3}$$

In this scheme, we merge two directions' edge attributions into the node features from layer $(l-1)$ to become two directions' edge augmented node features $\tilde{\mathbf{h}}_{i,out}$ and $\tilde{\mathbf{h}}_{i,in}$. We use the rectified linear unit (ReLU) as the activation function to introduce the non-linearity and sparsity. We leverage edge-augmented features to give diffusion graph convolution the capability for capturing edge information. To align with the message-passing paradigm, the new edge-augmented graph diffusion convolution can be expressed by the node-wise

[9]Under the directional convolution framework, the risk spread-over process is considered to occur in two different transition states, reflecting both the leading and participating roles of banks in the syndicated loan market. These two transition matrices are designed to capture distinct patterns.

[10]Incorporating loan features as edge attributes is supported by existing research on lending, which identifies various loan characteristics related to bank risks, such as the syndicated structure (Ivashina, 2009), loan concentration (Gao et al., 2023), and loan covenants (Demiroglu & James, 2010).

representation as:

$$\mathbf{h}_i^{(l)} = \sum_{k=0}^{K-1} \left( \sum_i \tilde{\mathbf{a}}_{out,i}^k \, \tilde{\mathbf{h}}_{out,i}^{(l-1)} \mathbf{W}_{k,out} + \sum_i \tilde{\mathbf{a}}_{in,i}^k \, \tilde{\mathbf{h}}_{in,i}^{(l-1)} \mathbf{W}_{k,in} \right), \tag{4}$$

where $\tilde{\mathbf{a}}_{out/in,i} = 1/d_{out/in,i}$ are out/in direction transition probability, with $d_{out/in,i} = \mathbf{D}_{out/in,ii}$. In both schemes (3) and (4), the weight matrices $\mathbf{W}$, $\mathbf{W}_{k,in}$ and $\mathbf{W}_{k,in}$ are layer dependent. For simplicity, we omit the layer index $l$ in the notation.

**[Insert Figure 2 about here]**

Figure 2 visually displays the CoLGNN framework, which includes a graph diffusion convolution module with edge embeddings. At each graph, we prepare high-dimensional node features and edge features. We utilize the message-passing paradigm and calculate the node-wise representation of each layer of CoLGNN. We use the softmax function to introduce non-linearity into the output of a neuron. Finally, the parameter is estimated by minimizing the binary cross-entropy (BCE) loss function.

# 4 Data and Empirical Deisgn

## 4.1 Measuring Bank Co-Lending Network Risk

To construct the bank-level co-lending network risk score (CLN score), we apply the CoLGNN framework detailed in Section 3.2.2 to a comprehensive sample of bank characteristics from Form FR Y-9C and bank loans from Dealscan. Specifically, at each quarter end, we construct an attributed graph $G_t$, where nodes represent bank holding companies (BHC, or bank hereafter) and edges indicate co-lending relationships. This process yields a dynamic series of co-lending networks, $\mathcal{G} = \{G_1, ..., G_T\}$, spanning from 1991Q1 to 2020Q4, using a five-year rolling window.[11] As indicated in Section 3.1, each directed edge points from a lead

[11]The results are robust to using a three-year rolling window alternative.

bank to another bank within the syndicate structure. For multiple co-lending relationships pointing from one bank to another bank within the window, we aggregate them into a single edge. For each co-lending network $G_t$, we use the BHC characteristics at $t-1$ as the node feature and the lending activities and loan characteristics from $t-1$ to $t-20$ as the edge features. Appendix IA.1 provides detailed information on how we pre-process the graph data.

In our semi-supervised learning framework, we leverage quarterly stock returns to assign "safe" or "risky" labels to banks based on their recent stock market performance. Stock returns are an ideal metric for this purpose, as stock prices are forward-looking and incorporate all publicly available information about a firm's financial health, market outlook, and risk exposure under the assumption of market efficiency. The top 25% of public banks by stock returns are categorized as "safe" (label $Y_{i,t} = 0$), while the bottom 25% are classified as "risky" (label $Y_{i,t} = 1$). The remaining 50% of public banks, along with all private banks, are treated as unlabeled, facilitating out-of-sample analysis.

We frame the prediction task to a node classification task, using the probability of the "risky" class as the co-lending network risk score. Utilizing the parameters derived from the labeled subset, we predict the co-lending risk score (i.e., the probability of the "risky" class) for all unlabeled banks. At the same time, we reassess the co-lending network risk scores for labeled banks using these generated parameters to account for dynamic changes in risk influenced by spillover effects within the network. We measure the bank-level CLN score using the output from the last layer of CoLGNN. For bank $i$ in a given co-lending network $G_t$, the CLN score is estimated as:

$$CLN\ score_i = \sigma\left(\sum_{k=0}^{K-1}\left(\sum_i \tilde{\mathbf{a}}^k_{out,i}\,\tilde{\mathbf{h}}^{(L)}_{out,i}\mathbf{W}_{k,out} + \sum_i \tilde{\mathbf{a}}^k_{in,i}\,\tilde{\mathbf{h}}^{(L)}_{in,i}\mathbf{W}_{k,in}\right)\right) \tag{5}$$

where $\sigma(\cdot)$ represents the sigmoid activation function. The output for each graph $G_t$ is a vector representing the estimated probabilities of the "risky" class for all banks in the co-lending network. At each year-quarter $t$, we construct a co-lending network $G_t$, resulting in an estimated *CLN score*$_{i,t}$ for bank $i$ at year-quarter $t$.

## 4.2 Sample Construction

Our cross-sectional analysis of bank-level co-lending network risk employs a sample of U.S. bank holding companies from 1991Q1 to 2020Q4. We collect syndicated loan data from Refinitiv LoanConnector DealScan, stock returns from Center for Research in Security Prices (CRSP), and bank characteristics from Form FR Y-9C. We combine bank characteristics and stock return data using the CRSP-FRB link table supplied by the Federal Reserve Bank of New York. We match data on syndicated loans from DealScan with bank characteristics based on the parent company of the lenders, using hand-matched bank name concordance files combined at the BHC level. Following (Ivashina, 2009), we identify the lead bank in the DealScan database if the bank's 'Primary Role' is listed as one of 'Arranges', 'Co-arranger', 'Co-lead arranger', 'Lead arranger', 'Mandated Lead arranger', 'Mandated arranger', 'Lead manager' or if the lender's name is listed in the 'Lead Arranger' column.

**[Insert Table 1 about here]**

Table 1 reports the summary statistics of loan-level and bank-level characteristics. Definitions of variables and data sources are provided in Table A1 in the Appendix. All continuous variables are winsorized at the 1$^{st}$ and 99$^{th}$ percentiles for each year-quarter. Our key variable, the bank co-lending network risk score (*CLN score*), ranges from 0 to 1 by construction, with a value of 0 indicating the safest bank and 1 indicating the riskiest bank based on co-lending network risk. Table IA.4.1 lists the top seven banks with the highest and lowest *CLN* scores. Our primary dependent variable, bank loan loss provisions, is measured as the ratio of loan loss provisions (BHCK4230) to total assets (BHCK2170), expressed in percentage points. The mean (median) loan loss provisions is 0.362 (0.163). The average (median) natural logarithm of the total asset size of banks in our sample is 16.164 (15.935). The average (median) return on assets is 0.540 (0.525). The average (median) loan size (total loans to total assets) is 0.628 (0.669). These bank characteristics align with prior literature (e.g. Ellul & Yerramilli, 2013; Dou et al., 2018; Gao et al., 2023). For example, Dou et al. (2018) report average (median) loan loss provisions of 0.335 (0.177). Ellul and Yerramilli (2013) report an average (median) bank size of 16.631 (16.66) and loan size of 0.626 (0.672).

# 5 The Risk Predictability of Bank CLN Score

## 5.1 Predicting future bank-specific risks

In this section, we examine whether the bank-level co-lending network risk score (*CLN score*) has predictive power for future bank risks. Specifically, we use loan loss provisions, scaled by total assets, as the dependent variable in our baseline analysis. Loan loss provisions represent the reserves banks allocate to account for potential uncollectable loans, serving as a proxy for expected loan losses. These provisions not only affect the income statement through expense recognition but also act as a critical indicator of banks' expectations regarding loan performance. To empirically assess whether the *CLN score* predicts future bank risk, we estimate the following $h$-quarter-ahead predictive regression:

$$LLP_{i,t+h} = \beta_1 CLN\ score_{i,t} + \gamma X_{i,t} + \text{Fixed Effects} + \varepsilon_{i,t+h} \tag{6}$$

where $LLP_{i,t+h}$ is the loan loss provisions for bank $i$ at time $t+h$, $CLN\ score_{i,t}$ refers to the bank-level co-lending network risk score (Equation (5)) generated by our CoLGNN model (detailed in Section 3.2.2) of bank $i$ at time $t$, $X_{i,t}$ is the vector of bank characteristics at time $t$ and we control for year-quarter fixed effects to capture the unobservable heterogeneity at each year-quarter level.[12] For bank-level characteristics, we control for bank size, equity capital ratio, deposits, return on assets, loan portfolio size, growth rate of loan portfolio, loan loss allowance, and liquidity ratio. To maintain consistency between the samples of public

[12] We do not control for bank fixed effects since our baseline results focus the cross-sectional risk predictive power of bank co-lending network risk score. However, our results remain robust after including the bank fixed effects, as shown in Table IA.4.3 in the Appendix.

and private banks, we exclude market-based control variables due to the limited availability of such data for private banks.[13] Standard errors are clustered at the bank level. Table 2 reports the results.

**[Insert Table 2 about here]**

We present the results of predictive regressions ranging from 1-quarter-ahead to 8-quarter-ahead horizons. The results show that banks with higher co-lending network risk scores (*CLN score*) tend to have higher loan loss provisions in the future. Specifically, in Panel A of Table 2, which includes all sample banks, the coefficient estimates of $\textit{CLN score}_{i,t}$ are consistently positive across all horizons and statistically significant at the 1% level for the first 7 quarters. For instance, column (1) shows that a one-standard-deviation increase in $\textit{CLN score}_{i,t}$ corresponds to a 0.047-percentage-point rise (i.e., 0.444*0.105) in scaled quarterly loan loss provisions in the subsequent quarter. Economically, considering an average natural logarithm of bank size of 16.164 (in thousands), this translates to an expected annual loan loss of about $20 million. These findings strongly support Hypothesis 1 that higher co-lending network risk of a bank predicts greater future bank risks.

Panels B and C of Table 2, focusing on unlabeled and private banks respectively, validate the $\textit{CLN score}_{i,t}$'s out-of-sample prediction efficacy.[14] The coefficient estimates of $\textit{CLN score}_{i,t}$ remain positive and mostly significant at the 1% level. Notably, consistent with Hypothesis 2, the results in Panel C of Table 2 show that $\textit{CLN score}_{i,t}$ exhibits robust predictive ability for private banks' future risks. This underscores the effectiveness of our semi-supervised GNN model in capturing risk transmission within the co-lending network, providing a solid foundation for risk assessment among private banks based on the topological structure of the co-lending network and the influence of neighboring banks' risk profiles. For instance, column (1) in Panel C reveals that a one-standard-deviation increase

[13] In untabulated results, we additionally include the market-to-book equity ratio and quarterly buy-and-hold stock returns for public banks. Our results remain qualitatively unchanged.

[14] Panel B includes the remaining 50% unlabeled public banks and all private banks. Panel C includes only private banks.

in $CLN\ score_{i,t}$ increases the scaled quarterly loan loss provisions by 0.045-percentage-point (i.e., 0.444*0.101), which is around 12% of the sample mean loan loss provisions of 0.362. Economically, it indicates a sizeable expected loan loss of about $19 million per year.

## 5.2 Robustness checks

We conduct several robustness checks. First, recognizing that co-lending network formation and banks' risk profiles may be influenced by their reputation and lending specialization, we include controls for these factors. Specifically, a bank's reputation is measured by its market share of lead-arranged syndicated loans, while lending specialization is assessed based on the focus of its lending activities within the syndicated loan market and its concentration in specific industries. Table IA.4.2 shows that even after accounting for these factors, the coefficient estimates of $CLN\ score_{i,t}$ remain both statistically and economically significant.

Second, we conduct horse-racing tests to compare the in-sample and out-of-sample predictive performance of the CLN score against the bank Z-score, following the empirical approach of Barwick et al. (2023). The accounting-based bank Z-score is one of the few risk metrics available for banks regardless of stock market information availability and has been widely used to evaluate bank risk in recent literature (e.g. Laeven & Levine, 2009; Houston et al., 2010).[15]

Specifically, we perform the horse-racing tests on four-quarters-ahead predictions using three samples: a) labeled banks only, b) unlabeled banks, c) private banks. We consider four model specifications: 1) banks-level controls only, 2) bank Z-score and control variables, 3) bank CLN score and control variables, and 4) all variables combined. Table 3 reports the results.[16] From columns (1) to (4) in Table 3, we observe that the CLN score improves the in-sample adjusted $R^2$ from 0.515 to 0.520, while the inclusion of the bank Z-score does not

[15]Following Laeven and Levine (2009), we construct the bank Z-score as Bank Z-score $= \frac{(ROA+CAR)}{\sigma(ROA)}$, where $ROA$ is the return on assets, $CAR$ is the ratio of equity to assets, and $\sigma(ROA)$ is the standard deviation of $ROA$.

[16]In untabulated results, we also examine the predictive regressions of bank Z-score and CLN score on all other predictive horizons. The results remain qualitatively unchanged.

result in a significant change in the adjusted $R^2$.

Using the coefficient estimates from Panel A of Table 3, we evaluate the out-of-sample predictions among unlabeled banks and private banks, and report the results in Panel B and C of Table 3. We pool all prediction errors together to compute the root-mean-square-error (RMSE). The results show that the inclusion of the CLN score reduces the RMSE from 0.681 (controls only) to 0.669 (CLN score and controls) for all unlabeled banks and reduces the RMSE from 0.715 (controls only) to 0.703 (CLN score and controls) for private banks. In contrast, the inclusion of the bank Z-score increases the RMSE, indicating its limited contribution to predictive accuracy. These findings validate the CLN score's superior ability to forecast future bank risks and enhance prediction accuracy. They also illustrate the complexity-accuracy tradeoff, where adding additional variables, such as the Z-score, can introduce noise and undermine model performance. The CLN score's predictive strength likely stems from its integration of both accounting information and the topological dynamics of banks' co-lending network, as well as the risk profiles of neighboring banks.

**[Insert Table 3 about here]**

Lastly, we include bank-fixed effects to account for unobserved, time-invariant heterogeneity across banks. Table IA.4.3 in the Appendix confirms that the predictive power of $\textit{CLN score}_{i,t}$ remains robust and statistically highly significant even after controlling for bank-fixed effects. Furthermore, the economic magnitudes of the coefficient estimates are comparable to those in the baseline results presented in Table 2.

## 5.3 Comparison with traditional network centrality

This section evaluates our proposed co-lending network (CLN) risk score against traditional network centrality measures in predicting future bank risks. Metrics such as eigenvector centrality, closeness centrality, and degree centrality are widely utilized in financial network analysis (e.g., Battiston et al., 2012; El-Khatib et al., 2015; Rossi et al., 2018) and primarily measure a node's structural importance within a network. However, these metrics fall short in capturing dynamic risk spillovers from neighboring banks, a critical determinant of future bank risks. For instance, a highly central bank in the co-lending network may be more stable due to diversification or more vulnerable due to heightened exposure, depending on the characteristics of its connections. Likewise, a peripheral bank might seem insulated from systemic shocks but could be critically dependent on a few key relationships. Traditional centrality measures overlook the risk profiles of neighboring nodes, which are essential in understanding risk propagation within co-lending networks. By contrast, the CLN score integrates both the network's structural topology and the risk characteristics of connected banks, offering a more comprehensive and nuanced approach to assessing future bank risks.

We compare the predictive power of the CLN score against various centrality measures, including both in-degree and out-degree versions of these measures. Traditional centrality measures are typically designed for undirected networks. To account for the centrality of a bank from both the lead arranger and participant perspectives, we estimate both in-degree and out-degree versions of network centrality. Section IA.3 provides detailed discussions and definitions of all network centrality measures.

Figure 4 displays the coefficient estimates of the CLN score alongside these centrality measures in predicting bank loan loss provisions four quarters ahead.[17] The figure shows that the coefficient estimates of the CLN score remain statistically significant across all specifications and samples. In contrast, traditional network centrality measures fail to predict future bank risks. This limitation likely arises from their inability to differentiate the nature of connections—whether a bank is predominantly linked to risky or safe counterparties—or to account for the specific attributes of lending relationships.

**[Insert Figure 4 about here]**

[17]The results remain consistent across different predictive horizons.

## 5.4 Economic mechanism

We now move on to investigate the economic mechanisms underlying the bank co-lending network and subsequent bank risks. As discussed in Hypothesis 3, we expect the predictive power of bank $\textit{CLN score}_{i,t}$ to be stronger for banks that are inherently more vulnerable in the co-lending network—specifically, those with smaller sizes, negative performance, higher return volatility, lower capital adequacy, and greater complexity. Such banks are theoretically less equipped to internalize negative shocks. Risk management processes would be more challenging when facing risk propagation in the network. Consequently, these banks' risk profiles are likely more sensitive to the risk dynamics of their neighboring banks in the co-lending network.

### 5.4.1 Bank size

Smaller banks are often found to be more sensitive to risk spillover due to their relatively limited resources and narrower diversification in lending portfolios. This vulnerability is accentuated in co-lending networks, where the interconnectedness with other financial institutions can amplify the transmission of risk. Recent literature such as (Giannetti & Saidi, 2019) shows that smaller banks with less market share would be less likely to internalize negative spillovers. Specifically, smaller banks may lack the capital buffers and risk management sophistication to effectively mitigate the impact of negative shocks originating from their co-lending partners. Their risk profiles would be more sensitive to the risk profiles of surrounding banks in the co-lending network, resulting in a higher predictive power of $\textit{CLN score}_{i,t}$. We construct the indicator variable $\textit{Small bank dummy}_{i,t}$ to represent the group of small banks at each quarter. To investigate the predictive power of bank CLN score, we estimate the regression specification similar to Equation 6 by adding the $\textit{Small bank dummy}_{i,t}$ for each bank $i$ at time $t$ and its interaction term with $\textit{CLN score}_{i,t}$.

**[Insert Table 4 about here]**

We report the results of all samples in Panel A of Table 4, we find that the coefficient estimates of interaction terms between *CLN score*$_{i,t}$ and *Small bank dummy*$_{i,t}$ are all positive and mostly statistically significant indicating that the predictive power of bank CLN score is stronger for small banks. Similarly, in Panels B and C of Table 4, focusing on unlabeled and private banks respectively, respectively, we observe a consistent pattern, although results for private banks show slight variations, likely due to their generally smaller size compared to public banks. The findings are consistent with Hypothesis 3 that smaller banks' risk profiles are intricately tied to the risk profiles of their neighboring banks and also underscore the value of bank *CLN score*$_{i,t}$ in capturing such complex interdependencies. Furthermore, we show that our results are robust if we alternatively use the decile rank of bank size in Table IA.4.7 and use non-performing loans as the dependent variable in Table IA.4.8. The consistent predictive strength of the *CLN score*$_{i,t}$ across various specifications and samples highlights its utility in capturing the intricate web of risk interdependencies, especially for smaller banks more vulnerable to network-induced risk spillovers.

### 5.4.2 Performance shocks and earnings volatility

Next, we delve into the economic mechanism that banks experiencing poor performance or exhibiting volatile returns demonstrate higher sensitivity to risk transmission within their co-lending networks. We conjecture that banks' operational challenges and financial instability potentially magnify their vulnerability to network-induced risks. First, instead of using analyst earnings forecasts, we adopt accounting-based metrics to quantify performance shocks due to the data limitation of private banks. Specifically, we calculate unexpected earnings using the year-over-year change in quarterly earnings per share (EPS), with the unexpected earnings standardized (SUE) by dividing by its standard deviation.[18] We introduce a dummy variable, *Negative earnings shock*$_{i,t}$, set to 1 for bank $i$ at time $t$ if the SUE is negative, indicating a performance shock. We estimate the regression specification

[18]We define the unexpected earnings (UE) for bank $i$ at time $t$ by $UE_{i,t} = \mathrm{EPS}_{i,t} - \mathrm{EPS}_{i,t-4}$ and the standardized unexpected earnings (SUE) by $SUE_{i,t} = \frac{UE_{i,t}}{\sigma(UE_{i,t})}$.

similar to Equation 6 by adding the *Negative earnings shock*$_{i,t}$ for each bank $i$ at time $t$ and its interaction term with *CLN score*$_{i,t}$ to assess the impact of performance shocks on the predictive power of the CLN score.

**[Insert Table 5 about here]**

Table 5 presents the results. We find that the coefficient estimates of *CLN score*$_{i,t}$ remain positive and mostly significant. More importantly, the interaction term between performance shock and bank CLN score are positive and mostly statistically significant in the first five quarters. Conditional on banks' earning performance, the CLN score generates higher predictive power for banks with poorer performance. The results suggest that banks experiencing negative earnings shocks are more sensitive to the risk dynamics of neighboring banks in the network.

Second, we estimate the accounting-based earnings volatility by the ROA volatility, which is applicable for all banks regardless of stock market information availability. ROA volatility is calculated as the standard deviation of quarterly ROA over the preceding five-year period for each bank. We introduce a dummy variable, *High ROA volatility*, to identify banks experiencing significant earnings fluctuations at any given quarter. We hypothesize that the *CLN score*$_{i,t}$'s prediction power intensifies for banks marked by *High ROA volatility*, since banks with low-level financial stability are more vulnerable to negative spillovers within the co-lending network and may depend more heavily on the performance of their co-lending partners.

**[Insert Table 6 about here]**

Table 6 reports the results. We find that the coefficient estimates of interaction terms are positive across all prediction horizons and statistically significant in some future horizons. The results imply that CLN score has more predictive power for banks with volatile returns. Alternatively, we directly examine continuous variables for performance shock (*Standardized earnings change*$_{i,t}$) and earnings volatility (*ROA volatility*$_{i,t}$). We find even stronger results in Table IA.4.9 in the Appendix. The coefficient estimates of interaction terms between *Standardized earnings change*$_{i,t}$ and *CLN score*$_{i,t}$ are significantly negative across most prediction horizons, indicating that the prediction power of CLN score is stronger for banks facing earnings declines. The coefficient estimates of the interaction term between *ROA volatility*$_{i,t}$ and *CLN score*$_{i,t}$ are significantly positive, showing that CLN has greater predictive power for banks with higher earning volatility. Particularly, Panel B of Table IA.4.9 reveals that interaction terms exhibit the highest statistical significance within the subset of unlabeled banks. These institutions, often lacking extensive stock market information or entirely absent from stock exchanges, rely more heavily on accounting-based performance indicators. Such measures signal a pronounced dependency of these banks' financial health on their positioning within the co-lending network, underscoring the nuanced utility of the *CLN score*$_{i,t}$ in capturing risk exposure derived from network interdependencies.

#### 5.4.3 Bank complexity and opacity

Furthermore, we investigate whether the risk management complexity might amplify banks' sensitivity to their neighbors' performance in the co-lending network. Following (Gao et al., 2023), we measure the bank complexity using the ratio of non-missing to total BHCK items (variables with the prefix BHCK for bank holding companies) reported in the FR Y-9C filings. This approach is similar to the accounting "Disaggregation" measure proposed by Chen et al. (2015), which uses the ratio of non-missing items to the number of total items in Compustat. In the banking context, we argue that the proportion of non-missing items in FR Y-9C reflects more of the complexity of a bank's lending activities and risk management. The risk management arguably becomes more complex if banks have more items to report. The underlying premise is that a higher proportion of filled items signifies a broader spectrum of banking operations, from standard balance-sheet entries to detailed off-balance-sheet activities such as unused loan commitments and credit derivatives. Banks' risk management complexities are closely linked to the diversity of financial products offered and the risk mitigation strategies employed. The FR Y-9C, with its extensive list of 2,374 distinct line items with the prefix BHCK as of 2020, provides a comprehensive framework for capturing such complexity.[19] Each non-missing entry in these filings implies active engagement in a specific banking operation or service, making the ratio an effective proxy for the bank's business complexity.

This complexity, we argue, could lead to an increased reliance on the performance of co-lending partners, as managing the risks associated with a broad range of activities might expose banks to greater spillover effects from the network. Therefore, the predictive power of the $\textit{CLN score}_{i,t}$ is hypothesized to be stronger for these complex banks, as their intricate operations and extensive co-lending relationships make them more sensitive to the dynamics within the network. Following Gao et al. (2023), we orthogonalize the bank complexity measure by regressing the bank complexity on bank size.[20] The final bank complexity measure is the residuals from this regression. To identify the group of complex banks, we introduce the $\textit{High complexity}_{i,t}$ dummy variable, signifying banks within the top quartile of complexity each quarter. This variable, alongside its interaction with the $\textit{CLN score}_{i,t}$, is incorporated into our baseline regression model as outlined in Equation (6).

**[Insert Table 7 about here]**

Table 7 shows that the coefficient estimates of $\textit{CLN score}_{i,t}$ remain consistently positive and statistically significant. We find that the coefficients of interaction terms are predominantly statistically significant in Panels B and C of Table 7, suggesting that the predictive power of bank CLN score is stronger for complex banks with less informative or non-existent stock market performance. The observed variance in significance across different samples can be attributed to the differential information environments surrounding these banks. Unlabeled public banks and private banks in our sample, likely operating with less transparency,

[19] In contrast to the 974 items listed in Compustat as of 2024, FR Y-9C's broader array of entries shows the comprehensive scope and high complexity nature of banking operations.

[20] Larger banks might naturally maintain broader business and have a larger number of non-missing items reported in FR Y-9C.

might rely more on international and neighborhood information to navigate their risk landscapes. Our CLN score captures the nuances of lending topological structure and operational complexities become a more important indicator of risks for these entities.

Furthermore, we validate this argument by additionally examining the predictive power of the CLN score conditional on bank opacity. We employ discretionary loan loss provisions as a proxy for opacity following Jiang et al. (2016).[21] We estimate the predictive regression specification similar to Equation (6) by adding the bank opacity measure and its interaction term with the bank CLN score. Table IA.4.10 presents that the coefficients of interaction terms between $Bank\ Opacity_{i,t}$ and $CLN\ score_{i,t}$ are mostly significantly positive in all three samples, showing that the predictive power of CLN score is higher when banks have high opacity. The results further validate our findings in Table 7 and our arguments that risk management complexity and information opacity increase banks' risk exposure to their neighbors' performance and network-induced risk.

#### 5.4.4 Capital adequacy

Lastly, our analysis extends to the influence of capital adequacy on the predictive power of the bank $CLN\ score_{i,t}$. We employ the capital adequacy ratio (CAR) to gauge a bank's capital in relation to its risk-weighted assets. Theoretically, banks with lower capital adequacy are considered more vulnerable to financial distress because they possess thinner capital buffers to absorb losses. Limited loss-absorption capacity could potentially amplify the bank's sensitivity to risk spillovers within the co-lending network. This vulnerability suggests that the financial health and lending decisions of their network partners might sig-

[21] The bank opacity is determined by the natural logarithm of the absolute value of residuals from a regression of loan loss provisions on changes in non-performing assets, among other factors, incorporating state-quarter fixed effects. The regression specification is: $LLP_{i,t,j} = \alpha_1 \Delta NPA_{i,j,t+1} + \alpha_2 \Delta NPA_{i,j,t} + \alpha_3 \Delta NPA_{i,j,t-1} + \alpha_4 Size_{i,j,t-1} + \alpha_5 \Delta Loan_{i,j,t} + \delta_{i,t} + \epsilon_{i,j,t}$ where $LLP_{i,j,t}$ represents loan loss provisions scaled by lagged total loans for bank $i$ in state $j$ at quarter $t$, $\Delta NPA_{i,j,t}$ denotes the change in non-performing assets for bank $i$ in state $j$ from quarter $t-1$ to $t$, scaled by lagged total loans. $Size_{i,j,t-1}$ is the natural logarithm of the bank's total assets at $t-1$, and $\Delta Loan_{i,j,t}$ captures the change in total loans from $t-1$ to $t$. $\delta_{j,t}$ represents state-quarter fixed effects, capturing regional and temporal variations. This model includes both lead and lag of $\Delta NPA_{i,j,t}$ to reflect the banks' use of forward-looking and historical non-performing asset data in provisioning for loan losses.

nificantly impact these banks, enhancing the *CLN score*$_{i,t}$'s predictive relevance for their risk exposure. We introduce a dummy variable, *Low capital adequacy*$_{i,t}$, identifying banks within the lowest quartile of capital adequacy each quarter. We estimate the similar prediction regressions as Equation (6) by adding the *Low capital adequacy*$_{i,t}$ and its interaction term with *CLN score*$_{i,t}$.

**[Insert Table 8 about here]**

Table 8 presents the findings, indicating that the interaction term is significantly positive, particularly in the short term, within the first three quarters. The temporal impact of *Low capital adequacy*$_{i,t}$ on bank *CLN score*$_{i,t}$ is more pronounced. Several factors might account for this short-term effect. First, due to their limited capital buffers, banks with low capital adequacy are immediately more exposed to any fluctuations in the co-lending network, which are quickly overwhelmed, making short-term predictions particularly relevant. Second, over time, banks may adjust their capital management strategies in response to perceived risks, thereby mitigating some of the initial vulnerabilities. More importantly, banks facing low capital adequacy might be subject to regulatory scrutiny, prompting swift corrective measures. Such adjustments could dilute the initial sensitivity to co-lending network risks over longer horizons.

## 5.5 Further Results

### 5.5.1 Predicting other bank risk metrics

Our early-warning risk prediction results are robust to other bank risk metrics beyond bank loan loss provisions. To address concerns regarding loan loss provisions—which may be influenced by managerial discretion and accounting standards—we utilize non-performing loans (NPLs) as an alternative dependent variable. Unlike loan loss provisions based on predicted future losses based on predictive models and may reflect managerial bias, NPLs represent loans where borrowers are no longer making interest payments or principal repayments, directly signaling the bank's credit risk exposure. NPLs reflect actual loan performance and are less susceptible to managerial discretion. In Panel A of Table 9, the coefficients of $CLN\ score_{i,t}$ are positive at the 1% significance level across all prediction horizons. For example, a one-standard-deviation increase in $CLN\ score_{i,t}$ two years ahead correlates with a 0.238-percentage-point rise in the NPL to total assets ratio—equivalent to 23.4% of the sample mean NPL ratio of 1.019, translating into an estimated annual increase in non-performing loans of approximately \$99.7 million.[22] Panels B and C further illustrate the $CLN\ score_{i,t}$'s out-of-sample predictive strength for unlabeled and private banks, with results remaining qualitatively consistent.

**[Insert Table 9 about here]**

Additionally, we validate the predictive power of bank-level CLN score on well-known and widely-used public bank risk metrics, including the default probability (Merton, 1974), the modified default probabilities (Nagel & Purnanandam, 2020),[23] and the natural logarithm of stock return idiosyncratic volatility (ln(IVOL)). We limit the sample to public banks due to the stock market information availability and employ similar specifications as Equation (6) by replacing the dependent variables with public bank risk metrics. In Table 10, we find the coefficient estimates of bank CLN score are negative and statistically significant at the 1% level across most specifications. The results validate the predictability of bank CLN score on future bank risks.

**[Insert Table 10 about here]**

[22]The coefficient in column (8) of Panel A in Table 9 is 0.238. This leads to an annual increase of non-performing loans by $e^{16.164} \times 0.238\% \times 4 \approx \$99,700$ thousands

[23]The default probability is calculated via a KMV iterative approach. The modified default probability is estimated using the data and code from https://voices.uchicago.edu/stefannagel/code-and-data/

### 5.5.2 Predicting future bank profitability

We further investigate the predictive power of bank CLN score on future bank profitability. Consistent with our views in Hypothesis 1, we argue banks' profitability might be affected by the "homophily" effect within the co-lending network, where banks with similar characteristics might affiliate more closely. This effect implies that lead banks, characterized by high profitability and stringent operational standards, are likely to select co-lenders that mirror these attributes. Second, the directional nature of these lending relationships—from lead arrangers to participant banks—facilitates the transmission of rigorous monitoring and due diligence practices, leading to a positive spillover effect on the participants. Empirically, we estimate the following predictive regression models as in Equation (6) by replacing the bank risk measures with the bank profitability measure, ROA. Table 11 reports the results.

**[Insert Table 11 about here]**

We find that the coefficient estimates of bank $\textit{CLN score}_{i,t}$ are negative and statistically significant at the 1% level in all specifications of panel A and in most specifications of panels B and C. The results show that low bank-level $\textit{CLN score}_{i,t}$ predicts high bank profitability for the future 8 quarters. Economically, column (8) of Panel A shows that a one-standard-deviation decrease in $\textit{CLN score}_{i,t}$ leads to a 0.127-percentage-point higher ROA, for 8 quarters ahead, which is around 23.5% of the sample mean ROA of 0.54%.

### 5.5.3 The predictability after controlling for bank stock performance

We further examine the predictive power of the bank CLN score after controlling for bank stock performance, measured by quarterly stock returns, within the subsample of public banks. In Panel A of Table IA.4.6, we find that controlling for quarterly bank stock returns does not reduce the predictive strength of $\textit{CLN score}_{i,t}$. Its coefficient estimates remain significant across all dependent variables—loan loss provisions, non-performing loans, and ROA—yielding results that are qualitatively consistent with our baseline findings.

Importantly, if the predictive power of the CLN score were due to banks with high CLN values experiencing shared negative shocks (e.g., common major borrower defaults), then controlling for the stock returns of these banks should significantly diminish the CLN score's predictive power. However, our findings suggest otherwise. Specifically, the bank CLN score continues to significantly predict future loan loss provisions, non-performing loans, and ROA after controlling for bank stock returns, consistent with its role in capturing spillover risk through the co-lending network.

Notably, Panel B of Table IA.4.6 further shows that the bank CLN score significantly outperforms quarterly bank stock returns in predicting future non-performing loans and ROA for the subsample of unlabeled public banks. Our CoLGNN model is trained using information from the top and bottom quartiles of stock performance and the co-lending network, which are considered to provide more informative signals than middle-ranked stock returns. Consequently, the results in Panel B indicate that stock returns offer limited predictive value for unlabeled public banks, while our CLN score continues to effectively anticipate future risks. These findings underscore the utility of the CLN score in capturing bank risks beyond conventional stock performance metrics.

### 5.5.4 Time-varying predictability of bank co-lending network risk score

Given that we train the CoLGNN model separately for each co-lending network, we alleviate the concern of different model convergence capabilities by performing subsample analysis. Table IA.4.4 presents the sub-sample analysis in which we equally split the sample into two halves to validate the predictability of early warning of the bank-level CLN score. The $\mathit{CLN\ score}_{i,t}$ exhibits robust and significant predictive power across all samples and all dependent variables.

Furthermore, Table IA.4.5 specifically examines the predictive power of the bank-level CLN score during the Global Financial Crisis (GFC). We introduce a GFC dummy variable which equals 1 if the year-quarter is between mid-2007 and the end of 2008, and 0 otherwise. The stronger statistical and economic significance of the interaction term between $CLN\ score_{i,t}$ and the GFC dummy shows that the predictive power of bank CLN score becomes particularly potent during periods of increased financial instability. During the GFC, the interconnected nature of financial institutions exacerbated the transmission of risk across the banking sector. As banks faced simultaneous liquidity constraints and credit defaults, the risk propagated through their co-lending relationships, making the risk spillover more prominent during GFC period. The result highlights the increased relevance of network-based risk measures during crisis periods, where traditional risk metrics may not capture the full extent of interconnected risks.

# 6 The Risk Transmission of Credit Rating Downgrade

In this section, we validate the bank CLN score and risk transmission mechanism in the co-lending network by a quasi-natural experiment of credit rating downgrade events. Credit ratings, pivotal in assessing the creditworthiness of financial institutions, signal a perceived increase in risk associated with the bank and directly impact the bank's operation capabilities. Institutional investors rely on credit rating to determine their investment in a bank's debt securities and Basel capital requirements for holding such securities on their balance sheets. Bank downgrade, as a first-order concern in a bank's access to funding, could significantly affect a bank's ability to source funding, especially from wholesale funding and public bond markets (Adelino & Ferreira, 2016). Banks facing a downgrade would experience restricted market access, heightened collateral requirements, and rising funding costs. Given the severe consequence, we study the negative shocks from a focal bank's credit rating downgrade propagates through the co-lending network, affecting connected banks' risk profiles and lending behaviors. The rationale is twofold: first, downgraded banks may curtail screening efforts to conserve capital and manage increased funding costs, thereby transmitting financial stress to co-lenders through reduced due diligence and monitoring efforts. Second, the interconnected nature of financial institutions means that the focal banks' risk perceptions and funding conditions can influence the broader cluster of co-lending partners, leading to a reassessment of risk and potentially tighter lending conditions across the surrounding neighborhood.

Following Adelino and Ferreira (2016), we use standard and poor's (S&P's) rating history instead of other credit rating history. Because S&P tends to make rating revisions more actively and to lead other agencies in revisions (Kaminsky & Schmukler, 2002). S&P rating announcements also generate a greater impact on the stock market that is less likely to be fully anticipated by the market (Reisen & Von Maltzan, 1999; Adelino & Ferreira, 2016). We focus on long-term issuer ratings following the convention of the standard credit rating literature (e.g. Xia, 2014; Badoer et al., 2019). Specifically, we investigate the transmission of risk within co-lending networks after a focal bank's credit downgrade. Such downgrades, while indicative of the focal bank's deteriorating financial health, pose an exogenous shock to banks directly connected within the co-lending network. The information asymmetry between these connected banks and credit rating agencies makes the former external to the conditions precipitating the downgrade. Consequently, connected banks encounter an unforeseen shift in their risk landscape, exacerbated by regulatory scrutiny and market reevaluation of the focal bank. Moreover, post-downgrade, the focal bank's operational challenges might reduce its monitoring efforts of existing loan contracts, thereby introducing unforeseen moral hazard concerns for its co-lending partners.

The earlier section shows that bank-level CLN scores provide an early-warning measure for future bank risks. Given the exogenous shock to connected banks in a credit downgrade event, we expect banks directly connected to a downgraded entity (the focal bank) will experience a significant increase in their CLN scores relative to banks without direct connections, post-event. This reflects the risk of transmissions of credit downgrade events in the co-lending network. Therefore, we perform a stacked cohort difference-in-differences (DiD) estimation. Empirically, we consider each credit downgrade event $c$ for a focal bank at time $t$ as an event cohort (*Credit event*$_{c,t}$). We construct a subsample (event cohort) for each credit downgrade event. In our directed co-lending network, as described in Section 3.1, a directed linkage points from one lead bank to other banks within a syndicated loan. Within each event cohort, we classify banks with direct in-wards connections from the focal bank as treated banks.[24] We utilize propensity score matching to select control banks — those without direct ties to the focal bank yet exhibiting similar characteristics to the treated banks. One possible concern is that credit downgrade events might be driven by broader economic downturns or sector-wide risks, which could potentially compromise the parallel trends assumption by predisposing treated banks to heightened systematic risk prior to the downgrade. To ensure the credit downgrade events are indeed exogenous to treated banks, we maintain a clean sample by excluding the GFC period and quarters characterized by an aggregation of downgrade events.[25] More details of how we construct our treatment banks and control banks sample are shown in the Appendix.

For each cohort, we use an event window for each credit downgrade event, spanning from three quarters before to three quarters after the event (excluding the focal downgrade-event quarter). We then stack all the credit downgrade event cohorts together and estimate the following standard DiD regression specification:

$$CLN\ score_{i,c,t} = \beta_1 Treat_{i,c} + \beta_2 Post_{c,t} + \beta_3 Treat_{i,c} \times Post_{c,t} + \gamma_i + \theta_t + \varepsilon_{i,c,t} \tag{7}$$

where $CLN\ score_{i,c,t}$ denotes the co-lending network score of bank $i$ in cohort $c$ at time $t$, capturing the bank's embedded risk within the co-lending network from the CoLGNN model as Equation (5). $Treat_{i,c}$ is a dummy that equals 1 for banks $i$ that are directly connected to a focal bank of the cohort $c$, and 0 otherwise. $Post_{c,t}$ is a dummy that equals 1 for quarters following the credit downgrade event in cohort $c$, and 0 for quarters prior. $\gamma_i$ is bank fixed

[24] We use the bank size, loan size, loan growth and ROA to perform the PSM matching. The control banks are selected based on 1-to-5 nearest neighbour matching.

[25] We remove the credit downgrade events during the GFC period from mid-2007 (2007Q2) through early-2009 (2009Q1) and quarters with more than ten downgrade events to avoid the confounding effects of widespread economic adversities.

effects and $\theta_t$ is year-quarter fixed effects. For robustness, we also include specifications with cohort-bank fixed effects and cohort-year-quarter fixed effects. In this case, $Treat_{i,c}$ and $Post_{c,t}$ are absorbed by cohort-bank fixed effects and cohort-year-quarter fixed effects, respectively. Standard errors are double-clustered at both the cohort and bank levels.

Panel A of Table 12 presents the results. We find that the coefficients estimate of the interaction term, $Treat_{i,c} \times Post_{c,t}$, is positive and significant at the 5% level in column (1) and at the 1% level in column (2), showing that treated banks have a significantly large increase in CLN score than control banks, post-credit downgrade. In columns (3) and (4), we specifically examine the DiD estimation for credit downgrade events below the rating class "A". The coefficient estimates of the DiD term are significantly positive at the 1% level and have large economic significance for downgrades to class "BBB" or below. Consistent with our findings in Section 5.4, this implies that the risk transmission has a larger impact on more vulnerable banks. Our DiD estimations validate that the bank-level CLN score generated by our CoLGNN model could capture the risk spillover in the financial network. In columns (5) and (6), we further include lagged one-quarter bank-level control variables as in Table 2, and the results remain qualitatively unchanged.

[**Insert Table 12 about here**]

Furthermore, we include the downgrade-event quarter in the sample to precisely identify the timing of the treatment effect using a dynamic DiD regression framework. To assess whether the observed treatment effects of credit downgrade risk spillover are influenced by potential nonparallel trends between the treatment and control banks prior to the downgrade event, we estimate dynamic DiD specifications, replacing the $Post_{c,t}$ dummy with quarter-specific dummy variables, using the first quarter of each cohort as the reference quarter.

Panel B of Table 12 shows no pre-existing differential trends in CLN scores between treatment and control banks. The treatment effect manifests only in the periods following the event quarter. The results indicate that treatment banks experience a significantly larger increase in their CLN score compared to control banks two and three quarters after the credit downgrade. These findings demonstrate the risk transmission mechanism within the co-lending network and further confirm that the bank CLN score effectively captures risk spillover in the co-lending network.

# 7 Conclusion

In this paper, we propose a novel directed network design that utilizes the co-lending relationship in syndicated loan markets. Leveraging the high-dimensional bank and loan characteristics and the topological structure of the co-lending network, we develop the Co-Lending Graph Neural Network (CoLGNN) model. CoLGNN not only facilitates our understanding of risk distribution across financial networks but also generates a novel bank-level co-lending network risk score, which predicts future bank risks and performance across both public and private banking sectors.

Our empirical investigations demonstrate the CLN score's robust predictive capability, extending its relevance to private banks—a sector traditionally obscured from public market-based risk assessments. In particular, we identify that the predictive power of bank co-lending network risk score is strong for more vulnerable banks, characterized by smaller size, negative performance, higher return volatility, lower capital adequacy, and greater complexity in risk management in the co-lending network. Moreover, using the S&P long-term credit rating history, we empirically validate the risk spillover mechanism in the co-lending network by studying the transmission of negative shocks. The results show that banks with in-ward connections with focal downgraded banks exhibit a significant increase in their co-lending network score relative to control banks after the downgrade events. The findings on the risk spillover highlight the complex interdependencies in the banking industry, illustrating how the risks associated with one financial institution can propagate through the network and affect others.

Finally, our findings are important for financial institutions and regulators alike, highlighting the necessity for a more comprehensive approach to risk management and monitoring that transcends individual bank assessments to encompass the broader financial ecosystem. For policymakers, our results advocate for the development and implementation of regulatory frameworks that consider the interconnected nature of banking networks, potentially leading to more resilient financial systems capable of withstanding shocks from individual entities. By acknowledging and addressing risk transmission in co-lending networks, regulators can better safeguard against systemic vulnerabilities, ensuring a more stable banking environment. The CoLGNN model, through its detailed accounting of network structures and risk dynamics, heralds a new era of financial analysis—one that equips stakeholders to navigate the complex interdependencies of risk with unprecedented precision and foresight.

# References


Acemoglu, D., Ozdaglar, A., & Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, *105*(2), 564–608.

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring systemic risk. *The review of financial studies*, *30*(1), 2–47.

Adelino, M., & Ferreira, M. A. (2016). Bank ratings and lending supply: Evidence from sovereign downgrades. *The Review of Financial Studies*, *29*(7), 1709–1746.

Allen, F., & Gale, D. (2000). Financial contagion. *Journal of political economy*, *108*(1), 1–33.

Anderson, H., Paddrik, M., & Wang, J. J. (2019). Bank networks and systemic risk: Evidence from the national banking acts. *American Economic Review*, *109*(9), 3125–61.

Badoer, D. C., Demiroglu, C., & James, C. M. (2019). Ratings quality and borrowing choice. *The Journal of Finance*, *74*(5), 2619–2665.

Barwick, P. J., Liu, Y., Patacchini, E., & Wu, Q. (2023). Information, mobile communication, and referral effects. *American Economic Review*, *113*(5), 1170–1207.

Battiston, S., Puliga, M., Kaushik, R., Tasca, P., & Caldarelli, G. (2012). Debtrank: Too central to fail? financial networks, the fed and systemic risk. *Scientific reports*, *2*(1), 541.

Benmelech, E., Dlugosz, J., & Ivashina, V. (2012). Securitization without adverse selection: The case of clos. *Journal of Financial Economics*, *106*(1), 91–113.

Bharath, S. T., Dahiya, S., Saunders, A., & Srinivasan, A. (2011). Lending relationships and loan contract terms. *The Review of Financial Studies*, *24*(4), 1141–1203.

Blickle, K., Fleckenstein, Q., Hillenbrand, S., & Saunders, A. (2020). The myth of the lead arranger's share. *FRB of New York Staff Report*, (922).

Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., & Vandergheynst, P. (2017). Geometric deep learning: Going beyond euclidean data. *IEEE Signal Processing Magazine*, *34*(4), 18–42.

Brunnermeier, M. K., & Oehmke, M. (2013). Bubbles, financial crises, and systemic risk. *Handbook of the Economics of Finance*, *2*, 1221–1288.

Cabrales, A., Gottardi, P., & Vega-Redondo, F. (2017). Risk sharing and contagion in networks. *The Review of Financial Studies*, *30*(9), 3086–3127.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance*, *63*(6), 2899–2939.

Cao, S., Jiang, W., Yang, B., & Zhang, A. L. (2023). How to talk when a machine is listening: Corporate disclosure in the age of ai. *The Review of Financial Studies*, *36*(9), 3603–3642.

Chami, I., Abu-El-Haija, S., Perozzi, B., Ré, C., & Murphy, K. (2022). Machine learning on graphs: A model and comprehensive taxonomy. *The Journal of Machine Learning Research*, *23*(1), 3840–3903.

Chen, S., Miao, B., & Shevlin, T. (2015). A new measure of disclosure quality: The level of disaggregation of accounting data in annual reports. *Journal of Accounting Research*, *53*(5), 1017–1054.

Cong, L. W., Tang, K., Wang, J., & Zhang, Y. (2021). Alphaportfolio: Direct construction through deep reinforcement learning and interpretable ai. *Available at SSRN 3554486*.

Demiroglu, C., & James, C. M. (2010). The information content of bank loan covenants. *The Review of Financial Studies*, *23*(10), 3700–3737.

Dou, Y., Ryan, S. G., & Zou, Y. (2018). The effect of credit competition on banks' loan-loss provisions. *Journal of Financial and Quantitative Analysis*, *53*(3), 1195–1226.

El-Khatib, R., Fogel, K., & Jandik, T. (2015). Ceo network centrality and merger performance. *Journal of Financial Economics*, *116*(2), 349–382.

Elliott, M., Golub, B., & Jackson, M. O. (2014). Financial networks and contagion. *American Economic Review*, *104*(10), 3115–3153.

Ellul, A., & Yerramilli, V. (2013). Stronger risk controls, lower risk: Evidence from us bank holding companies. *The Journal of Finance*, *68*(5), 1757–1803.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417.

Fama, E. F. (1991). Efficient capital markets: Ii. *The Journal of Finance*, *46*(5), 1575–1617.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22.

Gao, M., Hasan, I., Qiu, B., & Wu, E. (2023). Lone (loan) wolf pack risk. *SSRN Working Paper*.

Giannetti, M., & Saidi, F. (2019). Shock propagation and banking structure. *The Review of Financial Studies*, *32*(7), 2499–2540.

Glasserman, P., & Young, H. P. (2015). How likely is contagion in financial networks? *Journal of Banking & Finance*, *50*, 383–399.

Golub, B., & Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, *127*(3), 1287–1338.

Gopalan, R., Nanda, V., & Yerramilli, V. (2011). Does poor performance damage the reputation of financial intermediaries? evidence from the loan syndication market. *The Journal of Finance*, *66*(6), 2083–2120.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223–2273.

Haldane, A. G., & May, R. M. (2011). Systemic risk in banking ecosystems. *Nature*, *469*(7330), 351–355.

Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Representation learning on graphs: Methods and applications. *IEEE Data Eng. Bull.*, *40*(3), 52–74.

Holmstrom, B., & Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, 303–328.

Houston, J. F., Lin, C., Lin, P., & Ma, Y. (2010). Creditor rights, information sharing, and bank risk taking. *Journal of financial Economics*, *96*(3), 485–512.

Ivashina, V. (2009). Asymmetric information effects on loan spreads. *Journal of financial Economics*, *92*(2), 300–319.

Ivashina, V., & Scharfstein, D. (2010a). Bank lending during the financial crisis of 2008. *Journal of Financial economics*, *97*(3), 319–338.

Ivashina, V., & Scharfstein, D. (2010b). Loan syndication and credit cycles. *American Economic Review*, *100*(2), 57–61.

Jiang, J., Kelly, B., & Xiu, D. (2023). (re-) imag (in) ing price trends. *The Journal of Finance*, *78*(6), 3193–3249.

Jiang, L., Levine, R., & Lin, C. (2016). Competition and bank opacity. *The Review of Financial Studies*, *29*(7), 1911–1942.

Kaminsky, G., & Schmukler, S. L. (2002). Emerging market instability: Do sovereign ratings affect country risk and stock returns? *The World Bank Economic Review*, *16*(2), 171–195.

Laeven, L., & Levine, R. (2009). Bank governance, regulation and risk taking. *Journal of financial economics*, *93*(2), 259–275.

Li, K., Mai, F., Shen, R., & Yan, X. (2021). Measuring corporate culture using machine learning. *The Review of Financial Studies*, *34*(7), 3265–3315.

Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2017). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926*.

Lim, J., Minton, B. A., & Weisbach, M. S. (2014). Syndicated loan spreads and the composition of the syndicate. *Journal of Financial Economics*, *111*(1), 45–69.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, *29*(2), 449–470.

Nagel, S., & Purnanandam, A. (2020). Banks' risk dynamics and distance to default. *The Review of Financial Studies*, *33*(6), 2421–2467.

Pichler, P., & Wilhelm, W. (2001). A theory of the syndicate: Form follows function. *The Journal of Finance*, *56*(6), 2237–2264.

Reisen, H., & Von Maltzan, J. (1999). Boom and bust and sovereign ratings. *International Finance*, *2*(2), 273–293.

Richmond, R. J. (2019). Trade network centrality and currency risk premia. *The Journal of Finance*, *74*(3), 1315–1361.

Rossi, A. G., Blake, D., Timmermann, A., Tonks, I., & Wermers, R. (2018). Network centrality and delegated investment performance. *Journal of Financial Economics*, *128*(1), 183–206.

Tobias, A., & Brunnermeier, M. K. (2016). Covar. *The American Economic Review*, *106*(7), 1705.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Philip, S. Y. (2020). A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*.

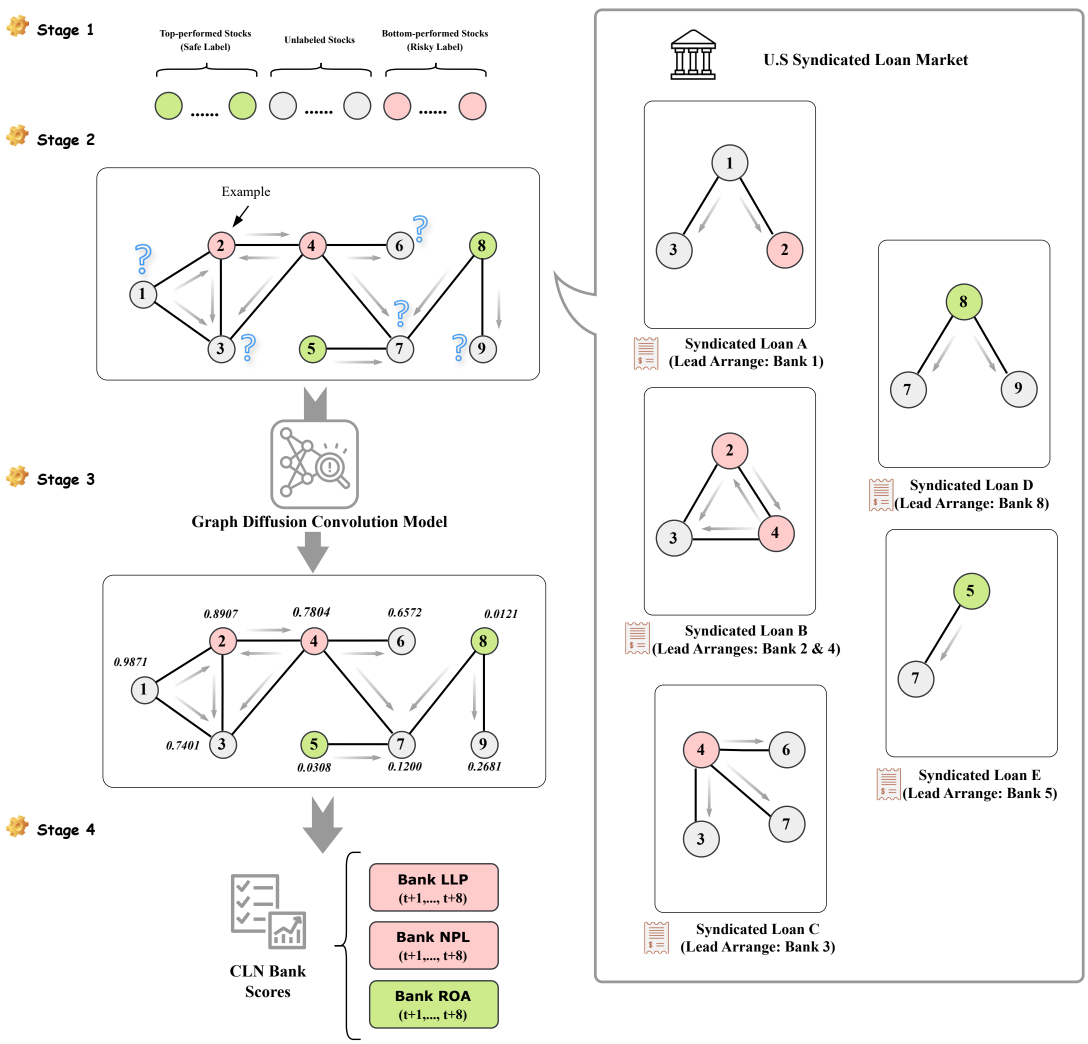
Xia, H. (2014). Can investor-paid credit rating agencies improve the information quality of issuer-paid rating agencies? *Journal of Financial Economics*, *111*(2), 450–468.

Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2019). How powerful are graph neural networks? *International Conference on Learning Representations*.

Yellen, J. (2013). *Interconnectedness and systemic risk: Lessons from the financial crisis and policy implications: A speech at the american economic association/american finance association joint luncheon, san diego, california, january 4, 2013* (tech. rep.). Board of Governors of the Federal Reserve System (US).

Figure 1: **Estimating Bank Risk via Co-Lending Networks and CoLGNN**

Figure 1 shows the flow chart of our estimation of bank co-lending network risk via Co-Lending Networks and CoLGNN. Specifically, the figure illustrates the co-lending network at time $t$. At stage 1, we classify banks with top-performed stock market performance to the “safe” group and classify banks with bottom-performed stock market performance to the “risky” group. The remaining public banks and private banks are treated as unlabeled samples. At stage 2, we construct the co-lending network by utilizing the all syndicated loans originated in the last five years (from $t-20$ to $t-1$). Each co-lending network is a directed network with edges points from each lead arrange to other banks within the syndicate. At stage 3, we estimate bank CLN score CoLGNN model as described in section 3.2.2. At stage 4, we perform empirical experiment showing the Bank CLN score effectively captures risk spillovers in the co-lending network, serving as an early-warning indicator of each bank’s future risk and performance.



## Figure 2: **Graph Diffusion Module**

Figure 2 visually illustrate the CoLGNN framework which includes a graph convolution with edge embedding. At stage 3.1, we prepare the high dimensional node features and edge features using bank characteristics from FR Y9-C and the loan characteristics from DealScan, respectively. At stage 3.2, we aggregate the all features at node level. To capture the risk spillover in the co-lending network, we utilize the message-passing paradigm design in the graph neural network by creating in-flow aggregation and out-flow aggregation. At stage 3.3, we calculate the node-wise representation for each layer of CoLGNN and use the softmax activation function to introduce non-linearity into the output of a neuron. At stage 3.4, the parameter matrix of CoLGNN is estimated by minimizing the binary cross-entropy (BCE) loss function. At stage 3.5, we consider the estimated probabilities to the "risky" class as the final co-lending network risk score for bank $i$ at time $t$.

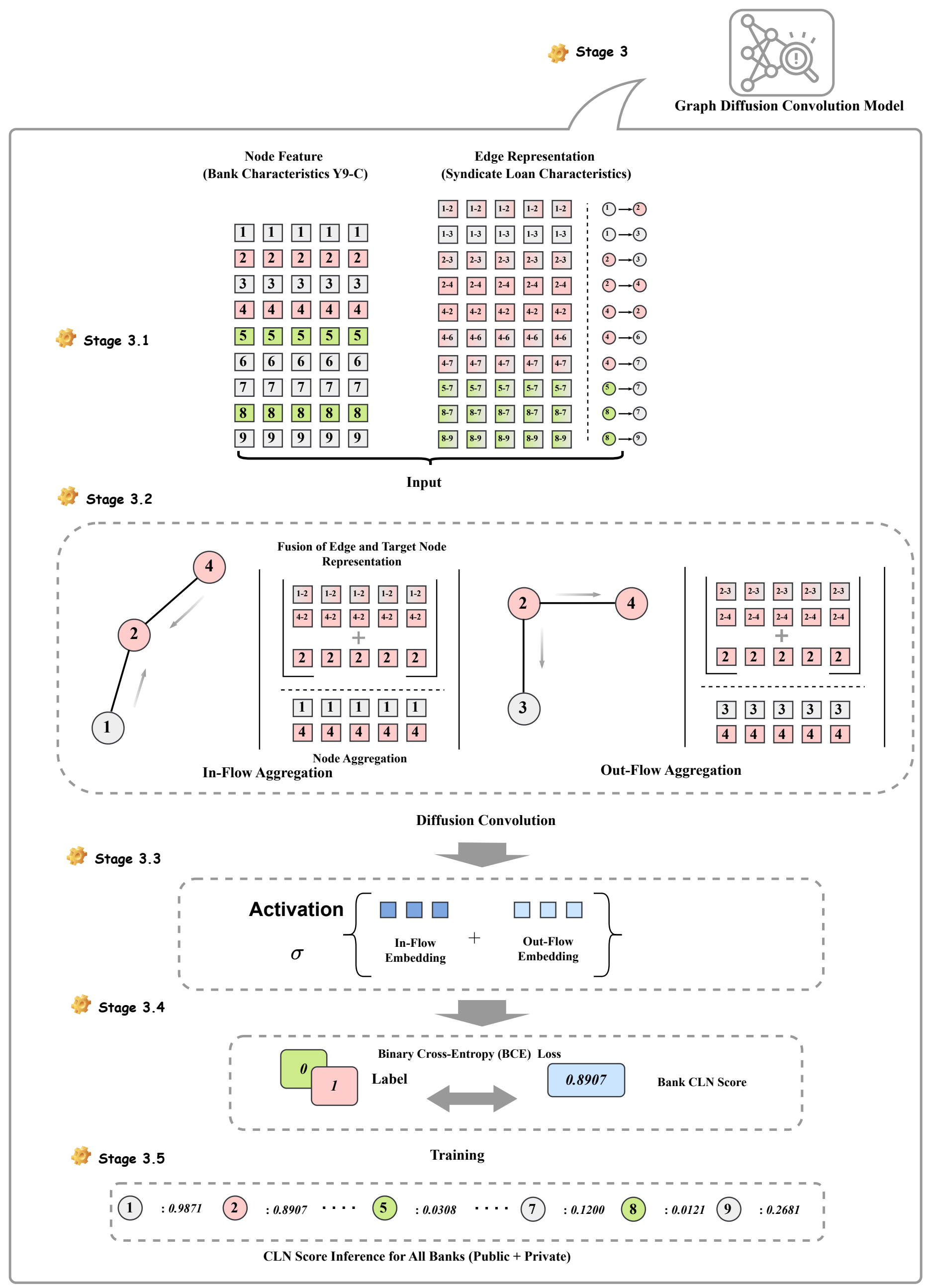


Figure 3: **Dynamic Treatment Effects: Credit Downgrade Spillover**

Figure 3 shows the coefficient estimates of the interaction of the time dummies and the treated dummy in the dynamic difference-in-differences regression in Table 12. The figure shows the 95% confidence interval of the coefficient estimates.

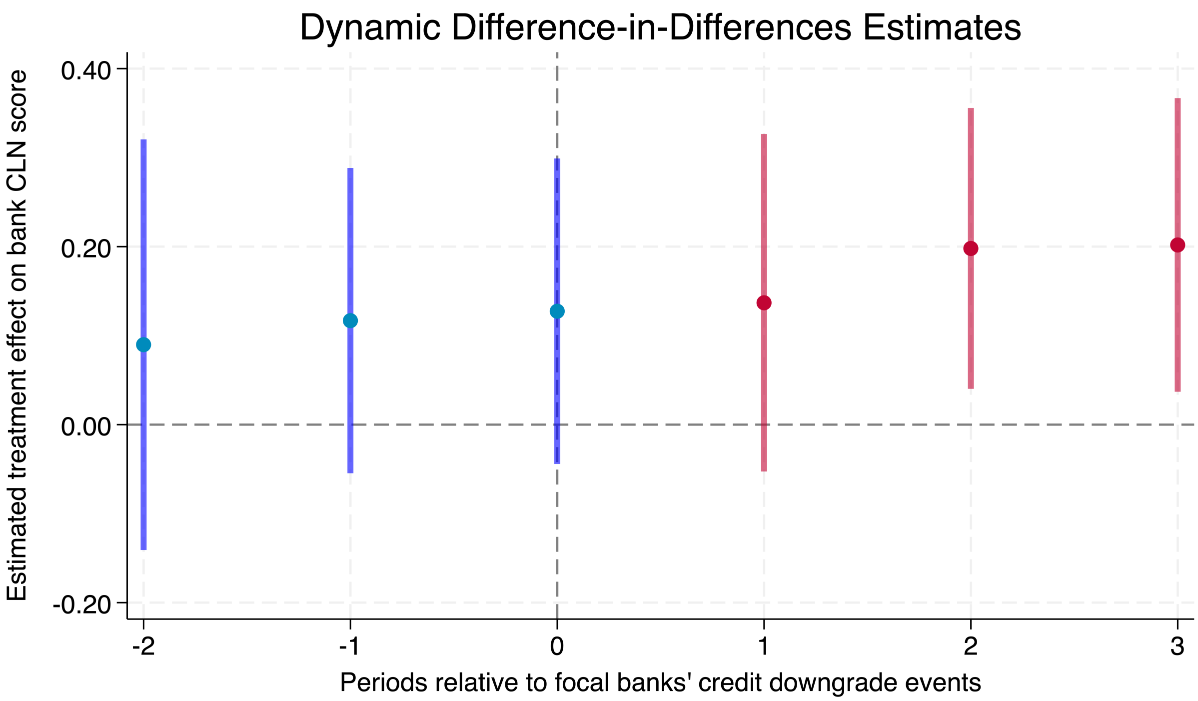
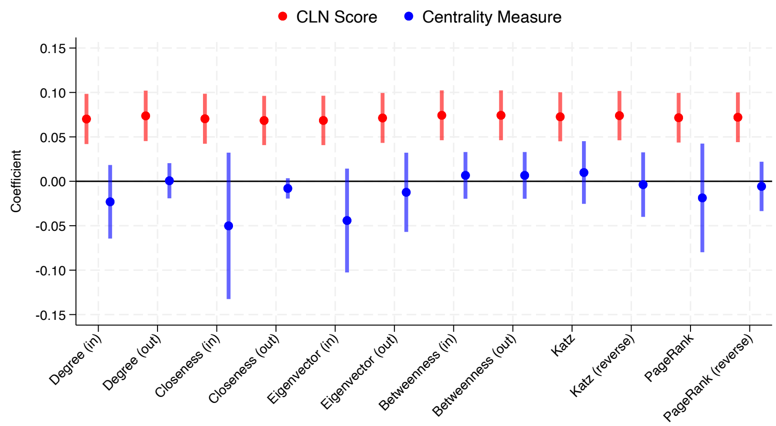
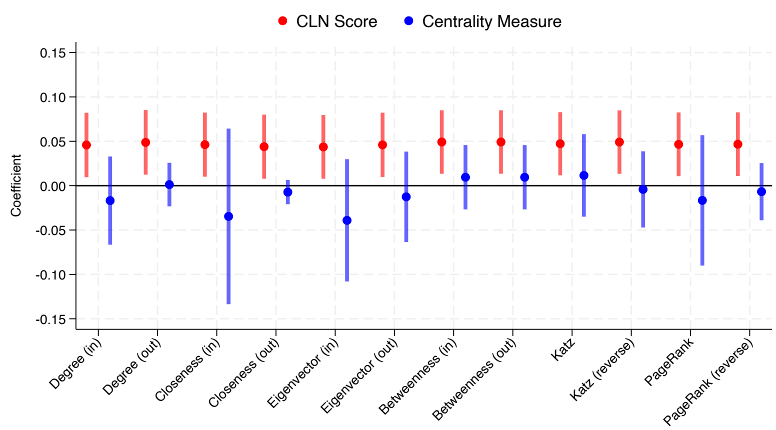


Figure 4: **CLN Score and Centrality Measure Coefficient Estimates for Different Samples**

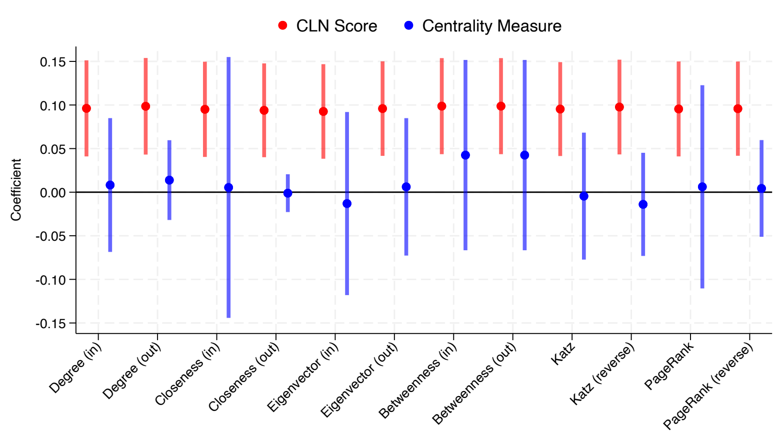
Figure 4 shows the coefficient estimates of the CLN score alongside the centrality measures in predicting bank loan loss provisions four quarters ahead. The centrality measures are defined in Appendix IA.3.




(a) All sample




(b) Out-of-sample (Unlabeled Banks)




(c) Out-of-sample (private banks)

Table 1: **Summary Statistics**

Table 1 presents the summary statistics of our study. The loan-level samples from January 1990 to December 2020. Definitions of the variables are provided in Table A1 in the Appendix. All continuous variables are winsorized by year at the $1^{st}$ and $99^{th}$ percentiles.

| | Observations | Mean | $10^{th}$ Percentile | Median | $90^{th}$ Percentile | Standard Deviation |
|---|---|---|---|---|---|---|
| ***Bank-level Samples*** | | | | | | |
| *CLN* score | 11662 | 0.468 | 0.000 | 0.373 | 1.000 | 0.444 |
| Loan loss provision | 11662 | 0.362 | 0.003 | 0.163 | 0.865 | 0.688 |
| Size | 11662 | 16.164 | 13.854 | 15.935 | 18.892 | 1.984 |
| Equity capital | 11662 | 0.104 | 0.069 | 0.098 | 0.139 | 0.042 |
| ROA | 11662 | 0.540 | 0.109 | 0.525 | 1.193 | 0.827 |
| Loan size | 11662 | 0.628 | 0.415 | 0.669 | 0.784 | 0.161 |
| Loan growth | 11662 | 0.035 | -0.023 | 0.015 | 0.068 | 1.165 |
| Loan loss allowance | 11662 | 0.946 | 0.431 | 0.869 | 1.550 | 0.521 |
| Liquidity | 11662 | 0.199 | 0.080 | 0.179 | 0.350 | 0.112 |
| Deposits | 11662 | 0.549 | 0.293 | 0.579 | 0.743 | 0.181 |
| Reputation | 11662 | 0.005 | 0.000 | 0.000 | 0.010 | 0.019 |
| Specialization in syndicated loan | 11662 | 0.007 | 0.000 | 0.000 | 0.001 | 0.311 |
| Specialization in industry | 11662 | 0.226 | 0.000 | 0.000 | 1.000 | 0.341 |
| Bank Z-score | 11662 | 0.323 | 0.150 | 0.297 | 0.505 | 0.187 |
| Bank nonperforming loans | 10586 | 1.019 | 0.140 | 0.567 | 2.160 | 1.581 |
| Standardized unexpected earnings | 11294 | 0.039 | -0.000 | 0.000 | 0.000 | 2.810 |
| Bank complexity | 11662 | 0.044 | -0.172 | 0.062 | 0.242 | 0.154 |
| Bank risk capital | 8309 | 0.143 | 0.108 | 0.132 | 0.181 | 0.066 |
| Modified default probability | 6252 | 0.209 | 0.089 | 0.174 | 0.369 | 0.134 |
| Default probability | 6252 | 0.131 | 0.018 | 0.065 | 0.316 | 0.174 |
| IVOL | 8190 | 1.594 | 0.762 | 1.245 | 2.649 | 1.224 |
| | | | | | | |
| ***Loan-level Samples*** | | | | | | |
| ln(Loan Size) | 66596 | 4.998 | 3.122 | 5.017 | 6.909 | 1.451 |
| Loan Spread (bps) | 66596 | 227.286 | 50.000 | 200.000 | 425.000 | 154.598 |
| Loan Maturity (months) | 66596 | 53.088 | 12.000 | 60.000 | 84.000 | 30.476 |
| Lead Banks | 66596 | 1.965 | 1.000 | 1.000 | 4.000 | 1.646 |
| Lenders | 66596 | 9.197 | 2.000 | 6.000 | 20.000 | 8.923 |
| General covenants | 66596 | 1.222 | 0.000 | 0.000 | 6.000 | 2.924 |
| Financial covenants | 66596 | 0.413 | 0.000 | 0.000 | 2.000 | 0.965 |
| Loan secured | 66596 | 0.400 | 0 | 0 | 1 | 0.490 |

Table 2: **Bank Co-Lending Network Risk and Bank Loan Loss Provisions**

Table 2 presents the h-quarter-ahead prediction results of bank-level *CLN*-score for bank loan loss provisions. The bank-level *CLN*-score and control variables are measured at time $t$, and the dependent variable is measured at time $t + h$. Panel A reports the results using all observations. Panel B shows the results using unlabeled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables in the Panel B and C. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Loan loss provisions** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All samples*** | | | | | | | | |
| *CLN* score | 0.105*** | 0.106*** | 0.105*** | 0.073*** | 0.092*** | 0.070*** | 0.063*** | 0.032* |
| | (7.958) | (8.197) | (7.637) | (5.186) | (5.700) | (4.260) | (3.769) | (1.689) |
| Size | 0.034*** | 0.035*** | 0.035*** | 0.039*** | 0.039*** | 0.039*** | 0.041*** | 0.045*** |
| | (3.628) | (3.665) | (3.571) | (3.756) | (3.623) | (3.501) | (3.639) | (3.819) |
| Equity capital | -0.671 | -0.534 | -0.408 | 0.148 | 0.242 | 0.270 | 0.388 | 0.625 |
| | (-0.698) | (-0.718) | (-0.610) | (0.278) | (0.488) | (0.539) | (0.769) | (1.234) |
| Deposits | -0.017 | -0.029 | -0.063 | -0.079 | -0.102 | -0.133 | -0.122 | -0.116 |
| | (-0.145) | (-0.248) | (-0.534) | (-0.633) | (-0.780) | (-0.998) | (-0.918) | (-0.843) |
| ROA | -0.130*** | -0.090*** | -0.068*** | -0.106*** | -0.057** | -0.051* | -0.056 | -0.073* |
| | (-4.288) | (-3.273) | (-2.684) | (-2.821) | (-2.248) | (-1.682) | (-1.482) | (-1.694) |
| Loan size | -0.492*** | -0.399*** | -0.298** | -0.161 | -0.076 | -0.011 | 0.041 | 0.114 |
| | (-3.553) | (-2.942) | (-2.101) | (-1.083) | (-0.505) | (-0.072) | (0.256) | (0.710) |
| Loan growth | 0.000 | 0.001 | -0.030 | -0.071 | -0.012 | -0.000 | 0.130* | 0.235** |
| | (0.059) | (1.365) | (-0.440) | (-1.216) | (-0.203) | (-0.007) | (1.706) | (2.212) |
| Loan loss allowance | 0.577*** | 0.531*** | 0.478*** | 0.401*** | 0.348*** | 0.316*** | 0.276*** | 0.234*** |
| | (10.712) | (8.974) | (7.474) | (6.284) | (5.331) | (4.524) | (3.964) | (3.361) |
| Liquidity | -0.274* | -0.289* | -0.318** | -0.307** | -0.309** | -0.296** | -0.289* | -0.273* |
| | (-1.721) | (-1.905) | (-2.121) | (-2.076) | (-2.132) | (-2.019) | (-1.915) | (-1.773) |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11308 | 10966 | 10631 | 10303 | 9980 | 9666 | 9365 | 9073 |
| Adjusted $R^2$ | 0.555 | 0.512 | 0.483 | 0.470 | 0.450 | 0.443 | 0.441 | 0.441 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | 0.082*** | 0.079*** | 0.073*** | 0.048*** | 0.078*** | 0.055*** | 0.049** | 0.019 |
| | (4.942) | (4.519) | (3.740) | (2.670) | (3.529) | (2.638) | (2.335) | (0.794) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6907 | 6699 | 6495 | 6294 | 6098 | 5901 | 5702 | 5518 |
| Adjusted $R^2$ | 0.511 | 0.477 | 0.456 | 0.448 | 0.424 | 0.429 | 0.430 | 0.435 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.101*** | 0.123*** | 0.131*** | 0.095*** | 0.140*** | 0.114*** | 0.118*** | 0.067* |
| | (3.848) | (4.550) | (4.168) | (3.406) | (4.045) | (3.529) | (3.618) | (1.765) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3882 | 3742 | 3607 | 3476 | 3348 | 3223 | 3103 | 2988 |
| Adjusted $R^2$ | 0.496 | 0.469 | 0.447 | 0.437 | 0.418 | 0.417 | 0.412 | 0.411 |

Table 3: **Incremental $R^2$ and Prediction Exercise**

Table 3 presents the prediction results of bank-level *CLN* score for bank loan loss provisions. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are heteroskedasticity-robust standard errors clustered at the borrower level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| | (1) Controls only | (2) Z-score | (3) CLN score | (4) Both |
|---|---|---|---|---|
| ***Panel A: In-sample Predictions (Labeled Public Banks)*** | | | | |
| *CLN* score | | | 0.112*** | 0.112*** |
| | | | (4.835) | (4.813) |
| Bank Z-score | | 0.029 | | 0.014 |
| | | (0.333) | | (0.160) |
| Size | 0.045*** | 0.045*** | 0.041*** | 0.041*** |
| | (4.116) | (4.217) | (3.632) | (3.697) |
| Equity capital | -0.464 | -0.533 | -0.618 | -0.650 |
| | (-0.557) | (-0.632) | (-0.747) | (-0.772) |
| Deposits | 0.130 | 0.129 | 0.109 | 0.109 |
| | (1.130) | (1.126) | (0.955) | (0.954) |
| ROA | -0.142* | -0.142* | -0.125* | -0.125* |
| | (-1.953) | (-1.959) | (-1.802) | (-1.806) |
| Loan size | -0.040 | -0.038 | -0.041 | -0.040 |
| | (-0.223) | (-0.217) | (-0.225) | (-0.223) |
| Loan growth | -0.001 | -0.001 | -0.001 | -0.001 |
| | (-1.277) | (-1.263) | (-0.845) | (-0.839) |
| Loan loss allowance | 0.231*** | 0.234*** | 0.239*** | 0.240*** |
| | (3.610) | (3.497) | (3.735) | (3.602) |
| Liquidity | -0.324* | -0.320* | -0.301 | -0.300 |
| | (-1.758) | (-1.740) | (-1.597) | (-1.590) |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes |
| Observations | 4009 | 4009 | 4009 | 4009 |
| Adjusted $R^2$ | 0.515 | 0.515 | 0.520 | 0.520 |
| ***Panel B: Out-of-sample Predictions (Unlabeled Banks)*** | | | | |
| RMSE | 0.681 | 0.688 | 0.669 | 0.672 |
| Controls and Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes |
| ***Panel C: Out-of-sample Predictions (Private Banks)*** | | | | |
| RMSE | 0.715 | 0.721 | 0.703 | 0.706 |
| Controls and Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes |

Table 4: **Bank Co-Lending Network Risk, Bank Size and Bank Loan Loss Provisions**

Table 4 examines the heterogeneous effects of bank co-lending network (CLN) risk score on predicting h-quarter-ahead bank loan loss provisions for different size of banks. The Small bank dummy variable equals to 1 (0) if the bank asset size is lower than (greater than or equal to) the median by each year-quarter. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** **Loan loss provisions** | (1) h=1 | (2) h=2 | (3) h=3 | (4) h=4 | (5) h=5 | (6) h=6 | (7) h=7 | (8) h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Panel A: All samples*** | | | | | | | | |
| *CLN* score | 0.087*** | 0.078*** | 0.079*** | 0.047*** | 0.052*** | 0.035** | 0.027 | -0.002 |
| | (5.616) | (5.382) | (5.103) | (3.084) | (3.386) | (2.110) | (1.366) | (-0.116) |
| Small bank dummy | 0.006 | 0.015 | 0.032 | 0.036 | 0.027 | 0.035 | 0.037 | 0.036 |
| | (0.167) | (0.422) | (0.841) | (0.927) | (0.665) | (0.860) | (0.865) | (0.829) |
| *CLN* score × Small bank dummy | 0.040 | 0.063** | 0.060* | 0.062** | 0.094*** | 0.083** | 0.089** | 0.084** |
| | (1.482) | (2.310) | (1.922) | (2.077) | (2.961) | (2.524) | (2.440) | (2.064) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11308 | 10966 | 10631 | 10303 | 9980 | 9666 | 9365 | 9073 |
| Adjusted $R^2$ | 0.555 | 0.512 | 0.483 | 0.470 | 0.450 | 0.443 | 0.441 | 0.441 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | 0.057*** | 0.045** | 0.042** | 0.009 | 0.028 | 0.016 | 0.007 | -0.022 |
| | (3.159) | (2.341) | (2.040) | (0.457) | (1.350) | (0.743) | (0.311) | (-0.884) |
| Small bank dummy | 0.005 | 0.010 | 0.016 | 0.015 | 0.014 | 0.021 | 0.015 | 0.004 |
| | (0.101) | (0.214) | (0.309) | (0.279) | (0.257) | (0.394) | (0.282) | (0.066) |
| *CLN* score × Small bank dummy | 0.051 | 0.070** | 0.065 | 0.082** | 0.105*** | 0.082** | 0.090** | 0.087* |
| | (1.529) | (2.084) | (1.625) | (2.284) | (2.721) | (2.242) | (2.073) | (1.756) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6907 | 6699 | 6495 | 6294 | 6098 | 5901 | 5702 | 5518 |
| Adjusted $R^2$ | 0.511 | 0.478 | 0.457 | 0.448 | 0.426 | 0.430 | 0.431 | 0.436 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.053* | 0.062* | 0.085** | 0.041 | 0.069** | 0.065* | 0.088** | 0.025 |
| | (1.695) | (1.827) | (2.533) | (1.216) | (2.041) | (1.732) | (2.213) | (0.638) |
| Small bank dummy | -0.042 | -0.033 | -0.004 | 0.003 | 0.002 | 0.020 | 0.037 | -0.018 |
| | (-0.449) | (-0.361) | (-0.038) | (0.026) | (0.023) | (0.213) | (0.381) | (-0.169) |
| *CLN* score × Small bank dummy | 0.082* | 0.105** | 0.081 | 0.095* | 0.126** | 0.091* | 0.055 | 0.074 |
| | (1.668) | (2.056) | (1.350) | (1.772) | (2.285) | (1.679) | (0.891) | (1.004) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3882 | 3742 | 3607 | 3476 | 3348 | 3223 | 3103 | 2988 |
| Adjusted $R^2$ | 0.497 | 0.469 | 0.447 | 0.438 | 0.419 | 0.417 | 0.412 | 0.411 |

Table 5: **Bank Co-Lending Network Risk, Negative Earnings Shock and Bank Loan Loss Provision**

Table 5 presents the h-quarter-ahead prediction results of bank-level *CLN*-score for bank loan loss provisions, as moderated by banks' Negative earnings shock . The negative earning shock indicator equals to 1 if the bank has negative earning growth comparing to the last quarter and 0 otherwise. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

**All samples**

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Loan loss provisions** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All samples*** | | | | | | | | |
| *CLN* score | 0.055*** | 0.045*** | 0.053*** | 0.034** | 0.042*** | 0.044** | 0.047** | 0.022 |
| | (3.885) | (3.355) | (3.783) | (2.204) | (2.645) | (2.243) | (2.307) | (1.077) |
| Negative earnings shock | 0.005 | -0.003 | 0.014 | 0.016 | 0.016 | 0.042* | 0.040* | 0.026 |
| | (0.274) | (-0.178) | (0.725) | (0.760) | (0.792) | (1.855) | (1.751) | (1.060) |
| *CLN* score × Negative earnings shock | 0.117*** | 0.148*** | 0.121*** | 0.092*** | 0.117*** | 0.055* | 0.030 | 0.018 |
| | (4.175) | (5.380) | (3.931) | (3.380) | (3.955) | (1.701) | (0.848) | (0.539) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11308 | 10966 | 10631 | 10303 | 9980 | 9666 | 9365 | 9073 |
| Adjusted $R^2$ | 0.558 | 0.516 | 0.487 | 0.473 | 0.454 | 0.445 | 0.442 | 0.442 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | 0.044*** | 0.029* | 0.048** | 0.026 | 0.030 | 0.025 | 0.030 | -0.000 |
| | (2.855) | (1.765) | (2.587) | (1.406) | (1.546) | (1.085) | (1.209) | (-0.003) |
| Negative earnings shock | 0.030 | 0.011 | 0.043* | 0.048* | 0.036 | 0.054* | 0.052* | 0.046 |
| | (1.515) | (0.512) | (1.762) | (1.865) | (1.379) | (1.964) | (1.868) | (1.591) |
| *CLN* score × Negative earnings shock | 0.096** | 0.130*** | 0.065 | 0.057 | 0.124*** | 0.078* | 0.049 | 0.048 |
| | (2.438) | (3.451) | (1.421) | (1.528) | (2.806) | (1.858) | (1.040) | (1.116) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6907 | 6699 | 6495 | 6294 | 6098 | 5901 | 5702 | 5518 |
| Adjusted $R^2$ | 0.514 | 0.481 | 0.459 | 0.450 | 0.429 | 0.433 | 0.432 | 0.437 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.062** | 0.058** | 0.090*** | 0.049 | 0.066** | 0.073* | 0.088** | 0.036 |
| | (2.410) | (2.256) | (3.056) | (1.639) | (2.084) | (1.960) | (2.228) | (1.042) |
| Negative earnings shock | 0.030 | 0.012 | 0.055 | 0.053 | 0.040 | 0.081** | 0.091** | 0.094** |
| | (1.053) | (0.385) | (1.621) | (1.445) | (1.149) | (2.101) | (2.202) | (2.254) |
| *CLN* score × Negative earnings shock | 0.094 | 0.161*** | 0.102 | 0.110* | 0.184*** | 0.100 | 0.068 | 0.070 |
| | (1.510) | (2.690) | (1.438) | (1.968) | (2.906) | (1.468) | (0.904) | (1.008) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3882 | 3742 | 3607 | 3476 | 3348 | 3223 | 3103 | 2988 |
| Adjusted $R^2$ | 0.498 | 0.473 | 0.451 | 0.441 | 0.425 | 0.422 | 0.416 | 0.416 |

Table 6: **Bank Co-Lending Network Risk, ROA Volatility and Bank Loan Loss Provisions**

Table 6 presents the h-quarter-ahead prediction results of bank-level *CLN*-score for bank loan loss provisions, as moderated by banks' return on assets (ROA) volatility. The High ROA volatility dummy equals to 1 if the bank's ROA standard deviation in the last five years are ranked in the top 25% percentile within each year-quarter and 0 otherwise. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

**All samples**

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Loan loss provisions** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All Samples*** | | | | | | | | |
| *CLN* score | 0.081*** | 0.084*** | 0.089*** | 0.056*** | 0.077*** | 0.058*** | 0.049** | 0.015 |
| | (6.184) | (5.994) | (5.678) | (3.580) | (4.356) | (3.172) | (2.482) | (0.665) |
| High ROA volatility | 0.038 | 0.022 | 0.020 | 0.014 | 0.001 | -0.023 | -0.041 | -0.042 |
| | (1.255) | (0.675) | (0.584) | (0.405) | (0.033) | (-0.650) | (-1.123) | (-1.059) |
| *CLN* score × High ROA volatility | 0.116*** | 0.113** | 0.085* | 0.102** | 0.089* | 0.076** | 0.072* | 0.067 |
| | (2.919) | (2.461) | (1.905) | (2.186) | (1.948) | (1.981) | (1.934) | (1.635) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 10382 | 10054 | 9732 | 9420 | 9118 | 8826 | 8547 | 8275 |
| Adjusted $R^2$ | 0.575 | 0.531 | 0.499 | 0.485 | 0.466 | 0.459 | 0.453 | 0.452 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | 0.051*** | 0.048*** | 0.055** | 0.024 | 0.051** | 0.038 | 0.032 | 0.001 |
| | (3.262) | (2.784) | (2.567) | (1.154) | (2.050) | (1.536) | (1.232) | (0.018) |
| High ROA volatility | 0.054 | 0.026 | 0.030 | 0.032 | 0.004 | -0.028 | -0.014 | -0.010 |
| | (1.358) | (0.629) | (0.681) | (0.712) | (0.107) | (-0.634) | (-0.282) | (-0.183) |
| *CLN* score × High ROA volatility | 0.140** | 0.144** | 0.085 | 0.121* | 0.146** | 0.127** | 0.084 | 0.074 |
| | (2.465) | (2.299) | (1.488) | (1.815) | (2.225) | (2.464) | (1.609) | (1.177) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6315 | 6119 | 5925 | 5733 | 5549 | 5364 | 5179 | 5006 |
| Adjusted $R^2$ | 0.538 | 0.505 | 0.480 | 0.467 | 0.447 | 0.450 | 0.445 | 0.451 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.038 | 0.064** | 0.098*** | 0.053 | 0.103** | 0.091** | 0.112** | 0.041 |
| | (1.532) | (2.476) | (2.629) | (1.552) | (2.454) | (2.248) | (2.535) | (0.778) |
| High ROA volatility | 0.064 | 0.044 | 0.047 | 0.046 | 0.002 | -0.042 | -0.041 | -0.066 |
| | (1.086) | (0.711) | (0.690) | (0.670) | (0.026) | (-0.592) | (-0.523) | (-0.756) |
| *CLN* score × High ROA volatility | 0.188** | 0.173** | 0.084 | 0.126 | 0.137 | 0.110 | 0.018 | 0.075 |
| | (2.453) | (2.093) | (1.077) | (1.395) | (1.524) | (1.576) | (0.245) | (0.835) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3485 | 3356 | 3230 | 3107 | 2988 | 2872 | 2763 | 2658 |
| Adjusted $R^2$ | 0.527 | 0.501 | 0.475 | 0.461 | 0.443 | 0.440 | 0.431 | 0.427 |

Table 7: **Bank Co-Lending Network Risk, Bank Complexity and Bank Loan Loss Provisions**

Table 7 presents the h-quarter-ahead prediction results of bank-level *CLN*-score for bank loan loss provisions, interacting with *High complexity.* The High complexity equals to 1 if the bank's ROA standard deviation in the last five years are ranked in the top 25% percentile within each year-quarter and 0 otherwise. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

**All samples**

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Loan loss provisions** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All Samples*** | | | | | | | | |
| *CLN* score | 0.096*** | 0.091*** | 0.093*** | 0.059*** | 0.075*** | 0.056*** | 0.048*** | 0.018 |
| | (6.794) | (6.598) | (6.441) | (4.184) | (4.971) | (3.578) | (2.958) | (0.965) |
| High complexity | 0.007 | 0.002 | 0.019 | 0.016 | 0.015 | 0.012 | 0.016 | 0.021 |
| | (0.199) | (0.065) | (0.515) | (0.428) | (0.393) | (0.297) | (0.371) | (0.437) |
| *CLN* score × High complexity | 0.043 | 0.074** | 0.061 | 0.073 | 0.087* | 0.077 | 0.085 | 0.085 |
| | (1.297) | (2.062) | (1.448) | (1.629) | (1.663) | (1.433) | (1.460) | (1.278) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11308 | 10966 | 10631 | 10303 | 9980 | 9666 | 9365 | 9073 |
| Adjusted $R^2$ | 0.556 | 0.512 | 0.484 | 0.471 | 0.451 | 0.443 | 0.441 | 0.442 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | 0.066*** | 0.055*** | 0.051** | 0.020 | 0.049** | 0.031 | 0.019 | -0.001 |
| | (3.644) | (3.006) | (2.442) | (1.050) | (2.441) | (1.529) | (0.917) | (-0.049) |
| High complexity | 0.048 | 0.038 | 0.038 | 0.031 | 0.034 | 0.032 | 0.025 | 0.046 |
| | (1.051) | (0.784) | (0.767) | (0.595) | (0.647) | (0.594) | (0.434) | (0.694) |
| *CLN* score × High complexity | 0.070* | 0.100** | 0.094* | 0.121** | 0.126** | 0.108** | 0.138** | 0.100 |
| | (1.685) | (2.304) | (1.935) | (2.464) | (2.097) | (1.970) | (2.196) | (1.377) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6907 | 6699 | 6495 | 6294 | 6098 | 5901 | 5702 | 5518 |
| Adjusted $R^2$ | 0.512 | 0.479 | 0.458 | 0.450 | 0.426 | 0.431 | 0.432 | 0.437 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.064** | 0.081*** | 0.103*** | 0.052* | 0.086*** | 0.069** | 0.076** | 0.038 |
| | (2.141) | (2.692) | (2.813) | (1.696) | (2.721) | (2.061) | (2.327) | (0.888) |
| High complexity | 0.115 | 0.117 | 0.140* | 0.126 | 0.119 | 0.121 | 0.129 | 0.146 |
| | (1.516) | (1.437) | (1.761) | (1.454) | (1.376) | (1.335) | (1.327) | (1.306) |
| *CLN* score × High complexity | 0.106* | 0.123** | 0.090 | 0.131** | 0.164** | 0.144** | 0.135* | 0.101 |
| | (1.935) | (2.142) | (1.346) | (2.030) | (2.165) | (2.049) | (1.692) | (1.024) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3882 | 3742 | 3607 | 3476 | 3348 | 3223 | 3103 | 2988 |
| Adjusted $R^2$ | 0.501 | 0.474 | 0.453 | 0.443 | 0.425 | 0.423 | 0.418 | 0.417 |

Table 8: **Bank Co-Lending Network Risk, Bank Capital Risk and Bank Loan Loss Provisions**

Table 8 examines the heterogeneous effects of bank co-lending network (CLN) risk score on predicting h-quarter-ahead bank loan loss provisions for banks with different capital risk. The Low bank capital dummy equals to 1 (0) if the bank's risk capital is in the bottom 25% within each year-quarter. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

**All Samples**

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Loan loss provisions** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All samples*** | | | | | | | | |
| *CLN* score | 0.086*** | 0.084*** | 0.087*** | 0.068*** | 0.084*** | 0.063*** | 0.065*** | 0.036 |
| | (6.655) | (6.619) | (6.156) | (4.276) | (4.863) | (3.420) | (3.386) | (1.562) |
| Low capital adequacy | -0.017 | -0.018 | 0.001 | 0.034 | 0.016 | 0.020 | 0.034 | 0.035 |
| | (-0.608) | (-0.713) | (0.026) | (1.315) | (0.672) | (0.688) | (1.031) | (1.006) |
| *CLN* score × Low capital adequacy | 0.072** | 0.084*** | 0.064* | 0.016 | 0.029 | 0.025 | -0.009 | -0.017 |
| | (2.330) | (2.596) | (1.715) | (0.489) | (0.848) | (0.652) | (-0.243) | (-0.459) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11308 | 10966 | 10631 | 10303 | 9980 | 9666 | 9365 | 9073 |
| Adjusted $R^2$ | 0.556 | 0.512 | 0.484 | 0.471 | 0.450 | 0.443 | 0.441 | 0.441 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | 0.061*** | 0.048*** | 0.048** | 0.033 | 0.052** | 0.038* | 0.045* | 0.016 |
| | (3.932) | (2.947) | (2.476) | (1.647) | (2.356) | (1.662) | (1.791) | (0.544) |
| Low capital adequacy | -0.043 | -0.047 | -0.013 | 0.031 | -0.004 | -0.001 | 0.033 | 0.043 |
| | (-1.328) | (-1.470) | (-0.365) | (0.925) | (-0.135) | (-0.018) | (0.779) | (0.961) |
| *CLN* score × Low capital adequacy | 0.082** | 0.117*** | 0.093* | 0.051 | 0.094** | 0.062 | 0.013 | 0.007 |
| | (2.050) | (2.873) | (1.796) | (1.190) | (2.116) | (1.236) | (0.235) | (0.146) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6907 | 6699 | 6495 | 6294 | 6098 | 5901 | 5702 | 5518 |
| Adjusted $R^2$ | 0.511 | 0.478 | 0.457 | 0.449 | 0.425 | 0.430 | 0.430 | 0.436 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.069*** | 0.073*** | 0.086*** | 0.068** | 0.103*** | 0.091** | 0.108*** | 0.060 |
| | (2.767) | (2.891) | (2.813) | (2.132) | (2.873) | (2.481) | (2.687) | (1.191) |
| Low capital adequacy | -0.072 | -0.096* | -0.067 | -0.001 | -0.022 | -0.023 | 0.007 | 0.021 |
| | (-1.524) | (-1.976) | (-1.277) | (-0.013) | (-0.497) | (-0.388) | (0.112) | (0.309) |
| *CLN* score × Low capital adequacy | 0.120** | 0.181*** | 0.163** | 0.090 | 0.126* | 0.085 | 0.030 | 0.020 |
| | (1.992) | (2.947) | (2.050) | (1.361) | (1.939) | (1.112) | (0.357) | (0.258) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3882 | 3742 | 3607 | 3476 | 3348 | 3223 | 3103 | 2988 |
| Adjusted $R^2$ | 0.498 | 0.471 | 0.449 | 0.439 | 0.420 | 0.417 | 0.412 | 0.411 |

Table 9: **Bank Co-Lending Network Risk and Bank Non-performing Loans**

Table 9 presents the h-quarter-ahead prediction results of bank-level *CLN*-score for bank non-performing loans. The bank-level *CLN*-score and control variables are measured at time $t$, and the dependent variable is measured at time $t+h$. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables in the Panel B and C. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

**All Samples**

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Non-peforming Loans** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All samples*** | | | | | | | | |
| *CLN* score | 0.237*** | 0.261*** | 0.273*** | 0.272*** | 0.275*** | 0.255*** | 0.252*** | 0.238*** |
| | (4.091) | (4.529) | (4.665) | (4.641) | (4.413) | (4.215) | (3.899) | (3.458) |
| Size | -0.012 | -0.013 | -0.013 | -0.015 | -0.015 | -0.012 | -0.010 | -0.006 |
| | (-0.357) | (-0.378) | (-0.352) | (-0.393) | (-0.368) | (-0.307) | (-0.253) | (-0.148) |
| Equity capital | 0.436 | 0.205 | 0.095 | -0.315 | -0.641 | -0.745 | -0.818 | -0.846 |
| | (0.207) | (0.097) | (0.046) | (-0.142) | (-0.272) | (-0.322) | (-0.363) | (-0.394) |
| Deposits | 0.167 | 0.151 | 0.106 | 0.052 | 0.042 | 0.014 | -0.002 | 0.004 |
| | (0.413) | (0.360) | (0.248) | (0.116) | (0.089) | (0.028) | (-0.004) | (0.008) |
| ROA | -0.446*** | -0.406*** | -0.339*** | -0.249*** | -0.176* | -0.151* | -0.106 | -0.073 |
| | (-3.163) | (-3.166) | (-3.103) | (-2.702) | (-1.961) | (-1.827) | (-1.281) | (-0.971) |
| Loan size | -0.736 | -0.692 | -0.554 | -0.431 | -0.320 | -0.165 | -0.043 | 0.100 |
| | (-1.522) | (-1.350) | (-1.026) | (-0.739) | (-0.543) | (-0.287) | (-0.075) | (0.179) |
| Loan growth | -0.480** | -0.001 | -0.275 | -0.157 | -0.107 | 0.062 | 0.276 | 0.335 |
| | (-1.994) | (-0.332) | (-1.491) | (-0.861) | (-0.549) | (0.300) | (1.144) | (1.325) |
| Loan loss allowance | 1.296*** | 1.264*** | 1.189*** | 1.095*** | 0.993*** | 0.882*** | 0.791*** | 0.677*** |
| | (5.403) | (5.059) | (4.732) | (4.212) | (3.999) | (3.845) | (3.651) | (3.356) |
| Liquidity | -1.260*** | -1.344*** | -1.398*** | -1.481*** | -1.539*** | -1.554*** | -1.570*** | -1.574*** |
| | (-3.279) | (-3.358) | (-3.355) | (-3.356) | (-3.301) | (-3.175) | (-3.066) | (-2.924) |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 10279 | 10036 | 9797 | 9504 | 9270 | 9039 | 8812 | 8585 |
| Adjusted $R^2$ | 0.431 | 0.396 | 0.359 | 0.324 | 0.301 | 0.285 | 0.273 | 0.264 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | 0.217*** | 0.231*** | 0.241*** | 0.258*** | 0.276*** | 0.245*** | 0.237*** | 0.223** |
| | (3.265) | (3.306) | (3.381) | (3.477) | (3.410) | (3.230) | (2.884) | (2.383) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6340 | 6191 | 6040 | 5862 | 5714 | 5559 | 5400 | 5250 |
| Adjusted $R^2$ | 0.423 | 0.379 | 0.343 | 0.302 | 0.273 | 0.259 | 0.251 | 0.238 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.310*** | 0.349*** | 0.368*** | 0.403*** | 0.451*** | 0.402*** | 0.390*** | 0.372** |
| | (3.179) | (3.303) | (3.417) | (3.538) | (3.515) | (3.270) | (2.999) | (2.460) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3628 | 3520 | 3415 | 3301 | 3198 | 3093 | 2991 | 2889 |
| Adjusted $R^2$ | 0.451 | 0.403 | 0.359 | 0.308 | 0.269 | 0.246 | 0.226 | 0.207 |

Table 10: **Bank Co-Lending Network Risk and Other Public Bank Risk Metrics**

Table 10 presents the h-quarter-ahead prediction results of bank-level *CLN*-score for some other public bank risk measures. We use stock market-based risk measures including Merton (1974) default probability, Nagel and Purnanandam (2020) modified default probability, the natural logarithm of idiosyncratic volatility, and use only the sample of public banks. The bank-level *CLN*-score and control variables are measured at time $t$, and the dependent variable is measured at time $t+h$. In all specifications, we include the same set of controls as in the baseline and control for year-quarter fixed effects. For simplicity, we do not report the coefficient estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1)<br>h=1 | (2)<br>h=2 | (3)<br>h=3 | (4)<br>h=4 | (5)<br>h=5 | (6)<br>h=6 | (7)<br>h=7 | (8)<br>h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Dependent: Merton default probability*** | | | | | | | | |
| *CLN* score | 0.032*** | 0.037*** | 0.042*** | 0.039*** | 0.034*** | 0.029*** | 0.027*** | 0.025*** |
| | (6.434) | (7.281) | (7.905) | (7.033) | (5.824) | (4.996) | (4.448) | (4.081) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6077 | 5902 | 5728 | 5553 | 5377 | 5206 | 5042 | 4879 |
| Adjusted $R^2$ | 0.730 | 0.724 | 0.712 | 0.699 | 0.688 | 0.675 | 0.664 | 0.660 |
| ***Dependent: Modified default probability*** | | | | | | | | |
| *CLN* score | 0.040*** | 0.040*** | 0.041*** | 0.037*** | 0.033*** | 0.030*** | 0.028*** | 0.027*** |
| | (8.642) | (8.216) | (8.269) | (7.408) | (6.256) | (5.696) | (4.959) | (4.452) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6077 | 5902 | 5728 | 5553 | 5377 | 5206 | 5042 | 4879 |
| Adjusted $R^2$ | 0.690 | 0.679 | 0.658 | 0.639 | 0.622 | 0.601 | 0.585 | 0.575 |
| ***Dependent: ln(IVOL)*** | | | | | | | | |
| *CLN* score | 0.060*** | 0.064*** | 0.051*** | 0.044*** | 0.038*** | 0.032*** | 0.025** | 0.016 |
| | (5.529) | (5.915) | (4.340) | (3.833) | (3.233) | (2.603) | (2.009) | (1.176) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7976 | 7764 | 7554 | 7347 | 7138 | 6935 | 6740 | 6547 |
| Adjusted $R^2$ | 0.686 | 0.684 | 0.679 | 0.679 | 0.678 | 0.675 | 0.674 | 0.672 |

## Table 11: **Bank Co-Lending Network Risk and Bank Profitability**

Table 11 presents the h-quarter-ahead prediction results of bank-level *CLN*-score for bank return on assets (ROA). The bank-level *CLN*-score and control variables are measured at time $t$, and the dependent variable is measured at time $t + h$. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables in the Panel B and C. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Return on assets** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All samples*** | | | | | | | | |
| *CLN* score | -0.270*** | -0.244*** | -0.234*** | -0.213*** | -0.204*** | -0.185*** | -0.153*** | -0.127*** |
| | (-9.928) | (-8.904) | (-8.740) | (-7.287) | (-7.676) | (-6.770) | (-5.917) | (-4.530) |
| Size | 0.040** | 0.037* | 0.035* | 0.035* | 0.032 | 0.030 | 0.028 | 0.028 |
| | (2.188) | (1.939) | (1.853) | (1.792) | (1.650) | (1.514) | (1.369) | (1.349) |
| Equity capital | 4.320*** | 3.635*** | 3.394*** | 3.242*** | 3.269** | 3.166** | 3.059* | 2.969* |
| | (4.994) | (3.894) | (3.186) | (2.666) | (2.393) | (2.178) | (1.954) | (1.878) |
| Deposits | -0.004 | 0.011 | 0.063 | 0.104 | 0.122 | 0.146 | 0.169 | 0.215 |
| | (-0.021) | (0.053) | (0.305) | (0.493) | (0.561) | (0.643) | (0.725) | (0.922) |
| Loan size | 0.830*** | 0.710*** | 0.601*** | 0.477** | 0.349 | 0.276 | 0.193 | 0.121 |
| | (3.645) | (3.195) | (2.702) | (2.107) | (1.525) | (1.183) | (0.844) | (0.556) |
| Loan growth | 0.000 | 0.000 | 0.000 | 0.002** | 0.001 | -0.001 | -0.001 | -0.004** |
| | (0.683) | (1.178) | (0.355) | (2.320) | (0.830) | (-0.647) | (-0.725) | (-2.170) |
| Loan loss allowance | -0.408*** | -0.344*** | -0.299*** | -0.237*** | -0.171** | -0.150** | -0.118* | -0.103* |
| | (-4.355) | (-3.967) | (-3.680) | (-3.190) | (-2.527) | (-2.240) | (-1.934) | (-1.822) |
| Liquidity | 0.598*** | 0.558*** | 0.536** | 0.539** | 0.527** | 0.474** | 0.422* | 0.363 |
| | (3.036) | (2.740) | (2.585) | (2.557) | (2.464) | (2.164) | (1.879) | (1.625) |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11310 | 10968 | 10633 | 10305 | 9982 | 9667 | 9365 | 9073 |
| Adjusted $R^2$ | 0.310 | 0.277 | 0.263 | 0.248 | 0.241 | 0.235 | 0.232 | 0.226 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | -0.210*** | -0.186*** | -0.179*** | -0.160*** | -0.150*** | -0.131*** | -0.119*** | -0.083** |
| | (-6.537) | (-5.529) | (-4.578) | (-3.713) | (-3.957) | (-3.876) | (-3.338) | (-2.003) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6909 | 6701 | 6497 | 6296 | 6100 | 5902 | 5702 | 5518 |
| Adjusted $R^2$ | 0.298 | 0.252 | 0.246 | 0.228 | 0.218 | 0.221 | 0.226 | 0.231 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | -0.267*** | -0.231*** | -0.255*** | -0.245*** | -0.255*** | -0.208*** | -0.173*** | -0.114* |
| | (-5.535) | (-4.765) | (-4.223) | (-3.430) | (-4.049) | (-3.694) | (-3.015) | (-1.658) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3884 | 3744 | 3609 | 3478 | 3350 | 3224 | 3103 | 2988 |
| Adjusted $R^2$ | 0.274 | 0.241 | 0.237 | 0.231 | 0.226 | 0.215 | 0.203 | 0.194 |

Table 12: **Difference-in-Differences Estimation: Credit Downgrade Spillover**

Table 12 shows the results of the difference-in-differences estimation using the focal banks' credit downgrade as exogenous shocks to the *CLN* score of neighboring banks in the co-lending network. The treatment events are long-term credit downgrade at the entity level from S&P credit ratings for banks in our co-lending network. In each event cohort, treatment groups are banks that have a inward direction with the focal bank (from the focal bank to treatment banks) and control groups are banks that are not connected with the focal banks. *Treat* equals to 1 for treatment banks and 0 for control banks. In each cohort, we use a three quarters before and three quarters after ($t-3$, $t+3$) event window within each cohort. *Post* dummy equals to 0 (1) for all quarters before (after) the credit event in each cohort, and time dummies $d_j$ equals 1 for the year that is j quarter(s) after the treatment. Bank controls are same set of variables as in Table 2 and lagged-one-period. Definitions of other variables are provided in Table A1 in the Appendix. Numbers in parentheses are standard errors double-clustered at both cohort and bank levels. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

***Panel A: Standard DiD***

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Treat × Post | 0.114** | 0.121*** | 0.157*** | 0.159** | 0.110** | 0.113** |
| | (0.040) | (0.040) | (0.060) | (0.040) | (0.039) | (0.041) |
| Treat | -0.056 | | 0.051 | | -0.060 | |
| | (0.043) | | (0.068) | | (0.042) | |
| Post | 0.054 | | 0.020 | | 0.067 | |
| | (0.056) | | (0.060) | | (0.065) | |
| Bank Controls | No | No | No | No | Yes | Yes |
| Severe Downgrade Only | No | No | Yes | Yes | No | No |
| Year-Quarter Fixed Effects | Yes | No | Yes | No | Yes | No |
| Bank Fixed Effects | Yes | No | Yes | No | Yes | No |
| Cohort-YrQtr Fixed Effects | No | Yes | No | Yes | No | Yes |
| Cohort-Bank Fixed Effects | No | Yes | No | Yes | No | Yes |
| Observations | 3,393 | 3,393 | 1,204 | 1,204 | 3,392 | 3,392 |
| Adjusted $R^2$ | 0.197 | 0.217 | 0.232 | 0.259 | 0.210 | 0.227 |

***Panel B: Dynamic DiD***

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Treat × $d_{-2}$ | 0.090 | 0.090 | -0.023 | -0.021 | 0.092 | 0.093 |
| | (0.107) | (0.106) | (0.184) | (0.183) | (0.105) | (0.104) |
| Treat × $d_{-1}$ | 0.117 | 0.121 | 0.051 | 0.053 | 0.122 | 0.120 |
| | (0.079) | (0.079) | (0.138) | (0.131) | (0.077) | (0.078) |
| Treat × $d_0$ | 0.127 | 0.132 | 0.118 | 0.122 | 0.132 | 0.130 |
| | (0.079) | (0.079) | (0.137) | (0.134) | (0.080) | (0.078) |
| Treat × $d_1$ | 0.137 | 0.142 | 0.080 | 0.083 | 0.137 | 0.136 |
| | (0.088) | (0.088) | (0.143) | (0.138) | (0.088) | (0.087) |
| Treat × $d_2$ | 0.198** | 0.203** | 0.137 | 0.140 | 0.199** | 0.199** |
| | (0.073) | (0.072) | (0.092) | (0.082) | (0.072) | (0.072) |
| Treat × $d_3$ | 0.202** | 0.208** | 0.293** | 0.295** | 0.208** | 0.210** |
| | (0.076) | (0.075) | (0.100) | (0.098) | (0.078) | (0.079) |
| Treat and Time Dummies | Yes | No | Yes | No | Yes | No |
| Bank Controls | No | No | No | No | Yes | Yes |
| Severe Downgrade Only | No | No | Yes | Yes | No | No |
| Year-Quarter Fixed Effects | Yes | No | Yes | No | Yes | No |
| Bank Fixed Effects | Yes | No | Yes | No | Yes | No |
| Cohort-YrQtr Fixed Effects | No | Yes | No | Yes | No | Yes |
| Cohort-Bank Fixed Effects | No | Yes | No | Yes | No | Yes |
| Observations | 4,008 | 4,008 | 1,409 | 1,409 | 4,007 | 4,007 |
| Adjusted $R^2$ | 0.214 | 0.252 | 0.268 | 0.303 | 0.225 | 0.259 |

# Appendix

Table A1: **Variable Definition**

| Variable | Definition | Source |
|---|---|---|
| **Bank level variables** | | |
| CLN score | Bank-level Co-Lending Network (CLN) risk score estimated by the CoLGNN model using FR Y-9C bank features and DealScan loan features. The detailed estimation process is described in Section 3.2.2 and Section 4.1. | CoLGNN |
| Loan loss provisions | The ratio of loan loss provisions (BHCK4230) to total assets (BHCK2170) in percentage points. | FR Y-9C |
| Nonperforming loans | The ratio of nonperforming loans (we use the sum of BHCK5525 and BHCK5526 for the period before 2017Q4 and the sum of BHCK1403 and BHCK1407 after 2017) to total assets (BHCK2170) in percentage points. | FR Y-9C |
| ROA | Banks' net income (BHCK4340) divided by bank total assets (BHCK2170) | FR Y-9C |
| Size | The natural logarithm of bank total assets (BHCK2170) | FR Y-9C |
| Equity capital | Total equity capital (BHCK3210) divided by bank total assets (BHCK2170). | FR Y-9C |
| Deposits | Total interest-bearing deposits (BHCK3517) devided by bank total assets (BHCK2170). | FR Y-9C |
| Loan size | Total loans (BHCK2122) divided by bank total assets (BHCK2170). | FR Y-9C |
| Loan growth | The annual percentage growth rate of total loans (BHCK2122). | FR Y-9C |
| Loan loss allowance | The allowance for loan and lease losses (BHCK3123) divided by bank total assets (BHCK2170). | FR Y-9C |
| Liquidity | The sum of cash (BHCK0010) and short-term securities (BHCK1773) divided by bank total assets (BHCK2170). | FR Y-9C |
| Standardized earnings change | The year-over-year quarterly change of earnings per share (BHCK4340 divided by BHCK3459) normalized by the rolling-window standard deviation of the earning change. | FR Y-9C |
| Bank capital adequacy | The ratio of loan loss provisions (BHCK4230) to total assets (BHCK2170) in percentage points. | FR Y-9C |
| Reputation | The total dollar amount of syndicated loans lead-arranged by a bank holding company normalized by the total amount of all syndicated loans in the past 12 months. | DealScan |
| Specialization in syndicated loan | The total amount of the syndicated loans a bank participates over the past 12 months divided by the bank's total loans (BHCK2122). | DealScan & FR Y-9C |
| Specialization in industry | The borrower's industry concentration of a bank's syndicated loan portfolio over the past 12 months, measured by the Herfindahl–Hirschman (HHI) index based on the borrowers' 2-digit SIC codes and the total loan amounts. | DealScan |
| Merton default probability | The default probability estimated by Merton (1974) model via a KMV iterative approach. | Nagel and Purnanandam (2020) |
| Modified default probability | Bank modified default probability using the data and code from Nagel and Purnanandam (2020). | Nagel and Purnanandam (2020) |
| ln(IVOL) | The natural logarithm of idiosyncratic volatility (IVOL) of bank stock returns, calculated by the standard deviation of the residuals from the Fama-French three-factor model estimated for each year-quarter. | CRSP & Kenneth R. French Data Library |

# Internet Appendix

## IA.1 Estimating Bank Co-Lending Network Risk (CLN) Score

The CLN score is estimated by employing the CoLGNN framework in Section 3.2.2 on a series of rolling-window co-lending networks from 1991Q1 to 2020Q4 with a rich sample of bank features from Form FR Y-9C and loan features from Dealscan. This appendix details our semi-supervised estimation process. As shown in *stage 1* of Figure 1, we calculate the quarterly buy-and-hold returns for all public banks. Banks are classified according to their performance, with the top 25% labeled as "safe" and the bottom 25% as "risky".[26] The remaining public banks and all private banks are treated as unlabeled samples.

*Stage 2* of Figure 1 displays the design of co-lending network. For each year-quarter $t$, we construct a co-lending network using syndicated loans originated in the past five years (20 quarters) from $t-1$ to $t-20$. For example, the co-lending network $G_{t=1996Q1}$ includes the syndicated loan originated from 1991Q1 to 1995Q4. In total, we construct 100 co-lending networks ($\mathcal{G} = \{G_{t=1996Q1}, ..., G_{t=2020Q4}\}$). Each co-lending network is an attributed graph $G_t = (\mathcal{V}, \mathcal{E}, \mathbf{X}, \mathbf{E})$ where $\mathcal{V} = \{v_1, \ldots, v_{|\mathcal{V}|}\}$ is the set of bank holding companies (nodes), and $\mathbf{X}$ are the bank characteristics at year-quarter $t-1$. $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of co-lending relationships (edges). Each edge $e_{i,j} \in \mathcal{E}$ is the directed edge points from bank $i$ to bank $j$. $\mathbf{E}$ utilize the loan characteristics on the origination date. We aggregate multiple co-lending relationships between two banks within the rolling window into one co-lending relationship. The summary statistics of our co-lending network series are shown in Table IA.1.1.

Next, we implement the CoLGNN model separately for each co-lending network $G_t$. Figure 2 illustrates the framework of the graph diffusion convolution module in CoLGNN. As shown in *Section 3.1* of Figure 2, we include the common bank characteristics from FR Y9-C as the node features and high-dimensional edge characteristics from DealScan. Both numerical and categorical data types are utilized. For categorical features, we apply one-hot vector

[26] This classification is tested for robustness with alternative thresholds, such as the top and bottom 33%, as detailed in the Supplementary Materials.

encoding to transform attributes such as 'Seniority Type', 'Primary Purpose', and 'Repayment Type' into numerical vectors. When multiple co-lending relationships exist within the specified time window, we combine these by calculating the loan amount weighted average of individual features from each loan. Since the graph neural network framework requires complete feature matrices without missing values, we fill missing features appropriately: numerical variables are filled with mean or zero, and categorical variables are filled with "No" or "unknown".[27]

We employ a semi-supervised learning framework to leverage the available stock market data from public banks to infer risk dynamics in private banks, which typically lack transparent financial data. We transform the risk estimation into a binary node classification task, labelling the safe group as risky group as $Y_{i,t} = 0$ and the risky group as $Y_{i,t} = 1$, with remaining banks unlabeled ($Y_{i,t} = NaN$). For model training, we split the labeled data into training, validation, and testing subsets in proportions of 70%, 20%, and 10%, respectively.

The training is executed over 200 epochs, employing an Adam optimizer with a learning rate and weight decay as specified in the model configuration. This optimizer is known for its effective handling of sparse gradients and adaptive learning rate adjustments. We also incorporate a learning rate scheduler to adjust the learning rate based on the validation loss, enhancing model convergence. Our CoLGNN model architecture includes two diffusion graph convolution modules, with each node passing through two layers of diffusion graph convolution. These layers are designed to incorporate both node features and edge attributes effectively, allowing the model to capture complex dependencies in the data:

1. **Initial Feature Transformation:** Each node feature vector undergoes a transformation via the first diffusion graph convolution layer, which integrates incoming features

[27] We fill the numerical variables with either mean or zero and categorical variables with either category "No" or category "unknown", according to the nature of variable characteristics. 1) We fill in zero for variables such as "All in Spread Undrawn", "Assignment Minimum", variables related to fees, variables related to participation structures, etc. 2) We fill the loan-amount-weighted mean for the variables such as "All In Spread Drawn bps", "Maturity", etc. 3) We fill the category "No" for variables such as "Secured Type", "Collateral Type", "Secondary Purpose", etc. 4) We fill the category "unknown" for variables such as "Distribution Method", "Tranche Type", "Repayment Type", etc.

from connected nodes, applying a Rectified Linear Unit (ReLU) activation and dropout regularization to prevent overfitting.

2. **Feature Aggregation and Classification:** The transformed features are then processed through a second diffusion graph convolution layer, aggregating further neighborhood information and passing through a softmax function to yield a two-dimensional output vector per node. This vector represents the probability of each bank being "safe" or "risky".

The highest performing model is selected based on the combined accuracy across the training and validation sets, ensuring the model not only fits well to the training data but also generalizes effectively to unseen data. The CoLGNN model's output is a two-dimensional probability vector showing the probability associated with the "risky" classification. We use the probability associated with the "risky" class as the estimated bank CLN score. Overall, by iterating through 200 epochs, the model dynamically adjusts its weights and biases to minimize prediction errors, refining its ability to distinguish between "safe" and "risky" banks based on their embedded features and topological structure within the co-lending network.

**Algorithm 1:** CoLGNN module

```
input : Graph 𝒢(ℰ, 𝒱); node features {x_i, ∀i ∈ 𝒱}; edge features {e_ij, ∀ij ∈ ℰ};
        out/in transition probabilities {ã_{out/in,i}, ∀i ∈ 𝒱}; layers L; random-walk
        depth K; weight matrices W; activation ReLU(·); sigmoid activation σ(·);
h_i^(0) ← x_i, ∀i ∈ 𝒱;
for l = 1, ..., L do
    for i ∈ 𝒱 do
        h̃_{out,i}^(l-1) ← h_i^(l-1) + Σ_{j∈𝒩_out(i)} ReLU(h_j^(l-1) + e_{i,j} W);      ▷ Edge fusion (out)
        h̃_{in,i}^(l-1) ← h_i^(l-1) + Σ_{r∈𝒩_in(i)} ReLU(h_r^(l-1) + e_{r,i} W);        ▷ Edge fusion (in)
        h_i^(l) ← Σ_{k=0}^{K-1} ( Σ_i ã_{out,i}^k h̃_{out,i}^(l-1) W_{k,out} + Σ_i ã_{in,i}^k h̃_{in,i}^(l-1) W_{k,in} );  ▷ Diffusion conv
    end
end
CLN score_i ← σ(h_i^(L))
output: CLN score_i, ∀i ∈ 𝒱
```

Table IA.1.1: **Directed Bank Co-Lending Networks (1991-2020)**

Table IA.1.1 presents the summary statistics of bank co-lending networks from 1991Q1 to 2020Q4.

| 1991-2000 | | | 2001-2010 | | | 2011-2020 | | |
|---|---|---|---|---|---|---|---|---|
| **Time Frame** | **Nodes** | **Edges** | **Time Frame** | **Nodes** | **Edges** | **Time Frame** | **Nodes** | **Edges** |
| 1991Q1-1995Q4 | 77 | 530 | 2001Q1-2005Q4 | 164 | 1306 | 2011Q1-2015Q4 | 179 | 2279 |
| 1991Q2-1996Q1 | 80 | 543 | 2001Q2-2006Q1 | 164 | 1313 | 2011Q2-2016Q1 | 179 | 2300 |
| 1991Q3-1996Q2 | 81 | 587 | 2001Q3-2006Q2 | 162 | 1329 | 2011Q3-2016Q2 | 178 | 2296 |
| 1991Q4-1996Q3 | 80 | 596 | 2001Q4-2006Q3 | 163 | 1347 | 2011Q4-2016Q3 | 189 | 2485 |
| 1992Q1-1996Q4 | 81 | 631 | 2002Q1-2006Q4 | 165 | 1366 | 2012Q1-2016Q4 | 189 | 2574 |
| 1992Q2-1997Q1 | 84 | 657 | 2002Q2-2007Q1 | 167 | 1380 | 2012Q2-2017Q1 | 187 | 2625 |
| 1992Q3-1997Q2 | 85 | 707 | 2002Q3-2007Q2 | 167 | 1392 | 2012Q3-2017Q2 | 189 | 2764 |
| 1992Q4-1997Q3 | 86 | 730 | 2002Q4-2007Q3 | 169 | 1433 | 2012Q4-2017Q3 | 186 | 2790 |
| 1993Q1-1997Q4 | 87 | 751 | 2003Q1-2007Q4 | 170 | 1450 | 2013Q1-2017Q4 | 187 | 2838 |
| 1993Q2-1998Q1 | 85 | 734 | 2003Q2-2008Q1 | 171 | 1465 | 2013Q2-2018Q1 | 186 | 2859 |
| 1993Q3-1998Q2 | 84 | 760 | 2003Q3-2008Q2 | 172 | 1473 | 2013Q3-2018Q2 | 182 | 2925 |
| 1993Q4-1998Q3 | 87 | 780 | 2003Q4-2008Q3 | 173 | 1474 | 2013Q4-2018Q3 | 181 | 2945 |
| 1994Q1-1998Q4 | 91 | 811 | 2004Q1-2008Q4 | 171 | 1457 | 2014Q1-2018Q4 | 181 | 2999 |
| 1994Q2-1999Q1 | 98 | 861 | 2004Q2-2009Q1 | 171 | 1443 | 2014Q2-2019Q1 | 180 | 3059 |
| 1994Q3-1999Q2 | 106 | 973 | 2004Q3-2009Q2 | 171 | 1461 | 2014Q3-2019Q2 | 179 | 3092 |
| 1994Q4-1999Q3 | 112 | 1031 | 2004Q4-2009Q3 | 168 | 1466 | 2014Q4-2019Q3 | 178 | 3128 |
| 1995Q1-1999Q4 | 116 | 1095 | 2005Q1-2009Q4 | 165 | 1476 | 2015Q1-2019Q4 | 175 | 3164 |
| 1995Q2-2000Q1 | 118 | 1121 | 2005Q2-2010Q1 | 160 | 1458 | 2015Q2-2020Q1 | 174 | 3149 |
| 1995Q3-2000Q2 | 122 | 1187 | 2005Q3-2010Q2 | 157 | 1495 | 2015Q3-2020Q2 | 171 | 3111 |
| 1995Q4-2000Q3 | 121 | 1212 | 2005Q4-2010Q3 | 157 | 1494 | 2015Q4-2020Q3 | 170 | 3072 |
| 1996Q1-2000Q4 | 122 | 1223 | 2006Q1-2010Q4 | 158 | 1563 | 2016Q1-2020Q4 | 166 | 3028 |
| 1996Q2-2001Q1 | 124 | 1222 | 2006Q2-2011Q1 | 156 | 1576 | | | |
| 1996Q3-2001Q2 | 126 | 1233 | 2006Q3-2011Q2 | 150 | 1609 | | | |
| 1996Q4-2001Q3 | 127 | 1231 | 2006Q4-2011Q3 | 149 | 1615 | | | |
| 1997Q1-2001Q4 | 126 | 1242 | 2007Q1-2011Q4 | 145 | 1619 | | | |
| 1997Q2-2002Q1 | 124 | 1219 | 2007Q2-2012Q1 | 150 | 1618 | | | |
| 1997Q3-2002Q2 | 126 | 1202 | 2007Q3-2012Q2 | 154 | 1689 | | | |
| 1997Q4-2002Q3 | 135 | 1194 | 2007Q4-2012Q3 | 154 | 1725 | | | |
| 1998Q1-2002Q4 | 132 | 1210 | 2008Q1-2012Q4 | 156 | 1757 | | | |
| 1998Q2-2003Q1 | 140 | 1235 | 2008Q2-2013Q1 | 157 | 1789 | | | |
| 1998Q3-2003Q2 | 142 | 1244 | 2008Q3-2013Q2 | 155 | 1786 | | | |
| 1998Q4-2003Q3 | 148 | 1264 | 2008Q4-2013Q3 | 159 | 1838 | | | |
| 1999Q1-2003Q4 | 148 | 1268 | 2009Q1-2013Q4 | 160 | 1903 | | | |
| 1999Q2-2004Q1 | 151 | 1270 | 2009Q2-2014Q1 | 160 | 1952 | | | |
| 1999Q3-2004Q2 | 152 | 1247 | 2009Q3-2014Q2 | 162 | 2009 | | | |
| 1999Q4-2004Q3 | 156 | 1208 | 2009Q4-2014Q3 | 168 | 2073 | | | |
| 2000Q1-2004Q4 | 158 | 1257 | 2010Q1-2014Q4 | 171 | 2095 | | | |
| 2000Q2-2005Q1 | 161 | 1244 | 2010Q2-2015Q1 | 173 | 2127 | | | |
| 2000Q3-2005Q2 | 162 | 1289 | 2010Q3-2015Q2 | 174 | 2154 | | | |
| 2000Q4-2005Q3 | 164 | 1284 | 2010Q4-2015Q3 | 178 | 2240 | | | |

## IA.2 A Simplified Model Visualization

Figure IA.2.1: **CoLGNN One-Layer Graph Neural Network Visualization Example**

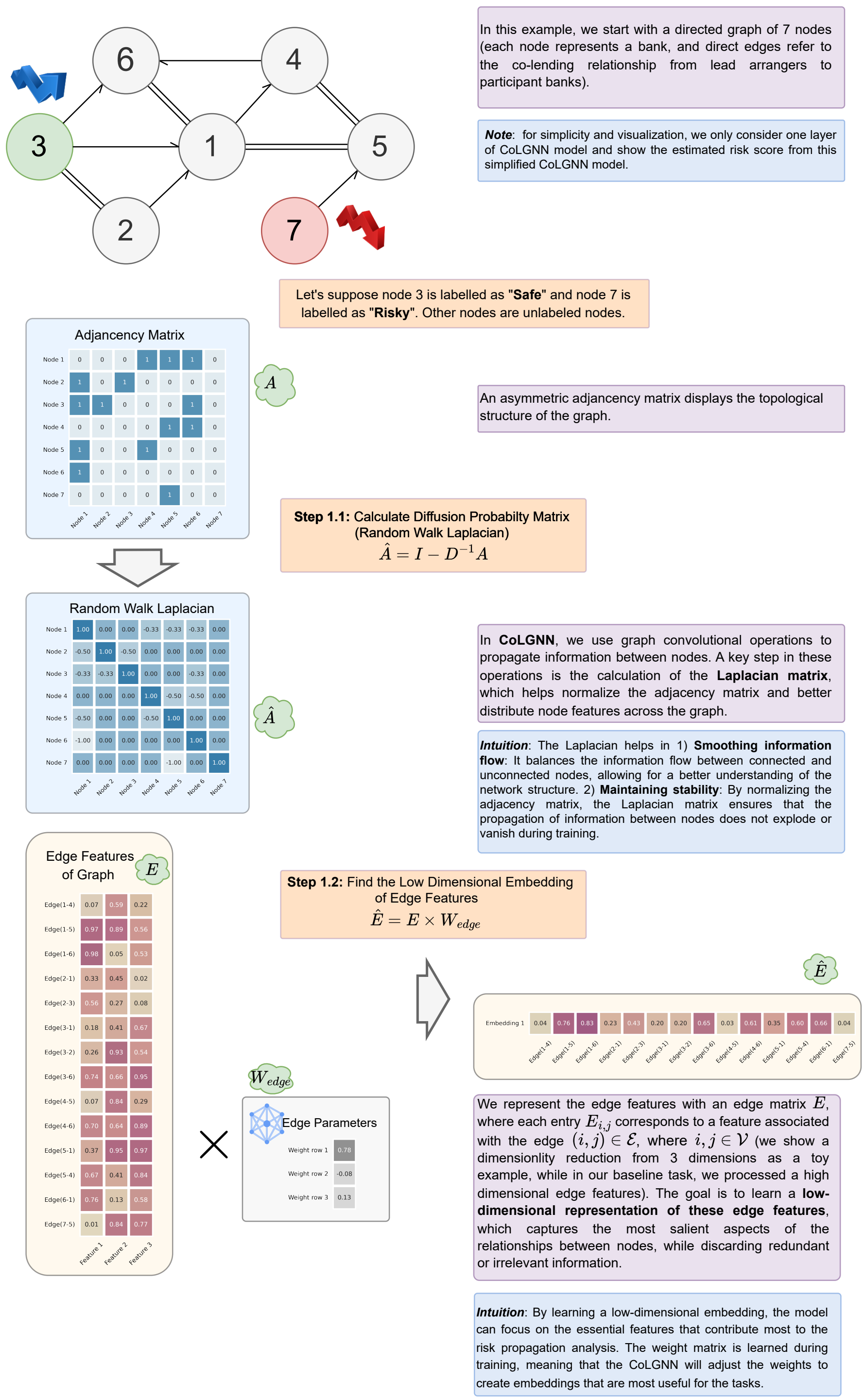
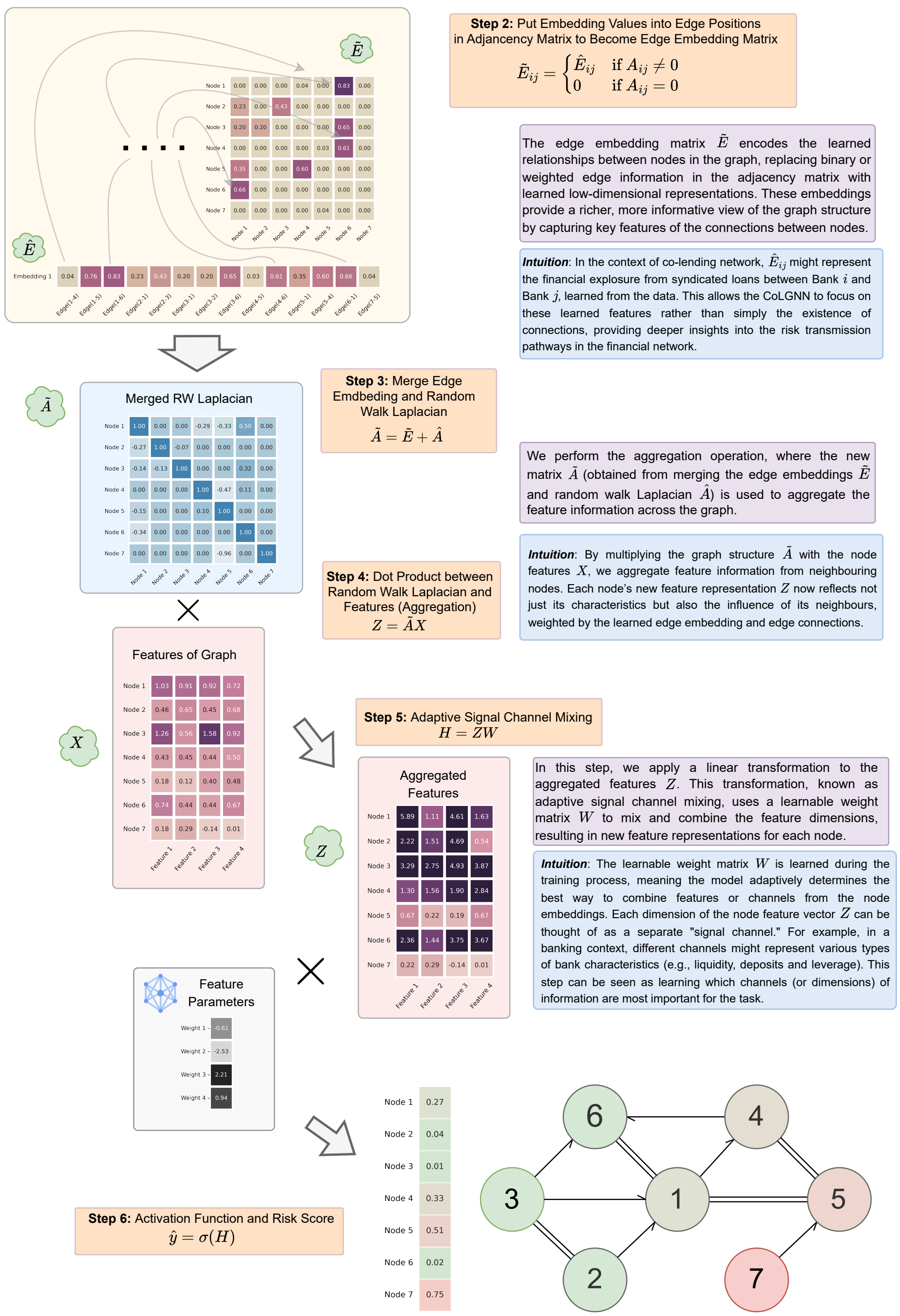


Figure IA.2.1: **Continued**



To provide insight on how the CoLGNN framework captures risk spillover within the co-lending network, we present a simplified illustration of how a single GNN layer operates within the CoLGNN framework in a co-lending network with seven banks, as shown in Figure IA.2.1. The process begins with constructing the adjacency matrix, where the directed edges represent the potential pathways for risk transmission from lead arrangers to participant banks. The random walk Laplacian matrix $\hat{A}$ is then calculated to smooth the information and maintain stability. This step helps balance information flow between connected and unconnected nodes, allowing for a better understanding of the network structure.

Next, edge features are represented with an edge matrix, where each entry corresponds to a feature associated with the edge between two banks.[28] By learning a low-dimensional embedding, the model can focus on the essential features that contribute the most to risk propagation analysis. These learned embeddings form the edge embedding matrix $\tilde{E}$, which is then incorporated into the adjacency matrix. This creates a combined matrix $\tilde{A} = \tilde{E} + \hat{A}$ that encodes the learned relationships between nodes in the graph, providing a richer, more informative view of the co-lending network structure.

In the feature aggregation step, the GNN layer uses the combined matrix $\tilde{A}$ to aggregate node features $X$, resulting in a new feature matrix $Z = \tilde{A}X$. This process allows information from neighboring nodes to influence each bank's feature representation, capturing the essence of risk spillover in the network. Following aggregation, adaptive signal channel mixing is applied, where a learnable weight matrix $W$ further transforms the aggregated features, yielding $H = ZW$. This step can be understood as learning which "channels" or dimensions of information (e.g., liquidity, deposits, leverage) are most important for the risk assessment task.

Lastly, the first dimension of $H$ is passed through a sigmoid activation function, producing a continuous risk score $\hat{y} = \sigma(H_1)$ for each bank, ranging from 0 (least risky) to 1 (most risky). This simplified visualization demonstrates how a basic GNN layer can transform co-lending

[28] We show a dimensionlity reduction from 3 dimensions as a toy example, while in our baseline task, we processed a high dimensional edge features.

network information into risk scores. While our CoLGNN model employs multiple such layers and additional specialized components, this example provides an intuitive understanding of the fundamental mechanics underlying graph neural networks in the context of banking risk assessment.

## IA.3 Network Centrality in the Bank Co-Lending Network

Traditional measures such as eigenvector centrality, closeness centrality, and eigenvector centrality do not account for the directionality of relationships, which is crucial in co-lending networks. To address this, we calculate both in-degree and out-degree versions of centrality measures. In the syndicated loan market, high **in-degree centrality** indicates a bank that frequently participates in loans arranged by others. Such banks may be more exposed to risks originating from multiple lead arrangers but may also benefit from diversification. High **out-degree centrality** suggests a bank that often acts as a lead arranger, initiating and structuring syndicated loans. These banks may have more control over loan terms but also bear greater responsibility for due diligence and potentially higher reputational risk.

Table IA.3.1: **Summary Statistics of Bank-level Network Centrality**

Table IA.3.1 presents the summary statistics bank-level directed network centrality at year-quarter level for each bank.

| | Observations | Mean | $10^{th}$ Percentile | Median | $90^{th}$ Percentile | Standard Deviation |
|---|---|---|---|---|---|---|
| ***Bank-level Samples*** | | | | | | |
| In-degree centrality | 11688 | 0.080 | 0.006 | 0.044 | 0.215 | 0.082 |
| Out-degree centrality | 11688 | 0.084 | 0.000 | 0.000 | 0.335 | 0.153 |
| Closeness centrality (in) | 11688 | 0.209 | 0.155 | 0.201 | 0.276 | 0.047 |
| Closeness centrality (out) | 11688 | 0.228 | 0.000 | 0.000 | 0.594 | 0.268 |
| Betweenness centrality (in) | 11688 | 0.003 | 0.000 | 0.000 | 0.009 | 0.009 |
| Betweenness centrality (in) | 11688 | 0.003 | 0.000 | 0.000 | 0.009 | 0.009 |
| Eigenvector centrality (in) | 11688 | 0.065 | 0.007 | 0.043 | 0.160 | 0.060 |
| Eigenvector centrality (out) | 11688 | 0.048 | 0.000 | 0.000 | 0.183 | 0.074 |
| Katz centrality | 11688 | 0.032 | -0.079 | 0.058 | 0.105 | 0.073 |
| Katz centrality (reverse) | 11688 | 0.016 | -0.126 | 0.053 | 0.072 | 0.082 |
| PageRank centrality | 11688 | 0.007 | 0.003 | 0.005 | 0.015 | 0.005 |
| PageRank centrality (reverse) | 11688 | 0.008 | 0.001 | 0.001 | 0.025 | 0.011 |

We estimate the network centrality using the `NetworkX` package for directed graphs[29].

[29]https://networkx.org/

**In-Degree Centrality** $C_D^{in}(v_i)$ measures the number of incoming edges to a node, normalized by the maximum possible in-degree. **Out-Degree Centrality** $C_D^{out}(v_i)$ measures the number of outgoing edges from a node, normalized by the maximum possible out-degree.

$$C_D^{in}(v_i) = \frac{\deg^{in}(v_i)}{|\mathcal{V}| - 1} \tag{IA.3.1}$$

$$C_D^{out}(v_i) = \frac{\deg^{out}(v_i)}{|\mathcal{V}| - 1} \tag{IA.3.2}$$

**Closeness Centrality (In)** $C_C^{in}(v_i)$ assesses how close a node is to all other nodes based on incoming paths. **Closeness Centrality (Out)** $C_C^{out}(v_i)$ assesses how close a node is to all other nodes based on outgoing paths.

$$C_C^{in}(v_i) = \frac{|\mathcal{V}| - 1}{\sum_{v_j \in \mathcal{V} \setminus \{v_i\}} d(v_j, v_i)} \tag{IA.3.3}$$

$$C_C^{out}(v_i) = \frac{|\mathcal{V}| - 1}{\sum_{v_j \in \mathcal{V} \setminus \{v_i\}} d(v_i, v_j)} \tag{IA.3.4}$$

**Eigenvector Centrality (In)** $C_{EV}^{in}(v_i)$ evaluates a node's influence based on the influence of its incoming neighbors. **Eigenvector Centrality (Out)** $C_{EV}^{out}(v_i)$ evaluates a node's influence based on the influence of its outgoing neighbors.

$$C_{EV}^{in}(v_i) = \frac{1}{\lambda} \sum_{v_j \in \mathcal{N}_{in}(v_i)} C_{EV}^{in}(v_j) \tag{IA.3.5}$$

$$C_{EV}^{out}(v_i) = \frac{1}{\lambda} \sum_{v_j \in \mathcal{N}_{out}(v_i)} C_{EV}^{out}(v_j) \tag{IA.3.6}$$

**Betweenness Centrality (In)** $C_B^{in}(v_i)$ quantifies the number of times a node acts as a bridge along the shortest paths directed towards it. **Betweenness Centrality (Out)** $C_B^{out}(v_i)$ quantifies the number of times a node acts as a bridge along the shortest paths originating from it.

$$C_B^{in}(v_i) = \sum_{s \neq v_i \neq t} \frac{\sigma_{st}(v_i)}{\sigma_{st}} \tag{IA.3.7}$$

$$C_B^{out}(v_i) = \sum_{s \neq v_i \neq t} \frac{\sigma_{st}(v_i)}{\sigma_{st}} \tag{IA.3.8}$$

**Katz Centrality** $C_K^{in}(v_i)$ considers all incoming walks to a node, with longer walks exponentially damped by a factor $\alpha$. **Katz Centrality (reverse)** $C_K^{reverse}(v_i)$ considers all outgoing walks from a node, with longer walks exponentially damped by a factor $\alpha$.

$$C_K^{in}(v_i) = \sum_{k=1}^{\infty} \alpha^k \cdot (\mathbf{A}^k)_{ji} \tag{IA.3.9}$$

$$C_K^{reverse}(v_i) = \sum_{k=1}^{\infty} \alpha^k \cdot (\mathbf{A}^k)_{ik} \tag{IA.3.10}$$

**PageRank Centrality** $C_{PR}^{reverse}(v_i)$ measures the probability of arriving at a node through a random walk that follows outgoing edges, incorporating a damping factor $\alpha$.

$$C_{PR}(v_i) = \frac{1-\alpha}{|\mathcal{V}|} + \alpha \sum_{v_j \in \mathcal{N}_{in}(v_i)} \frac{C_{PR}(v_j)}{\deg^{out}(v_j)} \tag{IA.3.11}$$

$$C_{PR}^{reverse}(v_i) = \frac{1-\alpha}{|\mathcal{V}|} + \alpha \sum_{v_j \in \mathcal{N}_{out}(v_i)} \frac{C_{PR}^{reverse}(v_j)}{\deg^{in}(v_j)} \tag{IA.3.12}$$

**Notes:** $\mathcal{V}$ denotes the set of all nodes in the graph. $\deg^{in}(v_i)$ and $\deg^{out}(v_i)$ represent the in-degree and out-degree of node $v_i$, respectively. $d(v_j, v_i)$ is the shortest path distance from node $v_j$ to node $v_i$. $\mathcal{N}_{in}(v_i)$ and $\mathcal{N}_{out}(v_i)$ denote the sets of nodes with edges directed towards and away from node $v_i$, respectively. $\sigma_{st}$ is the total number of shortest paths from node $s$ to node $t$, and $\sigma_{st}(v_i)$ is the number of those paths that pass through node $v_i$. $\mathbf{A}$ is the adjacency matrix of the graph. $\lambda$ is the largest eigenvalue of the adjacency matrix $\mathbf{A}$. $\alpha$ is the damping factor, typically set to 0.85 in PageRank calculations.

We estimate the following regression similar as Eq 6 with an additional network centrality control variable.

$$LLP_{i,t+h} = \beta_1 CLN\ score_{i,t} + \beta_1 Centrality_{i,t} + \gamma X_{i,t} + \text{Fixed Effects} + \varepsilon_{i,t+h} \qquad \text{(IA.3.13)}$$

where $LLP_{i,t+h}$ is the loan loss provisions for bank $i$ at time $t+h$, $CLN\ score_{i,t}$ refers to the bank-level co-lending network risk score of bank $i$ at time $t$. $Centrality_{i,t}$ is one of the network centrality measures defined above $Centrality_{i,t} \in \{\ C_D^{in}(v_i),\ C_D^{out}(v_i), ...,\ C_{PR}(v_i),\ C_{PR}^{reverse}(v_i)\}$.

## IA.4 Additional figures and tables

Table IA.4.1: **Banks with High and Low Co-Lending Network Risk**

Table IA.4.1 presents the top and bottom six banks in our sample, categorized based on their co-lending network risks, as measured by the *CLN* score. This table specifically includes banks with a presence spanning more than an economic cycle, defined as 7 years or 28 quarters. Notably, two RSSDIDs became inactive after the end of our sample period: People's United Finance, Inc.(RSSDID: 3650152) and Umpqua Holdings Corporation(RSSDID: 2747644). People's United Finance was acquired and fully integrated by the third quarter of 2022. Umpqua Holdings Corporation was acquired by Columbia Banking System, Inc., with the merger concluding on March 1, 2023.

| RSSDID | Bank | Headquarter State | Average CLN score |
|---|---|---|---|
| *Top Seven* | | | |
| 3650152 | People's United Finance, Inc. (Inactive) | Connecticut | 0.786 |
| 2333663 | Berkshire Hills Bancorp, Inc. | Massachusetts | 0.767 |
| 2747644 | Umpqua Holdings Corporation (Inactive) | Oregon | 0.755 |
| 2132932 | New York Community Bancorp, Inc. | New York | 0.717 |
| 1562859 | Ally Financial Inc. | Michigan | 0.672 |
| 1098303 | Old National Bancorp | Indiana | 0.669 |
| 1078846 | Synovus Financial Corp. | Georgia | 0.653 |
| *Bottom Seven* | | | |
| 2461016 | Enterprise Bancorp, Inc. | Massachusetts | 0.265 |
| 1107205 | Amarillo National Bancorp, Inc. | Texas | 0.259 |
| 3635319 | Servisfirst Bancshares, Inc. | Alabama | 0.228 |
| 1208906 | Lakeland Financial Corporation | Indiana | 0.221 |
| 1399073 | Heartland Banccorp | Ohio | 0.211 |
| 1059715 | American National Corporation | Nebraska | 0.202 |
| 1862036 | Guaranty Bancshares, Inc. | Texas | 0.199 |

Table IA.4.2: **Bank Co-Lending Network Risk and Bank Loan Loss Provisions: Controlling for Lending Specialization**

Table IA.4.2 presents the h-quarter-ahead prediction results of bank-level *CLN*-score for bank loan loss provisions. The bank-level *CLN*-score and control variables are measured at time $t$, and the dependent variable is measured at time $t+h$. Panel A reports the results using all observations. Panel B shows the results using unlabeled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables in the Panel B and C. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Loan loss provisions** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All samples*** | | | | | | | | |
| *CLN* score | 0.103*** | 0.106*** | 0.106*** | 0.078*** | 0.096*** | 0.075*** | 0.069*** | 0.038** |
| | (7.801) | (8.425) | (7.677) | (5.494) | (6.042) | (4.613) | (4.113) | (2.002) |
| Size | 0.029*** | 0.029*** | 0.029*** | 0.032*** | 0.029** | 0.029** | 0.030** | 0.033*** |
| | (3.161) | (3.088) | (2.908) | (2.979) | (2.586) | (2.445) | (2.525) | (2.659) |
| Equity capital | 0.211 | 0.171 | 0.100 | 0.439 | 0.318 | 0.353 | 0.481 | 0.720 |
| | (0.420) | (0.378) | (0.209) | (0.870) | (0.649) | (0.710) | (0.954) | (1.414) |
| Deposits | -0.038 | -0.050 | -0.082 | -0.097 | -0.125 | -0.157 | -0.146 | -0.140 |
| | (-0.341) | (-0.446) | (-0.712) | (-0.776) | (-0.959) | (-1.183) | (-1.103) | (-1.016) |
| ROA | -0.139*** | -0.098*** | -0.074*** | -0.108*** | -0.057** | -0.051* | -0.056 | -0.072* |
| | (-4.658) | (-3.566) | (-2.909) | (-2.857) | (-2.246) | (-1.665) | (-1.460) | (-1.669) |
| Loan size | -0.481*** | -0.370*** | -0.255* | -0.105 | -0.021 | 0.054 | 0.115 | 0.196 |
| | (-3.637) | (-2.818) | (-1.836) | (-0.721) | (-0.140) | (0.342) | (0.728) | (1.220) |
| Loan growth | 0.000 | 0.000 | -0.000 | -0.001 | -0.000 | -0.000 | 0.001* | 0.002** |
| | (0.377) | (1.648) | (-0.752) | (-1.378) | (-0.303) | (-0.097) | (1.676) | (2.196) |
| Loan loss allowance | 0.568*** | 0.521*** | 0.469*** | 0.393*** | 0.344*** | 0.311*** | 0.270*** | 0.228*** |
| | (10.573) | (8.832) | (7.343) | (6.165) | (5.257) | (4.446) | (3.869) | (3.257) |
| Liquidity | -0.261* | -0.252* | -0.266* | -0.243* | -0.241* | -0.215 | -0.194 | -0.170 |
| | (-1.795) | (-1.773) | (-1.853) | (-1.684) | (-1.719) | (-1.523) | (-1.330) | (-1.138) |
| Reputation | 1.163 | 1.420 | 1.644* | 1.916* | 2.436** | 2.602** | 2.775** | 2.946** |
| | (1.358) | (1.552) | (1.695) | (1.800) | (2.071) | (2.124) | (2.175) | (2.237) |
| Specialization in syndicated loan | -0.333*** | -0.351*** | -0.371*** | -0.473*** | -4.694* | -4.415 | -4.176 | -3.817 |
| | (-11.904) | (-13.146) | (-12.438) | (-10.311) | (-1.700) | (-1.538) | (-1.381) | (-1.198) |
| Specialization in industry | 0.013 | -0.013 | -0.017 | -0.013 | -0.012 | -0.022 | -0.036* | -0.033 |
| | (0.769) | (-0.711) | (-0.916) | (-0.712) | (-0.686) | (-1.140) | (-1.731) | (-1.647) |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11308 | 10966 | 10631 | 10303 | 9980 | 9666 | 9365 | 9073 |
| Adjusted $R^2$ | 0.569 | 0.522 | 0.491 | 0.474 | 0.452 | 0.445 | 0.444 | 0.444 |
| | | | | | | | | |
| *CLN* score | 0.080*** | 0.081*** | 0.076*** | 0.055*** | 0.083*** | 0.061*** | 0.055*** | 0.025 |
| | (4.840) | (4.771) | (3.855) | (3.079) | (3.881) | (3.036) | (2.632) | (1.059) |
| Controls and Lending Specializations | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6907 | 6699 | 6495 | 6294 | 6098 | 5901 | 5702 | 5518 |
| Adjusted $R^2$ | 0.511 | 0.477 | 0.456 | 0.448 | 0.424 | 0.429 | 0.430 | 0.435 |
| | | | | | | | | |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.098*** | 0.125*** | 0.136*** | 0.103*** | 0.145*** | 0.121*** | 0.124*** | 0.074** |
| | (3.778) | (4.973) | (4.409) | (3.875) | (4.415) | (3.947) | (3.984) | (1.982) |
| Controls and Lending Specializations | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3882 | 3742 | 3607 | 3476 | 3348 | 3223 | 3103 | 2988 |
| Adjusted $R^2$ | 0.496 | 0.469 | 0.447 | 0.437 | 0.418 | 0.417 | 0.412 | 0.411 |

Table IA.4.3: **Controlling for Bank Fixed Effects**

Table IA.4.3 presents the h-quarter-ahead prediction results of bank-level *CLN*-score with bank fixed effects. The bank-level *CLN*-score and control variables are measured at time $t$, and the dependent variable is measured at time $t+h$. Panel A reports the results using all observations. Panel B shows the results using unlabeled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. For simplicity, we do not report the coefficients estimates of the control variables in the Panel B and C. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1)<br>h=1 | (2)<br>h=2 | (3)<br>h=3 | (4)<br>h=4 | (5)<br>h=5 | (6)<br>h=6 | (7)<br>h=7 | (8)<br>h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Panel A: All samples*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.080*** | 0.080*** | 0.079*** | 0.050*** | 0.072*** | 0.052*** | 0.049*** | 0.021 |
| | (7.445) | (7.719) | (6.509) | (4.113) | (5.273) | (3.939) | (3.507) | (1.323) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11303 | 10963 | 10628 | 10297 | 9973 | 9658 | 9360 | 9065 |
| Adjusted $R^2$ | 0.655 | 0.616 | 0.597 | 0.586 | 0.575 | 0.573 | 0.575 | 0.577 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.126*** | 0.148*** | 0.159*** | 0.154*** | 0.156*** | 0.146*** | 0.138*** | 0.143*** |
| | (3.925) | (4.345) | (4.405) | (4.095) | (3.867) | (3.683) | (3.438) | (3.273) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 10263 | 10022 | 9782 | 9496 | 9262 | 9028 | 8806 | 8576 |
| Adjusted $R^2$ | 0.747 | 0.731 | 0.715 | 0.701 | 0.695 | 0.689 | 0.690 | 0.692 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.174*** | -0.151*** | -0.145*** | -0.126*** | -0.123*** | -0.104*** | -0.083*** | -0.068*** |
| | (-6.962) | (-6.310) | (-6.869) | (-5.352) | (-6.276) | (-4.710) | (-3.928) | (-3.318) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11305 | 10965 | 10630 | 10300 | 9975 | 9658 | 9360 | 9065 |
| Adjusted $R^2$ | 0.439 | 0.417 | 0.413 | 0.409 | 0.413 | 0.419 | 0.426 | 0.448 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.066*** | 0.063*** | 0.055*** | 0.032** | 0.066*** | 0.043** | 0.037** | 0.008 |
| | (4.241) | (3.950) | (2.973) | (2.005) | (3.447) | (2.515) | (2.145) | (0.371) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6894 | 6689 | 6488 | 6285 | 6089 | 5894 | 5695 | 5507 |
| Adjusted $R^2$ | 0.637 | 0.607 | 0.594 | 0.582 | 0.568 | 0.577 | 0.576 | 0.577 |

Table IA.4.3: **Continued**

| **Dependent variable (t+h):** | (1)<br>h=1 | (2)<br>h=2 | (3)<br>h=3 | (4)<br>h=4 | (5)<br>h=5 | (6)<br>h=6 | (7)<br>h=7 | (8)<br>h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.103** | 0.119** | 0.136*** | 0.143*** | 0.161*** | 0.146*** | 0.127** | 0.135** |
| | (2.282) | (2.398) | (2.606) | (2.747) | (3.022) | (2.876) | (2.505) | (2.347) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6321 | 6176 | 6024 | 5854 | 5702 | 5548 | 5392 | 5239 |
| Adjusted $R^2$ | 0.747 | 0.733 | 0.723 | 0.710 | 0.705 | 0.699 | 0.695 | 0.692 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.118*** | -0.101*** | -0.102*** | -0.092** | -0.087*** | -0.065*** | -0.062** | -0.040 |
| | (-5.020) | (-3.992) | (-3.344) | (-2.576) | (-3.041) | (-2.858) | (-2.303) | (-1.440) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6896 | 6691 | 6490 | 6288 | 6091 | 5894 | 5695 | 5507 |
| Adjusted $R^2$ | 0.466 | 0.425 | 0.431 | 0.412 | 0.415 | 0.434 | 0.460 | 0.539 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.083*** | 0.103*** | 0.113*** | 0.071*** | 0.115*** | 0.091*** | 0.095*** | 0.041 |
| | (3.513) | (4.345) | (3.917) | (3.034) | (4.031) | (3.521) | (3.668) | (1.221) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3875 | 3737 | 3603 | 3472 | 3342 | 3216 | 3101 | 2982 |
| Adjusted $R^2$ | 0.632 | 0.608 | 0.595 | 0.581 | 0.568 | 0.571 | 0.570 | 0.563 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.176** | 0.206*** | 0.238*** | 0.248*** | 0.278*** | 0.244*** | 0.217** | 0.222** |
| | (2.549) | (2.654) | (2.945) | (3.028) | (3.279) | (2.969) | (2.613) | (2.278) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3617 | 3509 | 3405 | 3294 | 3189 | 3085 | 2988 | 2883 |
| Adjusted $R^2$ | 0.765 | 0.751 | 0.742 | 0.727 | 0.716 | 0.705 | 0.704 | 0.703 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.142*** | -0.113*** | -0.155*** | -0.148*** | -0.164*** | -0.113*** | -0.099** | -0.064 |
| | (-4.211) | (-3.646) | (-3.577) | (-2.621) | (-3.758) | (-3.433) | (-2.604) | (-1.372) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3877 | 3739 | 3605 | 3475 | 3344 | 3216 | 3101 | 2982 |
| Adjusted $R^2$ | 0.463 | 0.452 | 0.451 | 0.458 | 0.469 | 0.481 | 0.491 | 0.552 |

## Table IA.4.4: **Subperiod Analysis**

Table IA.4.4 presents the h-quarter-ahead prediction results of bank-level *CLN*-score on different sample Periods. We equally split the sample to two halves. The bank-level *CLN*-score and control variables are measured at time $t$, and the dependent variable is measured at time $t+h$. Panel A reports the results using all observations. Panel B shows the results using unlabeled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. For simplicity, we do not report the coefficients estimates of the control variables in the Panel B and C. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1) h=1 | (2) h=2 | (3) h=3 | (4) h=4 | (5) h=5 | (6) h=6 | (7) h=7 | (8) h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Panel A: All samples*** | | | | | | | | |
| ***First-half Subperiod*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.105*** | 0.121*** | 0.149*** | 0.105*** | 0.120*** | 0.100*** | 0.115*** | 0.073** |
| | (7.162) | (7.876) | (7.411) | (5.169) | (4.924) | (3.783) | (3.750) | (2.033) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 5416 | 5308 | 5204 | 5101 | 4994 | 4891 | 4796 | 4706 |
| Adjusted $R^2$ | 0.468 | 0.476 | 0.510 | 0.484 | 0.464 | 0.458 | 0.455 | 0.448 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.003*** | 0.003*** | 0.004*** | 0.004*** | 0.004*** | 0.004*** | 0.003*** | 0.003*** |
| | (6.642) | (6.439) | (5.752) | (5.171) | (4.727) | (3.865) | (3.284) | (2.700) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 5162 | 5064 | 4966 | 4868 | 4767 | 4668 | 4635 | 4551 |
| Adjusted $R^2$ | 0.459 | 0.438 | 0.408 | 0.384 | 0.367 | 0.352 | 0.337 | 0.325 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.248*** | -0.257*** | -0.281*** | -0.253*** | -0.244*** | -0.233*** | -0.226*** | -0.180*** |
| | (-6.971) | (-7.232) | (-7.872) | (-6.046) | (-6.227) | (-5.820) | (-5.042) | (-3.584) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 5418 | 5310 | 5206 | 5103 | 4996 | 4892 | 4796 | 4706 |
| Adjusted $R^2$ | 0.305 | 0.276 | 0.260 | 0.243 | 0.232 | 0.226 | 0.219 | 0.210 |
| ***Second-half Subperiod*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.169*** | 0.131*** | 0.080*** | 0.076*** | 0.076*** | 0.044** | 0.012 | 0.001 |
| | (7.754) | (6.278) | (3.835) | (3.819) | (3.828) | (2.512) | (0.802) | (0.100) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 5892 | 5658 | 5427 | 5202 | 4986 | 4775 | 4569 | 4367 |
| Adjusted $R^2$ | 0.579 | 0.528 | 0.430 | 0.416 | 0.397 | 0.363 | 0.268 | 0.273 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.005*** | 0.004*** | 0.004*** | 0.003*** | 0.003*** | 0.002** | 0.002** | 0.002* |
| | (4.055) | (3.715) | (3.456) | (3.095) | (2.771) | (2.413) | (2.153) | (1.940) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 4515 | 4289 | 4070 | 3857 | 3655 | 3466 | 3282 | 3101 |
| Adjusted $R^2$ | 0.333 | 0.301 | 0.274 | 0.240 | 0.211 | 0.184 | 0.163 | 0.140 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.280*** | -0.220*** | -0.172*** | -0.156*** | -0.149*** | -0.118*** | -0.062*** | -0.061*** |
| | (-8.994) | (-7.723) | (-6.777) | (-5.902) | (-6.341) | (-4.510) | (-3.054) | (-3.024) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 5892 | 5658 | 5427 | 5202 | 4986 | 4775 | 4569 | 4367 |
| Adjusted $R^2$ | 0.316 | 0.290 | 0.313 | 0.296 | 0.297 | 0.302 | 0.328 | 0.326 |

Table IA.4.4: **Continued**

| **Dependent variable (t+h):** | (1) h=1 | (2) h=2 | (3) h=3 | (4) h=4 | (5) h=5 | (6) h=6 | (7) h=7 | (8) h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| ***First-half Subsamples*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.068*** | 0.072*** | 0.081*** | 0.047* | 0.073** | 0.055 | 0.070* | 0.036 |
| | (3.717) | (3.971) | (2.821) | (1.829) | (2.162) | (1.622) | (1.801) | (0.752) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3302 | 3231 | 3164 | 3099 | 3030 | 2963 | 2891 | 2829 |
| Adjusted $R^2$ | 0.436 | 0.463 | 0.488 | 0.466 | 0.436 | 0.452 | 0.450 | 0.453 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.002*** | 0.002*** | 0.003*** | 0.003*** | 0.003*** | 0.003** | 0.002* | 0.002 |
| | (3.860) | (3.485) | (3.188) | (3.076) | (2.864) | (2.216) | (1.699) | (1.161) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3164 | 3097 | 3036 | 2970 | 2905 | 2841 | 2802 | 2744 |
| Adjusted $R^2$ | 0.413 | 0.391 | 0.366 | 0.350 | 0.336 | 0.325 | 0.313 | 0.297 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.184*** | -0.188*** | -0.207*** | -0.195*** | -0.179*** | -0.156*** | -0.176*** | -0.126* |
| | (-5.548) | (-4.822) | (-3.940) | (-3.145) | (-3.140) | (-3.080) | (-2.722) | (-1.672) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3304 | 3233 | 3166 | 3101 | 3032 | 2964 | 2891 | 2829 |
| Adjusted $R^2$ | 0.340 | 0.272 | 0.254 | 0.227 | 0.209 | 0.216 | 0.220 | 0.222 |
| ***Second-half Subsamples*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.143*** | 0.115*** | 0.074*** | 0.070*** | 0.084*** | 0.047** | 0.015 | -0.005 |
| | (4.897) | (3.900) | (2.778) | (2.802) | (3.477) | (2.218) | (0.733) | (-0.279) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3605 | 3468 | 3331 | 3195 | 3068 | 2938 | 2811 | 2689 |
| Adjusted $R^2$ | 0.536 | 0.482 | 0.401 | 0.398 | 0.385 | 0.343 | 0.271 | 0.282 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.003*** | 0.003*** | 0.003*** | 0.003*** | 0.003*** | 0.003** | 0.002* | 0.002 |
| | (3.925) | (3.510) | (3.146) | (3.020) | (2.715) | (2.162) | (1.699) | (1.159) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3242 | 3172 | 3111 | 3045 | 2978 | 2913 | 2874 | 2813 |
| Adjusted $R^2$ | 0.430 | 0.405 | 0.379 | 0.360 | 0.344 | 0.329 | 0.315 | 0.297 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.218*** | -0.162*** | -0.126*** | -0.099*** | -0.096*** | -0.076** | -0.034 | -0.022 |
| | (-5.334) | (-4.328) | (-3.967) | (-3.268) | (-3.698) | (-2.341) | (-1.173) | (-0.786) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3605 | 3468 | 3331 | 3195 | 3068 | 2938 | 2811 | 2689 |
| Adjusted $R^2$ | 0.265 | 0.252 | 0.295 | 0.294 | 0.295 | 0.297 | 0.311 | 0.307 |

Table IA.4.4: **Continued**

| **Dependent variable (t+h):** | (1) h=1 | (2) h=2 | (3) h=3 | (4) h=4 | (5) h=5 | (6) h=6 | (7) h=7 | (8) h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| ***First-half Subsamples*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.101*** | 0.129*** | 0.164*** | 0.103*** | 0.143*** | 0.124** | 0.163*** | 0.086 |
| | (3.218) | (4.406) | (3.586) | (2.810) | (2.845) | (2.517) | (3.010) | (1.131) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1852 | 1798 | 1746 | 1696 | 1647 | 1598 | 1552 | 1510 |
| Adjusted $R^2$ | 0.443 | 0.460 | 0.481 | 0.468 | 0.445 | 0.454 | 0.448 | 0.438 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.003*** | 0.004*** | 0.004*** | 0.005*** | 0.005*** | 0.005** | 0.004* | 0.003 |
| | (3.553) | (3.461) | (3.308) | (3.162) | (3.125) | (2.278) | (1.850) | (1.272) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1795 | 1744 | 1696 | 1648 | 1600 | 1551 | 1520 | 1480 |
| Adjusted $R^2$ | 0.440 | 0.412 | 0.378 | 0.354 | 0.329 | 0.309 | 0.287 | 0.263 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.257*** | -0.249*** | -0.309*** | -0.312*** | -0.317*** | -0.256*** | -0.237** | -0.163 |
| | (-4.475) | (-4.479) | (-3.712) | (-2.926) | (-3.183) | (-2.905) | (-2.238) | (-1.262) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1854 | 1800 | 1748 | 1698 | 1649 | 1599 | 1552 | 1510 |
| Adjusted $R^2$ | 0.317 | 0.282 | 0.266 | 0.256 | 0.241 | 0.231 | 0.207 | 0.183 |
| ***Second-half Subsamples*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.158*** | 0.150*** | 0.109** | 0.108** | 0.125*** | 0.074** | 0.035 | 0.011 |
| | (3.423) | (3.280) | (2.538) | (2.594) | (3.116) | (2.074) | (0.996) | (0.374) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2030 | 1944 | 1861 | 1780 | 1701 | 1625 | 1551 | 1478 |
| Adjusted $R^2$ | 0.515 | 0.470 | 0.400 | 0.386 | 0.373 | 0.345 | 0.296 | 0.301 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.007*** | 0.007*** | 0.006*** | 0.006*** | 0.005*** | 0.004** | 0.004** | 0.004** |
| | (3.498) | (3.121) | (3.045) | (2.977) | (2.702) | (2.401) | (2.212) | (2.052) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1583 | 1500 | 1423 | 1346 | 1271 | 1204 | 1139 | 1075 |
| Adjusted $R^2$ | 0.327 | 0.290 | 0.266 | 0.231 | 0.188 | 0.151 | 0.123 | 0.098 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.252*** | -0.182*** | -0.168*** | -0.143*** | -0.160*** | -0.112** | -0.062 | -0.028 |
| | (-4.242) | (-3.277) | (-3.355) | (-2.883) | (-3.772) | (-2.164) | (-1.509) | (-0.679) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2030 | 1944 | 1861 | 1780 | 1701 | 1625 | 1551 | 1478 |
| Adjusted $R^2$ | 0.254 | 0.237 | 0.269 | 0.258 | 0.267 | 0.272 | 0.295 | 0.302 |

Table IA.4.5: **Bank Co-Lending Network Risk and GFC**

Table IA.4.5 presents the h-quarter-ahead prediction results of bank-level *CLN*-score on GFC periods. The GFC dummy variable equals to 1 (0) if the year-quarter is within the mid of 2007 and the end of 2008. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1)<br>h=1 | (2)<br>h=2 | (3)<br>h=3 | (4)<br>h=4 | (5)<br>h=5 | (6)<br>h=6 | (7)<br>h=7 | (8)<br>h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Panel A: All samples*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.123*** | 0.113*** | 0.093*** | 0.050*** | 0.059*** | 0.045*** | 0.037** | 0.008 |
| | (8.363) | (7.834) | (6.491) | (3.779) | (3.899) | (2.695) | (2.272) | (0.479) |
| *CLN* score × GFC | 0.257*** | 0.256*** | 0.451*** | 0.750*** | 0.721*** | 0.567*** | 0.611*** | 0.607*** |
| | (2.998) | (2.983) | (4.018) | (5.177) | (5.659) | (5.503) | (5.135) | (4.084) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11308 | 10966 | 10631 | 10303 | 9980 | 9666 | 9365 | 9073 |
| Adjusted $R^2$ | 0.541 | 0.506 | 0.484 | 0.476 | 0.460 | 0.448 | 0.447 | 0.446 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.003*** | 0.003*** | 0.003*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** | 0.002** |
| | (4.885) | (4.658) | (4.400) | (4.368) | (3.807) | (3.383) | (2.940) | (2.400) |
| *CLN* score × GFC | 0.007*** | 0.011*** | 0.013*** | 0.015*** | 0.018*** | 0.017*** | 0.015*** | 0.014*** |
| | (3.203) | (4.047) | (4.125) | (4.149) | (4.273) | (4.302) | (3.814) | (3.801) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 9677 | 9353 | 9036 | 8725 | 8422 | 8134 | 7917 | 7652 |
| Adjusted $R^2$ | 0.403 | 0.378 | 0.352 | 0.328 | 0.312 | 0.294 | 0.280 | 0.270 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.232*** | -0.203*** | -0.182*** | -0.144*** | -0.151*** | -0.148*** | -0.107*** | -0.079*** |
| | (-10.402) | (-9.049) | (-8.134) | (-6.626) | (-7.025) | (-6.021) | (-5.133) | (-3.598) |
| *CLN* score × GFC | -0.639*** | -0.675*** | -0.844*** | -1.101*** | -0.823*** | -0.576*** | -0.717*** | -0.735*** |
| | (-3.015) | (-3.375) | (-5.413) | (-5.601) | (-4.076) | (-4.043) | (-4.348) | (-4.076) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11310 | 10968 | 10633 | 10305 | 9982 | 9667 | 9365 | 9073 |
| Adjusted $R^2$ | 0.316 | 0.284 | 0.274 | 0.267 | 0.251 | 0.240 | 0.240 | 0.235 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.096*** | 0.089*** | 0.067*** | 0.035** | 0.052*** | 0.037* | 0.025 | -0.010 |
| | (5.180) | (4.555) | (3.516) | (2.076) | (2.660) | (1.678) | (1.207) | (-0.477) |
| *CLN* score × GFC | 0.185 | 0.107 | 0.274** | 0.481*** | 0.514*** | 0.358*** | 0.460*** | 0.524** |
| | (1.626) | (1.077) | (2.072) | (2.627) | (2.885) | (2.801) | (2.762) | (2.538) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6907 | 6699 | 6495 | 6294 | 6098 | 5901 | 5702 | 5518 |
| Adjusted $R^2$ | 0.499 | 0.472 | 0.455 | 0.447 | 0.429 | 0.431 | 0.433 | 0.440 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.003*** | 0.003*** | 0.003*** | 0.002*** | 0.002*** | 0.002*** | 0.002** | 0.001 |
| | (3.895) | (3.532) | (3.530) | (3.573) | (3.097) | (2.702) | (2.213) | (1.563) |
| *CLN* score × GFC | 0.006* | 0.009** | 0.010** | 0.012** | 0.016*** | 0.015*** | 0.014** | 0.013** |
| | (1.847) | (2.442) | (2.325) | (2.448) | (2.725) | (2.855) | (2.533) | (2.566) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 5944 | 5745 | 5557 | 5362 | 5171 | 4985 | 4831 | 4665 |
| Adjusted $R^2$ | 0.369 | 0.342 | 0.320 | 0.297 | 0.278 | 0.263 | 0.255 | 0.242 |

Table IA.4.5: **Continued**

| **Dependent variable (t+h):** | (1) h=1 | (2) h=2 | (3) h=3 | (4) h=4 | (5) h=5 | (6) h=6 | (7) h=7 | (8) h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.186*** | -0.154*** | -0.131*** | -0.099*** | -0.110*** | -0.108*** | -0.080*** | -0.040 |
| | (-6.361) | (-5.229) | (-4.248) | (-3.346) | (-3.942) | (-3.700) | (-2.885) | (-1.304) |
| *CLN* score × GFC | -0.410*** | -0.533*** | -0.779*** | -0.974*** | -0.632*** | -0.362** | -0.594*** | -0.651*** |
| | (-3.750) | (-2.887) | (-3.482) | (-3.510) | (-2.596) | (-2.298) | (-2.714) | (-2.606) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6909 | 6701 | 6497 | 6296 | 6100 | 5902 | 5702 | 5518 |
| Adjusted $R^2$ | 0.300 | 0.256 | 0.255 | 0.241 | 0.223 | 0.222 | 0.230 | 0.237 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.112*** | 0.126*** | 0.105*** | 0.050* | 0.082*** | 0.067* | 0.050 | 0.001 |
| | (3.757) | (4.136) | (3.343) | (1.865) | (2.720) | (1.914) | (1.546) | (0.037) |
| *CLN* score × GFC | 0.282 | 0.227 | 0.548*** | 0.919*** | 0.867*** | 0.634*** | 0.935*** | 0.877*** |
| | (1.642) | (1.482) | (2.951) | (3.428) | (3.747) | (3.549) | (4.827) | (2.947) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3882 | 3742 | 3607 | 3476 | 3348 | 3223 | 3103 | 2988 |
| Adjusted $R^2$ | 0.487 | 0.465 | 0.451 | 0.449 | 0.432 | 0.424 | 0.427 | 0.424 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.005*** | 0.004*** | 0.004*** | 0.004*** | 0.003*** | 0.003*** | 0.003** | 0.002* |
| | (3.846) | (3.558) | (3.533) | (3.616) | (3.393) | (2.940) | (2.567) | (1.704) |
| *CLN* score × GFC | 0.010* | 0.016** | 0.019*** | 0.020** | 0.026*** | 0.024*** | 0.023** | 0.023** |
| | (1.968) | (2.603) | (2.641) | (2.569) | (2.825) | (2.744) | (2.379) | (2.545) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3378 | 3244 | 3119 | 2994 | 2871 | 2755 | 2659 | 2555 |
| Adjusted $R^2$ | 0.385 | 0.356 | 0.330 | 0.304 | 0.282 | 0.256 | 0.237 | 0.219 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.231*** | -0.194*** | -0.181*** | -0.154*** | -0.185*** | -0.170*** | -0.108** | -0.053 |
| | (-5.250) | (-4.367) | (-3.816) | (-3.282) | (-4.110) | (-3.517) | (-2.557) | (-1.096) |
| *CLN* score × GFC | -0.577*** | -0.570** | -1.144*** | -1.364*** | -1.038*** | -0.550** | -0.953*** | -0.883** |
| | (-3.390) | (-2.557) | (-3.971) | (-3.293) | (-2.781) | (-2.068) | (-2.768) | (-2.208) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3884 | 3744 | 3609 | 3478 | 3350 | 3224 | 3103 | 2988 |
| Adjusted $R^2$ | 0.277 | 0.245 | 0.252 | 0.253 | 0.239 | 0.218 | 0.213 | 0.203 |

Table IA.4.6: **Controlling for bank stock performance**

Table IA.4.6 examines the heterogeneous effects of bank co-lending network (CLN) risk score on predicting h-quarter-ahead bank loan loss provisions for different size of banks. The Small bank dummy variable equals to 1 (0) if the bank asset size is lower than (greater than or equal to) the median by each year-quarter. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All public banks*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.092*** | 0.071*** | 0.079*** | 0.055*** | 0.063*** | 0.034* | 0.034* | -0.000 |
| | (5.949) | (5.389) | (5.520) | (3.600) | (3.577) | (1.918) | (1.823) | (-0.004) |
| Quarterly stock returns | -0.047 | -0.401*** | -0.346*** | -0.204** | -0.110 | -0.329*** | -0.245*** | -0.266*** |
| | (-0.597) | (-4.216) | (-3.969) | (-2.398) | (-1.439) | (-3.987) | (-3.762) | (-3.211) |
| Observations | 7972 | 7753 | 7535 | 7323 | 7113 | 6909 | 6714 | 6525 |
| Adjusted $R^2$ | 0.603 | 0.553 | 0.524 | 0.507 | 0.490 | 0.485 | 0.485 | 0.489 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.002*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** | 0.002*** |
| | (3.651) | (3.614) | (3.635) | (3.700) | (3.275) | (3.380) | (3.263) | (3.309) |
| Quarterly stock returns | -0.004** | -0.005** | -0.006*** | -0.006*** | -0.007*** | -0.005* | -0.003* | -0.002 |
| | (-2.504) | (-2.250) | (-3.623) | (-2.791) | (-2.713) | (-1.879) | (-1.829) | (-1.031) |
| Observations | 7972 | 7753 | 7535 | 7323 | 7113 | 6909 | 6714 | 6525 |
| Adjusted $R^2$ | 0.603 | 0.553 | 0.524 | 0.507 | 0.490 | 0.485 | 0.485 | 0.489 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.102*** | -0.122*** | -0.150*** | -0.115*** | -0.124*** | -0.114*** | -0.105*** | -0.079*** |
| | (-7.781) | (-6.672) | (-6.179) | (-4.285) | (-5.314) | (-4.134) | (-4.617) | (-3.679) |
| Quarterly stock returns | 0.017 | 0.437*** | 0.380*** | 0.405*** | 0.328*** | 0.399*** | 0.173*** | 0.239** |
| | (0.237) | (2.873) | (2.928) | (3.554) | (3.544) | (4.327) | (2.606) | (2.599) |
| Observations | 7972 | 7753 | 7535 | 7323 | 7113 | 6909 | 6714 | 6525 |
| Adjusted $R^2$ | 0.595 | 0.411 | 0.330 | 0.317 | 0.292 | 0.288 | 0.290 | 0.289 |
| ***Panel B: Unlabeled public banks*** | | | | | | | | |
| ***Dependent variable: Loan loss provisions*** | | | | | | | | |
| *CLN* score | 0.042** | 0.021 | 0.024 | 0.010 | 0.010 | -0.004 | -0.002 | -0.026 |
| | (2.490) | (1.206) | (1.144) | (0.513) | (0.505) | (-0.212) | (-0.080) | (-1.107) |
| Quarterly stock returns | 0.177 | -0.025 | -0.052 | 0.107 | 0.209 | -0.141 | -0.228* | -0.199 |
| | (1.576) | (-0.141) | (-0.373) | (0.673) | (1.627) | (-1.230) | (-1.800) | (-1.634) |
| Observations | 3579 | 3494 | 3407 | 3322 | 3239 | 3152 | 3059 | 2977 |
| Adjusted $R^2$ | 0.581 | 0.534 | 0.517 | 0.497 | 0.480 | 0.494 | 0.497 | 0.516 |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.001** | 0.001* | 0.001** | 0.001** | 0.001** | 0.001** | 0.002** | 0.002** |
| | (2.035) | (1.953) | (2.086) | (2.241) | (2.015) | (2.353) | (2.061) | (2.002) |
| Quarterly stock returns | 0.002 | -0.001 | 0.002 | 0.002 | 0.003 | 0.004 | 0.003 | 0.004* |
| | (0.710) | (-0.139) | (0.649) | (0.747) | (1.208) | (1.442) | (1.203) | (1.664) |
| Observations | 3022 | 2940 | 2860 | 2775 | 2692 | 2606 | 2534 | 2460 |
| Adjusted $R^2$ | 0.332 | 0.323 | 0.324 | 0.317 | 0.307 | 0.304 | 0.311 | 0.314 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.074** | -0.083** | -0.073** | -0.034 | -0.019 | -0.020 | -0.051 | -0.036 |
| | (-2.509) | (-2.056) | (-1.991) | (-1.083) | (-0.716) | (-0.888) | (-1.582) | (-1.326) |
| Quarterly stock returns | -0.060 | 0.024 | -0.021 | 0.052 | -0.091 | 0.012 | -0.048 | 0.062 |
| | (-0.542) | (0.139) | (-0.154) | (0.323) | (-0.720) | (0.110) | (-0.420) | (0.613) |
| Observations | 3579 | 3494 | 3407 | 3322 | 3239 | 3152 | 3059 | 2977 |
| Adjusted $R^2$ | 0.583 | 0.373 | 0.330 | 0.291 | 0.270 | 0.292 | 0.318 | 0.393 |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Table IA.4.7: **Bank Co-Lending Network Risk, Bank Size Rank and Bank Loan Loss Provisions**

Table IA.4.7 examines the heterogeneous effects of bank co-lending network (CLN) risk score on predicting h-quarter-ahead bank loan loss provisions interacting with bank size rank. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Loan loss provisions** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All samples*** | | | | | | | | |
| *CLN* score | 0.158*** | 0.168*** | 0.160*** | 0.143*** | 0.193*** | 0.161*** | 0.148*** | 0.128** |
| | (5.295) | (5.273) | (4.314) | (3.835) | (4.206) | (3.463) | (2.946) | (2.201) |
| Bank size rank | -0.010 | -0.016 | -0.024 | -0.031* | -0.034** | -0.037** | -0.042** | -0.048** |
| | (-0.636) | (-0.986) | (-1.393) | (-1.816) | (-1.980) | (-2.087) | (-2.297) | (-2.542) |
| *CLN* score × Bank size rank | -0.009** | -0.010** | -0.009* | -0.011** | -0.017*** | -0.015** | -0.014* | -0.015* |
| | (-2.128) | (-2.283) | (-1.736) | (-2.200) | (-2.682) | (-2.301) | (-1.897) | (-1.887) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11308 | 10966 | 10631 | 10303 | 9980 | 9666 | 9365 | 9073 |
| Adjusted $R^2$ | 0.556 | 0.512 | 0.485 | 0.472 | 0.453 | 0.445 | 0.444 | 0.445 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | 0.132*** | 0.141*** | 0.128*** | 0.131*** | 0.175*** | 0.140*** | 0.133** | 0.107 |
| | (3.652) | (3.570) | (2.793) | (3.058) | (3.163) | (2.784) | (2.307) | (1.602) |
| Bank size rank | -0.013 | -0.017 | -0.017 | -0.024 | -0.028 | -0.027 | -0.031 | -0.038 |
| | (-0.573) | (-0.755) | (-0.706) | (-0.992) | (-1.156) | (-1.081) | (-1.204) | (-1.470) |
| *CLN* score × Bank size rank | -0.009* | -0.011* | -0.010 | -0.014** | -0.017** | -0.015** | -0.014* | -0.015 |
| | (-1.719) | (-1.864) | (-1.510) | (-2.376) | (-2.253) | (-2.022) | (-1.689) | (-1.606) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6907 | 6699 | 6495 | 6294 | 6098 | 5901 | 5702 | 5518 |
| Adjusted $R^2$ | 0.511 | 0.478 | 0.457 | 0.449 | 0.426 | 0.431 | 0.432 | 0.438 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.146*** | 0.179*** | 0.174*** | 0.163*** | 0.235*** | 0.191*** | 0.157** | 0.106 |
| | (3.132) | (3.558) | (2.928) | (2.964) | (3.276) | (2.928) | (2.126) | (1.199) |
| Bank size rank | 0.001 | -0.002 | -0.009 | -0.021 | -0.018 | -0.016 | -0.030 | -0.043 |
| | (0.018) | (-0.061) | (-0.242) | (-0.520) | (-0.472) | (-0.406) | (-0.765) | (-1.095) |
| *CLN* score × Bank size rank | -0.009 | -0.011 | -0.008 | -0.013 | -0.018* | -0.015 | -0.007 | -0.007 |
| | (-1.291) | (-1.403) | (-1.000) | (-1.617) | (-1.835) | (-1.492) | (-0.634) | (-0.576) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3882 | 3742 | 3607 | 3476 | 3348 | 3223 | 3103 | 2988 |
| Adjusted $R^2$ | 0.496 | 0.469 | 0.447 | 0.438 | 0.419 | 0.417 | 0.412 | 0.413 |

Table IA.4.8: **Robustness: Bank Co-Lending Network Risks on Small Banks**

Table IA.4.8 examines the heterogeneous effects of bank co-lending network (CLN) risk score on predicting h-quarter-ahead bank loan loss provisions for different size of banks. The Small bank dummy variable equals to 1 (0) if the bank asset size is lower than (greater than or equal to) the median by each year-quarter. Panel A reports the results using all observations. Panel B shows the results using unlabelled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we do not report the coefficients estimates of the control variables. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1)<br>h=1 | (2)<br>h=2 | (3)<br>h=3 | (4)<br>h=4 | (5)<br>h=5 | (6)<br>h=6 | (7)<br>h=7 | (8)<br>h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Panel A: All samples*** | | | | | | | | |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.115 | 0.132* | 0.147** | 0.149** | 0.158** | 0.163** | 0.168** | 0.155** |
| | (1.586) | (1.851) | (2.070) | (2.183) | (2.329) | (2.455) | (2.277) | (2.163) |
| Small bank dummy | -0.284** | -0.278** | -0.256** | -0.237* | -0.215* | -0.184 | -0.171 | -0.168 |
| | (-2.210) | (-2.200) | (-2.070) | (-1.935) | (-1.768) | (-1.507) | (-1.402) | (-1.283) |
| *CLN* score × Small bank dummy | 0.249** | 0.268** | 0.267** | 0.263** | 0.254* | 0.200 | 0.186 | 0.187 |
| | (2.201) | (2.232) | (2.154) | (2.035) | (1.899) | (1.474) | (1.251) | (1.162) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 10279 | 10036 | 9797 | 9504 | 9270 | 9039 | 8812 | 8585 |
| Adjusted $R^2$ | 0.434 | 0.398 | 0.361 | 0.326 | 0.302 | 0.286 | 0.274 | 0.265 |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.071*** | -0.084*** | -0.120*** | -0.086** | -0.073*** | -0.073*** | -0.057*** | -0.050** |
| | (-4.871) | (-4.727) | (-4.470) | (-2.587) | (-2.726) | (-2.786) | (-2.650) | (-2.096) |
| Small bank dummy | 0.028 | 0.041 | 0.022 | 0.023 | 0.045 | 0.036 | 0.017 | -0.009 |
| | (1.362) | (1.221) | (0.502) | (0.480) | (0.931) | (0.783) | (0.370) | (-0.191) |
| *CLN* score × Small bank dummy | -0.070*** | -0.125*** | -0.104** | -0.112** | -0.182*** | -0.156*** | -0.113** | -0.052 |
| | (-2.637) | (-3.635) | (-2.501) | (-2.163) | (-3.957) | (-3.398) | (-2.131) | (-0.845) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 11310 | 10968 | 10633 | 10305 | 9982 | 9667 | 9365 | 9073 |
| Adjusted $R^2$ | 0.569 | 0.382 | 0.309 | 0.304 | 0.267 | 0.254 | 0.253 | 0.256 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.057 | 0.061 | 0.070 | 0.079 | 0.096 | 0.102 | 0.112 | 0.102 |
| | (0.978) | (1.067) | (1.174) | (1.299) | (1.516) | (1.567) | (1.313) | (1.115) |
| Small bank dummy | -0.209* | -0.224* | -0.222* | -0.235* | -0.233* | -0.200 | -0.177 | -0.187 |
| | (-1.674) | (-1.720) | (-1.673) | (-1.713) | (-1.668) | (-1.433) | (-1.240) | (-1.171) |
| *CLN* score × Small bank dummy | 0.316** | 0.340** | 0.344** | 0.366** | 0.370** | 0.296* | 0.260 | 0.251 |
| | (2.220) | (2.233) | (2.234) | (2.290) | (2.275) | (1.761) | (1.340) | (1.145) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6340 | 6191 | 6040 | 5862 | 5714 | 5559 | 5400 | 5250 |
| Adjusted $R^2$ | 0.425 | 0.381 | 0.345 | 0.304 | 0.275 | 0.260 | 0.251 | 0.239 |

Table IA.4.8: **Continued**

| **Dependent variable (t+h):** | (1) h=1 | (2) h=2 | (3) h=3 | (4) h=4 | (5) h=5 | (6) h=6 | (7) h=7 | (8) h=8 |
|---|---|---|---|---|---|---|---|---|
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.053** | -0.047* | -0.067** | -0.015 | 0.005 | -0.036 | -0.019 | 0.003 |
| | (-2.113) | (-1.892) | (-2.253) | (-0.536) | (0.201) | (-1.187) | (-0.676) | (0.114) |
| Small bank dummy | 0.044 | 0.059 | 0.058 | 0.067 | 0.095 | 0.048 | 0.042 | 0.054 |
| | (1.381) | (1.160) | (0.954) | (1.003) | (1.393) | (0.737) | (0.700) | (0.891) |
| *CLN* score × Small bank dummy | -0.063 | -0.112*** | -0.103* | -0.138** | -0.217*** | -0.105* | -0.108 | -0.066 |
| | (-1.552) | (-2.646) | (-1.964) | (-2.156) | (-3.763) | (-1.964) | (-1.513) | (-0.818) |
| | (2.585) | (2.361) | | | | | | |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6909 | 6701 | 6497 | 6296 | 6100 | 5902 | 5702 | 5518 |
| Adjusted $R^2$ | 0.535 | 0.351 | 0.302 | 0.306 | 0.253 | 0.247 | 0.256 | 0.278 |
| | | | | | | | | |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| ***Dependent variable: Non-performing loans*** | | | | | | | | |
| *CLN* score | 0.033 | 0.037 | 0.057 | 0.088 | 0.128 | 0.154 | 0.158 | 0.155 |
| | (0.384) | (0.435) | (0.661) | (0.969) | (1.257) | (1.436) | (1.378) | (1.236) |
| Small bank dummy | -0.174 | -0.210 | -0.211 | -0.238 | -0.247 | -0.214 | -0.228 | -0.256 |
| | (-0.870) | (-0.973) | (-0.927) | (-1.004) | (-1.016) | (-0.904) | (-0.919) | (-0.914) |
| *CLN* score × Small bank dummy | 0.485** | 0.549** | 0.548** | 0.558** | 0.575** | 0.443* | 0.414 | 0.386 |
| | (2.260) | (2.363) | (2.337) | (2.269) | (2.285) | (1.707) | (1.420) | (1.147) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3628 | 3520 | 3415 | 3301 | 3198 | 3093 | 2991 | 2889 |
| Adjusted $R^2$ | 0.454 | 0.406 | 0.362 | 0.311 | 0.272 | 0.248 | 0.228 | 0.208 |
| | | | | | | | | |
| ***Dependent variable: Return on assets*** | | | | | | | | |
| *CLN* score | -0.045 | -0.050 | -0.133** | -0.053 | -0.057 | -0.103* | -0.102** | -0.049 |
| | (-1.197) | (-1.291) | (-2.559) | (-1.157) | (-1.327) | (-1.883) | (-2.008) | (-0.992) |
| Small bank dummy | 0.080 | 0.108 | 0.103 | 0.149 | 0.191 | 0.114 | 0.071 | 0.087 |
| | (1.417) | (1.047) | (0.838) | (1.224) | (1.538) | (0.928) | (0.614) | (0.788) |
| *CLN* score × Small bank dummy | -0.079 | -0.113* | -0.062 | -0.134 | -0.223*** | -0.074 | -0.009 | 0.023 |
| | (-1.360) | (-1.776) | (-0.775) | (-1.387) | (-2.639) | (-0.856) | (-0.088) | (0.180) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3884 | 3744 | 3609 | 3478 | 3350 | 3224 | 3103 | 2988 |
| Adjusted $R^2$ | 0.538 | 0.352 | 0.304 | 0.335 | 0.269 | 0.243 | 0.235 | 0.245 |

Table IA.4.9: **Robustness: Bank Co-Lending Network Risks and Bank Profitability**

Table IA.4.9 examines the heterogeneous effects of bank co-lending network (CLN) risk score on predicting h-quarter-ahead bank loan loss provisions interacting with bank profitability measure. Panel A reports the results using all observations. Panel B shows the results using unlabeled banks (banks who do not have assigned labels). Panel C shows the results using private banks only. Panel B and C have extactly same control variables as Panel A. For simplicity, we only report the interaction coefficient estimates. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are two-tailed t-statistics. Heteroskedasticity-robust standard errors are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Loan loss provisions** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All samples*** | | | | | | | | |
| *CLN* score × Standardized earnings change | 0.000 | 0.000 | -0.003** | -0.009*** | -0.002 | -0.003*** | -0.003 | -0.004*** |
| | (0.076) | (0.171) | (-2.076) | (-2.843) | (-0.661) | (-3.890) | (-1.631) | (-3.746) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 10951 | 10619 | 10296 | 9978 | 9666 | 9362 | 9071 | 8787 |
| Adjusted $R^2$ | 0.557 | 0.517 | 0.487 | 0.475 | 0.452 | 0.447 | 0.444 | 0.446 |
| | | | | | | | | |
| *CLN* score × ROA volatility | 0.061 | 0.050 | 0.024 | 0.043 | 0.094*** | 0.077*** | 0.082*** | 0.053** |
| | (0.963) | (0.606) | (0.371) | (1.016) | (2.964) | (2.649) | (3.145) | (2.077) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 10382 | 10054 | 9732 | 9420 | 9118 | 8826 | 8547 | 8275 |
| Adjusted $R^2$ | 0.572 | 0.529 | 0.498 | 0.483 | 0.465 | 0.459 | 0.453 | 0.453 |
| | | | | | | | | |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score × Standardized earnings change | 0.000 | 0.000 | -0.002 | -0.008*** | -0.001 | -0.003*** | -0.002 | -0.004*** |
| | (0.177) | (0.153) | (-1.566) | (-2.637) | (-0.521) | (-3.257) | (-1.421) | (-3.066) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6583 | 6385 | 6192 | 6000 | 5814 | 5627 | 5436 | 5258 |
| Adjusted $R^2$ | 0.512 | 0.485 | 0.461 | 0.454 | 0.428 | 0.433 | 0.432 | 0.439 |
| | | | | | | | | |
| *CLN* score × ROA volatility | 0.128*** | 0.139** | 0.102** | 0.139*** | 0.229*** | 0.183*** | 0.124*** | 0.079** |
| | (2.831) | (2.292) | (1.980) | (2.959) | (3.269) | (4.393) | (2.658) | (2.228) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 6315 | 6119 | 5925 | 5733 | 5549 | 5364 | 5179 | 5006 |
| Adjusted $R^2$ | 0.534 | 0.503 | 0.479 | 0.466 | 0.449 | 0.452 | 0.446 | 0.452 |
| | | | | | | | | |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score × Standardized earnings change | 0.000 | -0.000 | -0.003 | -0.008*** | -0.002 | -0.004*** | -0.003* | -0.004*** |
| | (0.226) | (-0.023) | (-1.626) | (-2.786) | (-0.575) | (-3.772) | (-1.887) | (-3.304) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3565 | 3435 | 3311 | 3189 | 3071 | 2955 | 2843 | 2734 |
| Adjusted $R^2$ | 0.496 | 0.478 | 0.453 | 0.447 | 0.423 | 0.422 | 0.415 | 0.417 |
| | | | | | | | | |
| *CLN* score × ROA volatility | 0.220*** | 0.254*** | 0.174** | 0.163* | 0.135* | 0.140** | 0.030 | 0.043 |
| | (3.487) | (3.415) | (2.296) | (1.787) | (1.817) | (2.171) | (0.633) | (0.912) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bank Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3485 | 3356 | 3230 | 3107 | 2988 | 2872 | 2763 | 2658 |
| Adjusted $R^2$ | 0.523 | 0.500 | 0.475 | 0.462 | 0.446 | 0.445 | 0.437 | 0.434 |

Table IA.4.10: **Bank Co-Lending Network Risk, Bank opacity and Bank Loan Loss Provisions**

Table IA.4.10 examines the heterogeneous effects of bank co-lending network (CLN) risk score on predicting h-quarter-ahead bank loan loss provisions for different banks. Definitions of the variables are provided in Table A1 in the Appendix. Numbers in parentheses are heteroskedasticity-robust standard errors clustered at the borrower level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

**All Samples**

| **Dependent variable (t+h):** | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| **Loan loss provisions** | h=1 | h=2 | h=3 | h=4 | h=5 | h=6 | h=7 | h=8 |
| ***Panel A: All Samples*** | | | | | | | | |
| *CLN* score | 0.592*** | 0.440*** | 0.363*** | 0.421*** | 0.376*** | 0.183* | 0.087 | 0.160 |
| | (5.041) | (4.157) | (3.640) | (3.823) | (3.920) | (1.966) | (0.950) | (1.522) |
| Bank Opacity | 0.021* | 0.020* | 0.016 | 0.020 | 0.017 | 0.025* | 0.030* | 0.034* |
| | (1.697) | (1.686) | (1.207) | (1.334) | (1.184) | (1.796) | (1.930) | (1.880) |
| *CLN* score × Bank opacity | 0.075*** | 0.051*** | 0.038*** | 0.053*** | 0.042*** | 0.017 | 0.003 | 0.019 |
| | (4.602) | (3.380) | (2.635) | (3.387) | (3.062) | (1.223) | (0.247) | (1.336) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 9086 | 8909 | 8735 | 8552 | 8375 | 8202 | 8039 | 7879 |
| Adjusted $R^2$ | 0.579 | 0.530 | 0.498 | 0.485 | 0.462 | 0.454 | 0.451 | 0.455 |
| ***Panel B: Out-of-sample (unlabeled banks)*** | | | | | | | | |
| *CLN* score | 0.558*** | 0.395*** | 0.332*** | 0.387*** | 0.413*** | 0.259** | 0.120 | 0.206 |
| | (3.695) | (2.784) | (2.683) | (2.626) | (3.572) | (2.347) | (1.096) | (1.437) |
| Bank opacity | 0.032** | 0.026* | 0.018 | 0.036** | 0.029* | 0.040** | 0.047** | 0.051** |
| | (2.037) | (1.728) | (1.129) | (2.037) | (1.699) | (2.351) | (2.534) | (2.297) |
| *CLN* score × Bank opacity | 0.074*** | 0.048** | 0.038** | 0.052** | 0.051*** | 0.031* | 0.011 | 0.028 |
| | (3.524) | (2.396) | (2.145) | (2.473) | (3.013) | (1.853) | (0.642) | (1.419) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 5608 | 5498 | 5392 | 5282 | 5169 | 5052 | 4937 | 4829 |
| Adjusted $R^2$ | 0.538 | 0.498 | 0.472 | 0.465 | 0.435 | 0.440 | 0.439 | 0.450 |
| ***Panel C: Out-of-sample (private banks)*** | | | | | | | | |
| *CLN* score | 0.731*** | 0.588*** | 0.554*** | 0.775*** | 0.766*** | 0.576*** | 0.353** | 0.420** |
| | (3.227) | (2.819) | (3.259) | (3.925) | (4.746) | (3.994) | (2.406) | (2.092) |
| Bank opacity | 0.053** | 0.043** | 0.022 | 0.052** | 0.050** | 0.064*** | 0.060** | 0.079** |
| | (2.563) | (2.000) | (0.966) | (2.282) | (2.243) | (2.688) | (2.310) | (2.569) |
| *CLN* score × Bank opacity | 0.099*** | 0.071** | 0.062** | 0.105*** | 0.097*** | 0.072*** | 0.037 | 0.055** |
| | (3.117) | (2.353) | (2.492) | (3.747) | (4.167) | (3.285) | (1.611) | (1.981) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Quarter Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3213 | 3128 | 3048 | 2967 | 2889 | 2810 | 2734 | 2659 |
| Adjusted $R^2$ | 0.529 | 0.494 | 0.464 | 0.466 | 0.439 | 0.438 | 0.425 | 0.434 |