A high frequency indicator of credit in Peru: A Random Forests and dynamic network connectedness approach

Rodrigo Grandez Renzo Pardo Figueroa

September 6, 2024

Abstract

This paper contributes to the literature on early warning indicators for financial stability by developing a daily credit expansion indicator for Peru. We implemented random forest (RF) models to impute the missing daily data from the monthly credit series, and quantile RF models to identify periods of excessive credit growth, using high-frequency financial variables such as spreads, interest rates, and stock market returns. This method integrates machine learning with mixed-frequency analysis to develop an early warning system for detecting excessive credit growth in Peru. By effectively managing nonlinear relationships and adapting to various scenarios, this approach provides substantial benefits to central bankers and macroprudential authorities in maintaining financial stability. Additionally, we analyze the financial connectivity among these series using a VAR model. To measure the transmission of shocks between these variables, we employ the Diebold & Yilmaz (2012) approach, which is widely used to study connectivity across different financial markets. We find that significant spillover effects occur during periods associated with specific events, such as the recent COVID-19 pandemic.

Keywords: Random Forest, credit.
JEL Classification: C32, ES1, GO1.

1 Introduction

One of the key variables that central banks monitor to maintain financial system stability is the volume of credit in the financial system. This indicator reflects the ability of banks to channel resources into the real economy. Abnormal variations in credit could be associated with disruptive events in the financial system and the economy. Unexpected and large-scale shocks such as those that occurred at the beginning of the Covid pandemic can potentially break the payment chain, requiring government credit support programs Hong & Lucas (2023). Conversely, the uncontrolled increase in loans can increase the probability of the occurrence of crises through credit (Caballero 2016). Financial risks during an economic expansion cannot be underestimated, prolonged accommodative monetary policies can increase financial fragility, which reinforces the need for constant vigilance even during periods of economic growth (Grimm et al. 2023).

Central banks in the world have emphasized the importance of monitoring credit indicators as part of their macroprudential policy framework (Borio 2014). Since credit data is traditionally reported on a monthly basis, its availability at high frequency (daily) would provide a valuable tool for central banks, allowing them to react more quickly to changes in financial conditions. This is especially relevant in contexts of financial stress, where rapid credit expansion or contraction can amplify economic cycles (Kiyotaki & Moore 1997, Claessens & Kose 2013).

The first objective of this study is to impute daily credit data to construct a daily frequency credit indicator. To achieve this, following Giraldo et al. (2024), we employ a random forest model, using high-frequency financial variables such as spreads, interest rates, and stock market returns as predictors. We use Quantile Random Forest to pinpoint periods of excessive credit growth that might worry regulators. Our focus is on the top 5 percent of credit growth quantiles, as these are most likely to signal future financial and macroeconomic instability. Data imputation with machine learning techniques has shown remarkable effectiveness, allowing for precise and robust imputation of missing data. In particular, the random forest method has been highlighted for its ability to capture nonlinear relationships and handle complex interactions between variables, making it a valuable tool in financial applications (Breiman & Breiman 2001, Hastie et al. 2009). Moreover, recent studies have underscored the relevance of these techniques in the financial context, highlighting their application in risk prediction, asset valuation, and missing financial data imputation Varian (2014).

Daily credit variation is related to various economic and financial factors. Yield spreads across different maturities reflect expectations about future interest rates, which can significantly influence the daily supply and demand for credit. Additionally, risk perceptions play a crucial role in credit dynamics. The Morgan Stanley Capital International (MSCI) index is widely used as an indicator of equity risk sentiment in emerging economies. In times of heightened uncertainty, lenders may restrict credit access or increase interest rates to compensate for perceived risk Adrian & Shin (2010). Libor rates are an important indicator of international liquidity conditions. Global liquidity conditions are considered key determinants of the cost of interbank dollar loans, particularly in emerging economies such as Peru Avdjiev et al. (2020). When these rates rise, the cost of dollar financing for local banks also increases, potentially restricting credit availability Obstfeld (2013).

As far as we know, this is the first study to create a practical early warning tool for detecting excessive credit growth in Peru through the use of machine learning techniques.

Additionally, the second objective of this study is to analyze the financial connectivity between high-frequency financial series and the credit indicator with imputed data using a VAR model. To measure the transmission of shocks among these variables, we employ the Diebold & Yilmaz (2012) approach, which has been widely used to study connectivity in different financial markets Demirer et al. (2018), Antonakakis et al. (2017). This approach is particularly relevant for understanding how financial conditions interact with monetary and credit policy in an emerging market like Peru, distinguishing events where certain types of shocks may prevail over others that affect credit.

The contribution of this study is twofold. On the one hand, we use a machine learning methodology to impute daily credit data, allowing for more detailed and timely analyses of credit dynamics in relation to other financial variables. On the other hand, it provides empirical evidence on financial connectivity in the Peruvian market, which has significant implications for both policymakers and investors.

The paper is organized as follows. Section 2 show a brief literature review, section 3 presents the methodology employed in this study, showing both the data imputation approach using machine learning techniques and the VAR model. Then, Section 4 provides the description and statistics of the data used. Section 5 presents the results and analysis. Finally, Section 6 presents the conclusions.

2 Literature review

Monitoring credit activity at high frequency is crucial for understanding the dynamics of financial markets and ensuring economic stability. Schwaab et al. (2014) highlights the significant role that the credit market has played in the years leading up to economic crises in various countries, underscoring the importance of continuously tracking credit behavior to prevent systemic risks. This view is supported by Trejo García et al. (2017) and Ma et al. (2023), who emphasize that accurate credit predictions are essential for mitigating financial risks, making informed lending decisions, and contributing to the overall stability of the financial system.

Recent advancements in econometric and machine learning methodologies have facilitated more precise real-time predictions, known as nowcasting. Kant et al. (2022) compare the accuracy of various econometric and machine learning models in real-time GDP prediction, finding that machine learning models, particularly Random Forest, outperform traditional econometric models in terms of forecasting accuracy. The Random Forest model, as they note, uses a more balanced and stable set of variables, especially financial and international trade data, compared to traditional models like the Dynamic Factor Model (DFM), which has shifted focus to production, sales, and finance variables.

While most studies have concentrated on GDP nowcasting, there is a growing interest in applying similar techniques to credit data nowcasting. Richardson et al. (2021) point out that machine learning models are particularly effective for nowcasting when dealing with large datasets that include numerous potential regressors. For example, Gosh and Ranjan (2023) successfully applied a combination of DFM and machine learning techniques, such as Random Forest and Prophet, to predict GDP growth in emerging markets like India. This methodology shows promise for extending nowcasting techniques to credit data, although literature on this specific application remains limited.

Bitetto et al. (2023) provide a compelling case for using historical ML models, particularly Random Forest, in credit risk estimation for small and medium-sized businesses (SMBs). Their study demonstrates that Random Forest models outperform traditional parametric models, particularly in the classification of credit risk, indicating the model's robustness in handling complex financial data. Ma et al. (2023) further explore the application of Explainable Artificial Intelligence (XAI) techniques to predict credit default risk in Chinese real estate companies, providing insights into the adaptability of machine learning models across different financial sectors.

Finally, Schwaab et al. (2014) developed a dynamic model for predicting systematic stress risk in financial institutions on an international scale. Their methodology, which includes a mixed dynamic factor approach to model credit risk dynamics, reveals that macroeconomic, industry-specific, and fragility factors significantly influence global credit risk conditions. This study underlines the importance of adaptable and comprehensive models, such as those offered by machine learning techniques, for assessing credit risk in a globalized financial environment.

Together, these studies highlight the growing role of machine learning, particularly Random Forest, in improving the accuracy of credit predictions and understanding the connectedness among financial variables. The application of these advanced methodologies to credit data nowcasting and risk prediction represents a significant step forward in enhancing financial stability and resilience.

3 Methodology

3.1 A daily credit growth indicator

In the first part, we tackle the challenge of mixed frequencies in our dataset by treating it as a problem of predicting missing values. For this, we adapt the approach described by Stekhoven and Bühlmann (2012), who used a Random Forest method for data imputation. However, whereas their method was designed for standard frequency data, our dataset includes both daily and monthly frequencies.

We assume $X = (X_1, X_2, ..., X_p)$ to be a $n \times p$-dimensional data matrix. We propose using a random forest to impute the missing values due to its earlier mentioned advantages as a regression method. The random forest algorithm has a built-in routine to handle missing values by weighting the frequency of the observed values in a variable with the random forest proximities after being trained on the initially mean imputed data set (Breiman (2001)). However, this approach requires a complete response variable for training the forest. Instead, we directly predict the missing values using a random forest trained on the observed parts of the data set.

For any lower frequency variable X_s,including missing points at entries $i_d^{NA} \epsilon \{1, \ldots, n\}$ the data set can be split into four sets:1)the non-missing values of X_s,which are denoted by y_s^{obs}; 2) the missing observations, y_s^{NA} ; 3) variables different from s , with higher frequencies, with observations $i_s^{obs} = \{1, \ldots, n\} \setminus i_s^{NA}$ denoted as X_s^{obs} , and 4) indicators different than X_s with observations i_s^{NA} ,denoted by x_s^{NA}.

Note that x_s^{obs} is typically not completely observed since the index i_s^{obs} corresponds to the observed values of the variable X_s. Likewise, x_s^{NA} is typically not completely missing.

The random forest algorithm makes an initial conjecture for the missing values in X using mean imputation or another imputation method. Then, sort the variables $X_s, s = 1, ..., p$ according to the amount of missing values starting with the lowest amount. For each variable X_s the missing values are imputed by first fitting a random forest with response y_s^{obs} and predictors X_s^{obs}, in a given day, which in our case correspond to the higher frequency variables. Then, predicting the missing values y_s^{NA} by applying the trained random forest to x_s^{NA}. The imputation procedure is repeated until a stopping criterion is met.

In the second part, Quantile regression forests (Meinshausen,2006) generalizes the standard random forests to provide information for the full conditional distribution of the response variable, not only about the conditional mean. Therefore quantile regression give a non-parametric and accurate way of estimating conditional quantiles for high-dimensional predictor variables.

The standard random forest model calculates the mean value of the target variable in each tree leaf. If we record all observed responses in the leaf, we will be able to calculate the percentiles. Assume in a random forest model there are 100 trees, which produce 100 predicted values for an input observation. The standard random forests get the conditional mean by taking the mean of the 100 predicted values. We can extend this to get the entire distribution thus the confidence intervals.

In our empirical analysis, we use the high quantile (i.e., 0.95) to represent scenarios of rapid credit growth, which might be linked to financial instability or bubbles in financial markets. We then compare these results with those from the low quantile (i.e., 0.05).

3.2 Connectedness methodology

A vector autoregressive (VAR) model with a lag length of p can be formulated as follows:

$$y_t = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i} + u_t \quad u_t \sim N(0, \Sigma)$$

where $y_t, t = 1, \ldots, T - p$, is an $k \times 1$ dimensional vector of the endogenous variables, β_0 is an $k \times 1$ dimensional vector of all intercepts, and β_i represents the $k \times k$ dimensional VAR coefficient matrix of the i th lag. The $k \times 1$ dimensional error vector, u_t, $t = 1, \ldots, T - p$ is multivariate normally distributed with means equal to zero and a variance-covariance matrix equals to Σ. Σ is a diagonal matrix as var $(u_i) = \sigma_i^2, i = 1, \ldots, k$ and cov $(\boldsymbol{u}_i, \boldsymbol{u}_j) = 0, i, j = 1, \ldots, k$ and $i \neq j$.

For the spillover index approach of Diebold & Yilmaz (2012), we use the generalized variance decomposition of Pesaran & Shin (1998), where the variance decompositions are invariant to the ordering of variables in the VAR. The variable j's contribution to variable i's H-step-ahead generalized forecast error variance, $\theta_{ij}^g(H)$, for $H = 1, 2, \ldots$, is defined as:

$$\theta_{ij}^g(H) = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} \left(e_i' A_h \Sigma e_j\right)^2}{\sum_{h=0}^{H-1} \left(e_i' A_h \Sigma A_h' e_i\right)} \tag{1}$$

Diebold & Yilmaz (2012) construct the total connectedness index as:

$$C(H) = \frac{\sum_{\substack{i,j=1 \\ i \neq j}}^{N} \tilde{\theta}_{ij}^g(H)}{\sum_{i,j=1}^{N} \tilde{\theta}_{ij}^g(H)} = \frac{\sum_{\substack{i,j=1 \\ i \neq j}}^{N} \tilde{\theta}_{ij}^g(H)}{N} \tag{2}$$

In similar fashion, directional spillovers transmitted by variable i to all other variables j is measured as:

$$C_{\bullet \leftarrow i} = \frac{\sum_{\substack{j=1 \\ j \neq i}}^{N} \tilde{\theta}_{ji}^g(H)}{\sum_{i,j=1}^{N} \tilde{\theta}_{ji}^g(H)} \times 100 = \frac{\sum_{\substack{j=1 \\ j \neq i}}^{N} \tilde{\theta}_{ji}^g(H)}{N} \times 100 \tag{3}$$

Next considering directional vulnerability, Diebold & Yilmaz (2012) define gross directional vulnerability received by variable i from all other variables j as:

$$C_{i \leftarrow \bullet} = \frac{\sum_{\substack{j=1 \\ j \neq i}}^{N} \tilde{\theta}_{ij}^g(H)}{\sum_{i,j=1}^{N} \tilde{\theta}_{ij}^g(H)} \times 100 = \frac{\sum_{\substack{j=1 \\ j \neq i}}^{N} \tilde{\theta}_{ij}^g(H)}{N} \times 100 \tag{4}$$

The set of directional spillovers can be viewed as a decomposition of total connectedness, revealing the contributions of each variable in the sample. This decomposition quantitatively assesses whether a variable's fluctuations are primarily due to external shocks or how influential the variable is in transmitting its own variations to others.

4 Data

Daily credit variation is influenced by a range of economic and financial factors. Yield spreads across different maturities reveal expectations about future interest rates, which can greatly impact the daily supply and demand for credit. Risk perceptions also play a vital role in credit behavior. The Morgan Stanley Capital International (MSCI) index is commonly used to gauge equity risk sentiment in emerging markets. Libor rates are key indicators of global liquidity conditions. Moreover, the interbank interest rate set by the Central Bank of Peru acts as a reference for financial institutions' borrowing and lending activities, reflecting the current liquidity conditions and monetary policy stance within Peru's financial system. Lastly, changes in total loans, reported by the Central Bank of Peru and labeled as "Credit," offer insights into the trends of credit expansion or contraction within the Peruvian economy. The sample period covers 12 years, from January 2012 to June 2024 (4508 daily observations). In the Annex, figure 7 and table 3 shows in more detail the variables considered in the exercise.

Table 1: Descriptive Statistics

Indicator	Frequency	Abreviation	Source	Mean	Median	Std. Dev	Max.	Min.
Spread 10 years- 2 years bond	Daily	10y-2y	Bloomberg	1,538	1,470	1,154	5,146	-0,607
Spread 10 years- 5 years bond	Daily	10y-5y	Bloomberg	1,568	1,183	0,833	3,848	0,308
Spread LIBOR USD 3 months - 2 years bond	Daily	LIBOR-2y	Bloomberg	-2,707	-2,997	1,282	0,138	-4,645
Spread LIBOR USD 3 months - 5 years bond	Daily	LIBOR-5y	Bloomberg	-2,677	-2,562	1,477	0,383	-6,300
Interbank Interest Rate	Daily	Interbank	Central Bank of Peru	3.757	4.000	1.921	7.9200	0.090
MSCI Peru	Daily	MSCI	Bloomberg	0.013	0,000	1.347	10.714	-12.682
Variation of Total Loans	Monthly	Credit	Central Bank of Peru	0,599	0,601	0,723	5,457	-2,061

5 Results

This section present the main results. Figure 1 compares the original credit series (left panel) with the imputed credit series (right panel), using a MIDAS approach. The imputed series closely mirrors the original, as anticipated. The key difference lies in the frequency: the original series is reported monthly, whereas the imputed series offers daily data. This added granularity enables policymakers to swiftly detect unusual credit growth. Both series exhibit considerable fluctuation and short-term persistence, indicating that credit growth tends to vary significantly over time. When credit growth is elevated today, it is likely to stay high in the near future (within the current month).

Figure 2 presents the key findings of this study by showing the daily credit growth series, along with the high (95th percentile) and low (5th percentile) quantiles of this growth. In the graph, the black curve represents the overall daily credit growth trajectory, while the orange curves denote the high and low quantiles (left panel for high quantiles, right panel for low quantiles). Several important observations emerge from this analysis. Both high and low quantiles of credit growth demonstrate temporal variability, indicating the impact of macroeconomic and financial cycles on identifying periods of unusually high or low credit growth. As a result, our model offers an accurate and internal method for assessing the significance of extreme credit growth on any given day. This feature aids central bankers in developing macro-prudential policies to address abnormal credit growth effectively.

Our approach enables the use of various percentage thresholds to pinpoint periods of unusually high or low credit growth. In this study, we use the 95th and 5th percentiles as cutoffs, but the flexibility of our method allows policymakers to choose other quantiles that they deem suitable.

An important aspect is selecting daily variables to complete the daily credit growth series. Figure

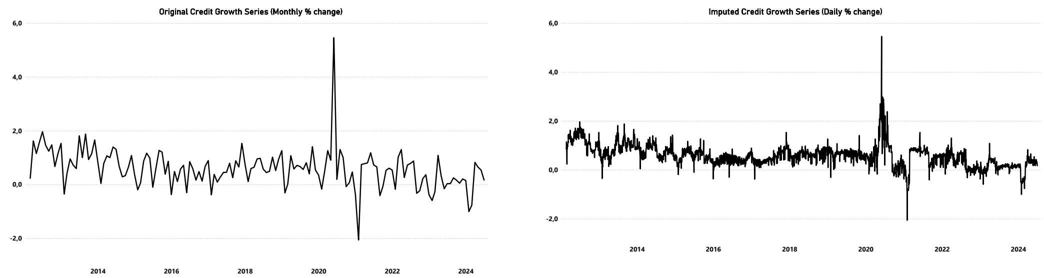

Figure 1: Original versus Imputed Credit Series using a MIDAS approach. Note: The figure presents the original series of monthly credit growth in Peru on the left, while it shows the constructed series of monthly credit growth with a daily frequency using a Mixed Data Sampling approach with Random Forest.

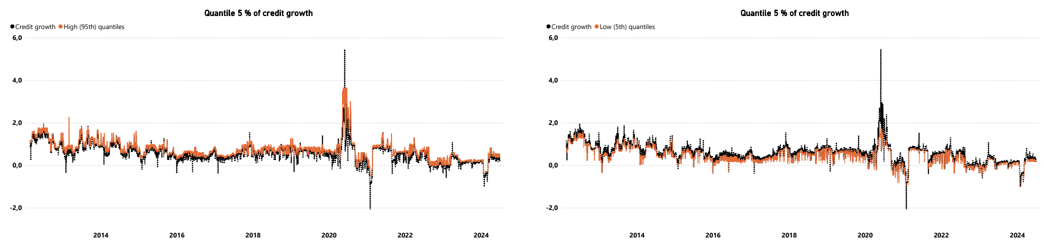

Figure 2: High (95th) and low (5th) quantiles of credit growth versus credit growth. Note: The figure presents the 95th and 5th quantile of credit growth estimated using quantile random forest against the series of credit growth (with a daily frequency).

2 illustrates the Variable Importance, based on node purity, of the credit growth at risk indicator throughout the entire conditional distribution of credit growth, including both high and unusually low growth scenarios. In this analysis, two key variables stand out—both their current and lagged values—namely, the spread between the 3-month LIBOR USD and the 5-year bond yields and the spread between the 10-year bond yields and the 5-year bond yields.

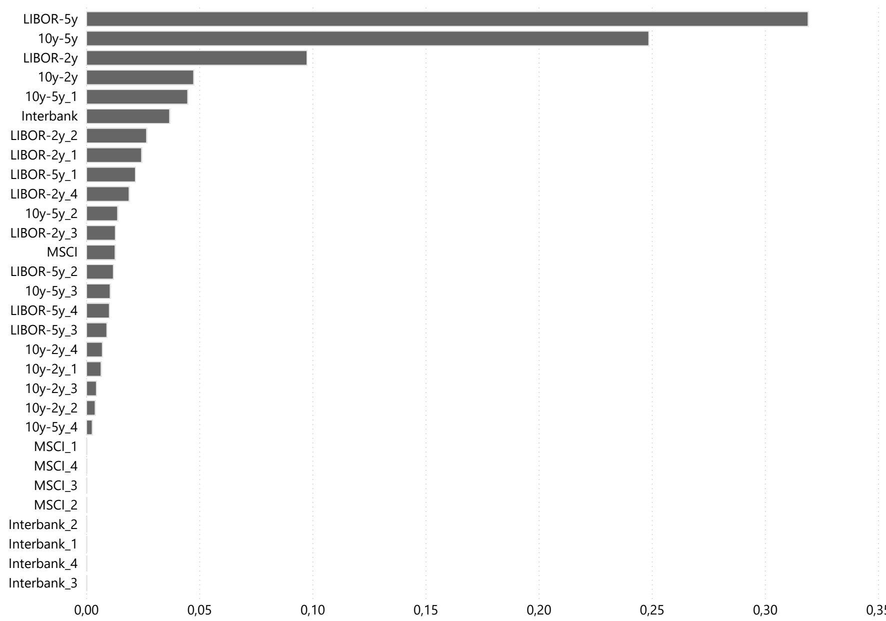

Figure 3: Variable Importance (node purity) of the credit growth at risk indicator. Note: The figure shows the variable importance from most important at the top to least important at the bottom. This graph correspond to the 95th quantile. The number in front of the variables correspond to the lag of the variable.

Table 2 delves deeper into these relationships by showing the values of predictive variables during normal periods (when the credit risk indicator is below the 95th percentile) and during stressed periods with unusually high credit growth (above the 95th percentile). This analysis reveals that most variables are generally higher during times of increased credit creation. In particular, we observe excessive credit creation coincide with higher market returns compared to regular periods.

	Regular	Stress
Credit	0.584	1.358
10y-2y	1.536	2.160
10y-5y	1.568	1.783
LIBOR-2y	-2.703	-2.613
LIBOR-5y	-2.671	-2.989
Interbank	0.000	0.000
MSCI	0.013	0.599

Table 2: This table displays the average values of the variables under two different states: when credit growth is below its 95th percentile (considered regular- left column) and when it exceeds this threshold (considered stressed- right column).

Now present the connectedness analysis, with the daily sample series and the imputed daily credit series. We estimate a VAR using 3 lags and generalized variance decompositions of 10-day-ahead forecast error. Aditionally, we employ regularization approaches and Bayesian shrinkage to offer comparative analysis. Table 3 summarizes the total connectedness index and the directional spillovers for the entire sample period. Our results reveal that the series exhibit a moderate average connectedness level of approximately 40 percent. Specifically, we find that the 10-year minus 2-year yield spread and the 2-year LIBOR spread are significant transmitters and receivers of shocks, reflecting the central role in the financial system. In contrast, the MSCI index acts primarily as a transmitter, while the interbank rate is mostly a receiver of shocks. The daily credit indicator, although less influential, also receives spillovers from other financial variables. This suggests that while the credit markets function as net transmitters of their own fluctuations, they are not entirely insulated from broader financial shocks.

Table 3: Connectedness and Spillover Table

Estimation Methods	Minnesota	Ridge	Lasso	Elastic Net
Total connectedness	36,33	36,00	46,94	44,13
Directional to All				
10y-2y	10,23	10,17	12,14	10,44
LIBOR-2y	4,66	4,46	14,45	13,00
Interbank	0,00	0,00	0,00	0,00
MSCI	20,43	16,69	10,26	8,74
Credit	10,10	13,68	21,82	22,98
Directional From All				
10y-2y	11,74	11,83	14,14	12,49
LIBOR-2y	9,45	9,42	10,88	9,34
Interbank	23,45	23,17	20,85	21,50
MSCI	0,03	0,03	0,01	0,01
Credit	0,73	0,56	12,79	11,81
Net Directional Spillover				
10y-2y	-1,52	-1,66	-2,01	-2,05
LIBOR-2y	-4,80	-4,96	3,57	3,66
Interbank	-23,45	-23,17	-20,85	-21,50
MSCI	20,40	16,66	10,25	8,72
Credit	9,37	13,12	9,03	11,17

Figure 4 shows the Total Connectedness Index calculated over time using a 200-day rolling window and a 10-day forecast horizon across the analyzed financial variables. The level of risk in the monetary and credit markets, as captured by the total connectedness, has exhibited significant fluctuations, with peak values surpassing 30 percent and 75 percent. The dynamics are largely similar between the Lasso and Elastic Net models. During periods of financial stress, the interconnectedness among variables intensifies, amplifying volatility transmission across the system, particularly during major economic events like the COVID-19 pandemic. We can distinguish peaks around april 2016, may 2020 among others.

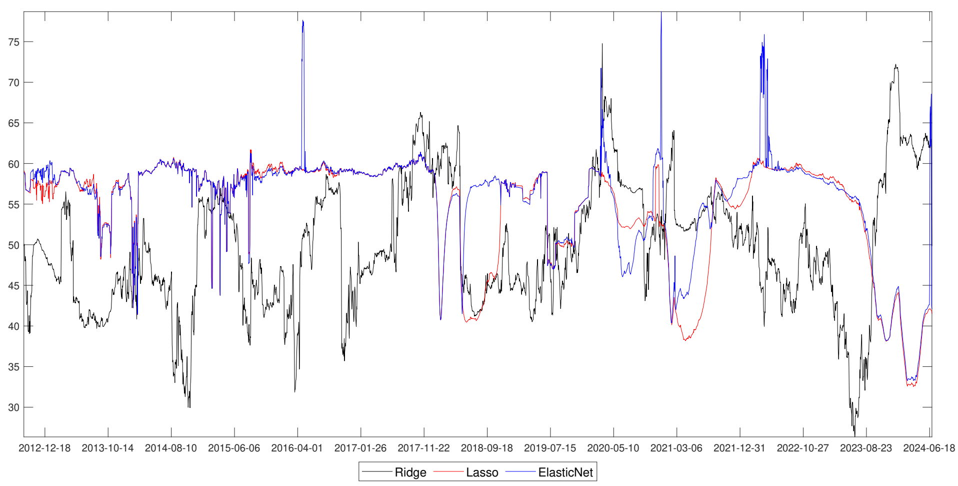

Figure 4: Dynamic total connectedness index. Note: Results are based on all 400-day rolling-window models with a 10-step-ahead forecast horizon.

Figure 5 presents the directional spillovers from one variable to others, illustrating how spillover intensity changes over time. Notably, fluctuations originating from the 10-year minus 2-year spread and the 2-year LIBOR propagate across the system. Significant movements can be observed around 2018 and 2020, affecting most of the variables, highlighting key periods of heightened financial activity and stress.

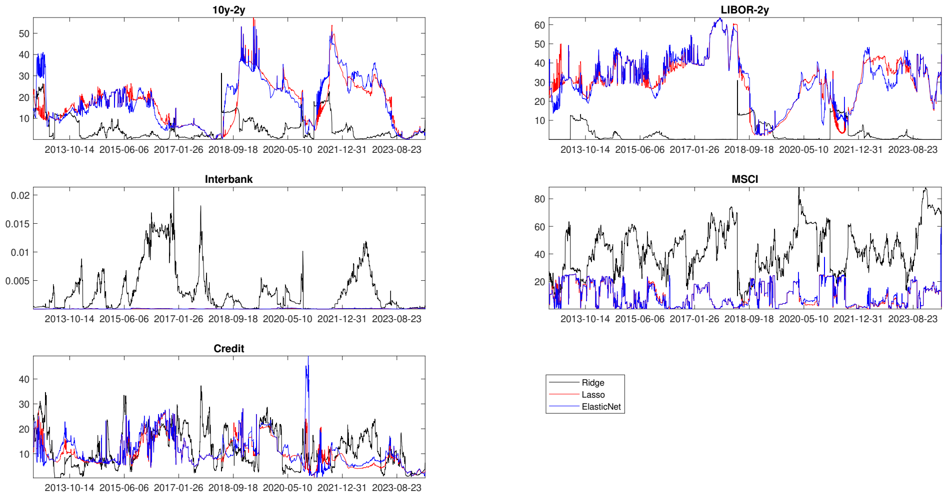

Figure 5: Directional volatility spillovers, FROM one variable to others.

Moreover, Figure 6 reports the directional spillovers from other variables to each specific variable. The findings reveal that certain variables, such as the interbank rate, primarily act as receivers of spillovers, while others, like the MSCI index, serve mainly as contributors. This highlights the asymmetric nature of shock transmission within the financial system, with some variables playing dominant roles in propagating volatility while others absorb it. Further, we find that the interbank rate does not propagate its changes, indicating that monetary policy operates over a longer horizon, with its effects not being visible in daily data. Additionally, daily credit shows limited responsiveness to the reference rate.

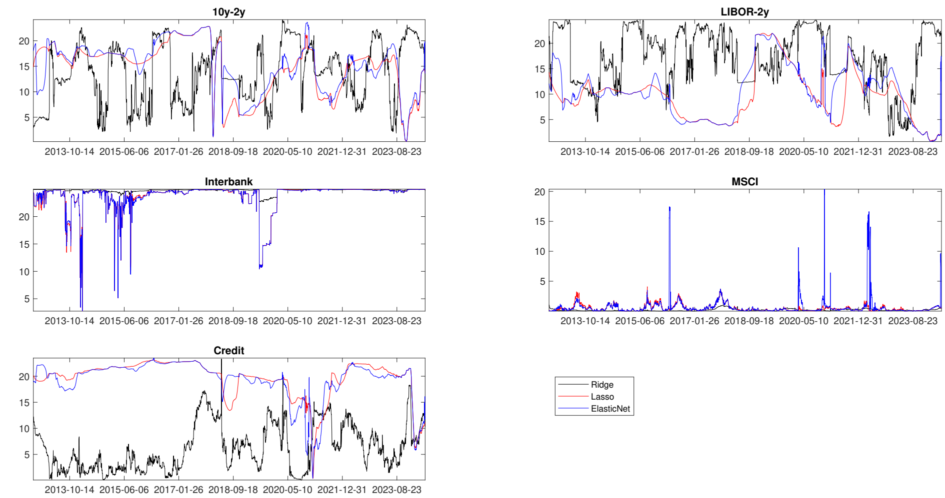

Figure 6: Directional volatility spillovers, from other variables to a variable.

6 Conclusions

This research presents a novel approach for developing an early warning system to detect unusually high or low credit growth using monthly data from Peru. We apply a two-step machine learning technique to generate a daily indicator that tracks changes in credit growth over time, enabling us to monitor daily variations and identify periods of rapid credit expansion.

In the first step, we utilize Random Forests to construct a daily credit growth indicator from the monthly data. Specifically, we address missing dates in the raw credit data by using Random Forests to interpolate these gaps, leveraging information from high-frequency indicators such as spreads, overnight interbank interest rates, and stock market returns. In the second step, we use a quantile random forest, which combines quantile regression with Random Forests, to identify periods of excessive credit creation, which are of regulatory concern. Our analysis focuses on the top 5 percent of credit growth quantiles, indicating instances of rapid credit expansion that could signal potential future risks to financial and macroeconomic stability

This study uncovers three key insights: Firstly, the estimated credit quantiles display clustering patterns that underscore the role of financial and macroeconomic cycles in pinpointing periods of notably high or low credit growth. As a result, relying on ad hoc definitions for these extremes, a common practice among central bankers, is less advisable. Our model offers an endogenously derived and precise method for assessing the significance of high and low credit growth on any given day, thereby aiding central bankers in implementing macroprudential policies. Additionally, our approach supports the use of various percentage cutoffs, providing greater flexibility for policymakers compared to the fixed 95 percent and 5 percent cutoffs used in this study. Secondly, instances where credit growth exceeds or falls below a certain threshold indicate unusually high or low daily credit growth values. Importantly, these periods are not randomly spread over time but instead form discernible clusters. For example, notably high credit growth days were particularly concentrated in 2012, a period characterized by elevated commodity prices. This suggests a positive correlation in Peru, and likely in other emerging markets dependent on commodities, between exceptionally high credit growth and times of rising commodity prices. Consequently, policymakers should closely monitor credit growth during commodity price booms. Another period of unusually high credit growth occurred during the Covid-19 pandemic, reflecting government efforts to boost credit availability and the high prices of copper.

Thirdly, our daily credit indicator shows variable behavior over time and demonstrates short-term clustering within the month. Additionally, during periods of high credit growth, two key factors become particularly important: the spread between the 3-month LIBOR USD and the 5-year bond yields, and the spread between the 10-year bond yields and the 5-year bond yields.

Additionally, we analyze the financial connectivity between high-frequency financial series and the credit indicator with imputed data using a VAR model. This approach reveals significant insights into spillovers among key financial variables and their impact on daily credit dynamics in Peru. The study finds that during economic and financial stress periods, such as commodities boom period and the COVID-19 pandemic, the interconnectedness between variables intensifies, leading to increased volatility spillovers. The Total Connectedness Index reveals significant fluctuations in risk within the monetary and credit markets, with peak values surpassing 30 percent and 75 percent. This index shows increased interconnectedness during financial stress, especially during major economic events like the COVID-19 pandemic, with notable peaks around April 2016 and May 2020. Analysis of directional spillovers indicates that fluctuations from the 10-year minus 2-year spread and the 2-year

LIBOR significantly impact the financial system, with important movements observed in 2018 and 2020. Additionally, certain variables, such as the interbank rate, mainly act as receivers of spillovers, while others, like the MSCI index, primarily contribute to movements in the market. This underscores the asymmetric nature of shock transmission, with some variables playing a larger role in spreading volatility than others. Furthermore, the interbank rate does not propagate its changes significantly, suggesting that monetary policy effects are more apparent over longer horizons, with daily credit showing limited responsiveness to the reference rate.The connectedness analysis emphasizes the need for continuous monitoring of volatility spillovers, particularly in times of financial uncertainty, to ensure the stability of the credit market and the broader economy.

A Appendix

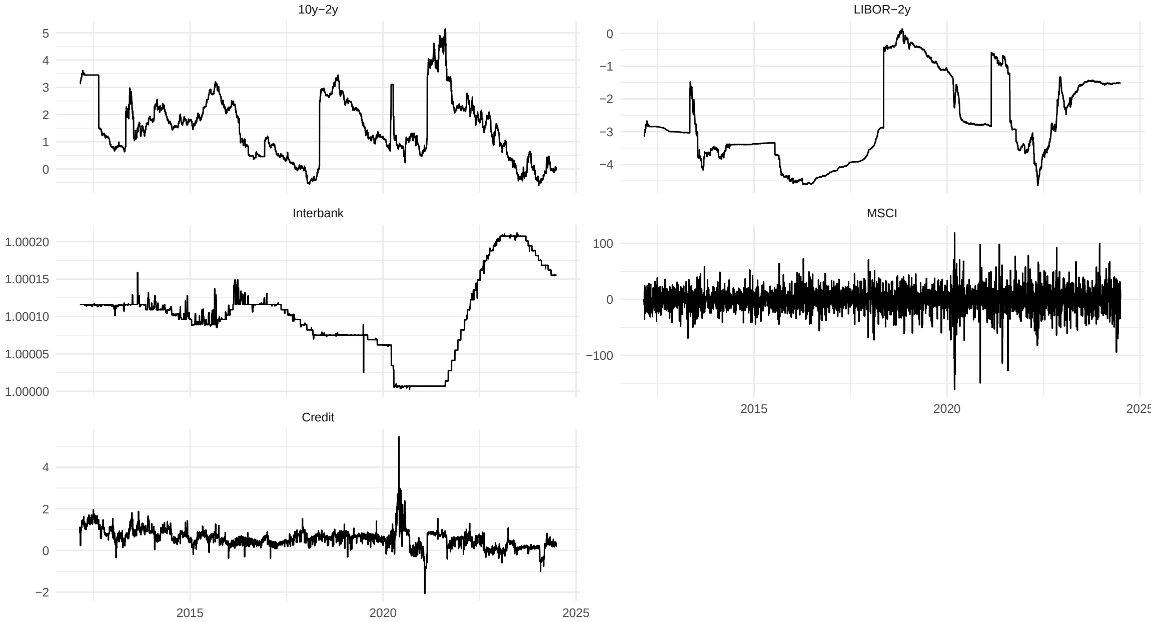

Figure 7: Data series.

Table 4: Descriptive Statistics

	10y-2y	10y-5y	LIBOR-2y	LIBOR-5y	Interbank	MSCI
Mean	1.538***	1.568***	-2.707***	-2.677***	1.000***	0.074
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.778)
Skewness	0.377***	0.769***	0.487***	-0.287***	0.065*	-0.478***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.076)	(0.000)
Ex. Kurtosis	-0.332***	-0.698***	-0.916***	-0.467***	0.004	8.444***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.908)	(0.000)
JB	127.739***	536.788***	336.187***	102.678***	3.140	13574.844***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.208)	(0.000)
ERS	-1.224	-2.542	-1.497	-0.037	-1.172	-29.359
	(0.221)	(0.011)	(0.134)	(0.970)	(0.241)	(0.000)
Q(20)	44846.599***	44552.364***	46435.384***	46528.767***	46884.764***	19.260**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.024)
Q2(20)	44053.968***	43487.481***	46635.403***	46464.937***	46884.763***	831.148***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Correlation matrix						
10y-2y	1.000***	0.364***	0.185***	-0.415***	-0.396***	-0.043***
10y-5y	0.364***	1.000***	-0.112***	0.183***	-0.394***	-0.012
LIBOR-2y	0.185***	-0.112***	1.000***	0.660***	-0.124***	-0.020
LIBOR-5y	-0.415***	0.183***	0.660***	1.000***	-0.021	0.010
Interbank	-0.396***	-0.394***	-0.124***	-0.021	1.000***	0.016
MSCI	-0.043***	-0.012	-0.020	0.010	0.016	1.000***

References

Adrian, T. & Shin, H. S. (2010), 'Liquidity and leverage', *Journal of financial intermediation* **19**(3), 418–437.

Antonakakis, N., Chatziantoniou, I. & Filis, G. (2017), 'Oil shocks and stock markets: Dynamic connectedness under the prism of recent geopolitical and economic unrest', *International Review of Financial Analysis* **50**, 1–26.

Avdjiev, S., Gambacorta, L., Goldberg, L. S. & Schiaffi, S. (2020), 'The shifting drivers of global liquidity', *Journal of International Economics* **125**, 103324.

Borio, C. (2014), 'The financial cycle and macroeconomics: What have we learnt?', *Journal of banking & finance* **45**, 182–198.

Breiman, L. & Breiman, L. (2001), 'Random forests', *null* .

Caballero, J. A. (2016), 'Do surges in international capital inflows influence the likelihood of banking crises?', *The Economic Journal* **126**(591), 281–316.

Claessens, M. S. & Kose, M. A. (2013), 'Financial crises explanations, types, and implications'.

Demirer, M., Diebold, F. X., Liu, L. & Yilmaz, K. (2018), 'Estimating global bank network connectedness', *Journal of Applied Econometrics* **33**(1), 1–15.

Diebold, F. X. & Yilmaz, K. (2012), 'Better to give than to receive: Predictive directional measurement of volatility spillovers', *International Journal of Forecasting* .

Giraldo, C., Giraldo, I., Gomez-Gonzalez, J. E. & Uribe, J. M. (2024), 'High frequency monitoring of credit creation: A new tool for central banks in emerging market economies', *Quarterly Review of Economics and Finance* .

Grimm, M., Jordà, Ò., Schularick, M. & Taylor, A. M. (2023), Loose monetary policy and financial instability, Technical report, National Bureau of Economic Research.

Hastie, T., Tibshirani, R., Friedman, J., Hastie, T., Tibshirani, R. & Friedman, J. (2009), 'Random forests', *The elements of statistical learning: Data mining, inference, and prediction* pp. 587–604.

Hong, M. G. H. & Lucas, D. (2023), *Evaluating the costs of government credit support programs during COVID-19: International evidence*, International Monetary Fund.

Kant, D., Pick, A. & de Winter, J. (2022), 'Nowcasting gdp using machine learning methods'.

Kiyotaki, N. & Moore, J. (1997), 'Credit cycles', *Journal of political economy* **105**(2), 211–248.

Ma, Y., Zhang, P., Duan, S. & Zhang, T. (2023), 'Credit default prediction of chinese real estate listed companies based on explainable machine learning', *Finance Research Letters* **58**, 104305.

Obstfeld, M. (2013), 'Finance at center stage: Some lessons of the euro crisis'.

Pesaran, H. H. & Shin, Y. (1998), 'Generalized impulse response analysis in linear multivariate models', *Economics letters* **58**(1), 17–29.

Schwaab, B., Koopman, S. J. & Lucas, A. (2014), 'Nowcasting and forecasting global financial sector stress and credit market dislocation', *International Journal of Forecasting* **30**(3), 741–758.

Trejo García, J. C., Martínez García, M. Á. & Venegas Martínez, F. (2017), 'Administración del riesgo crediticio al menudeo en méxico: una mejora econométrica en la selección de variables y cambios en sus características', *Contaduría y administración* **62**(2), 377–398.

Varian, H. R. (2014), 'Big data: New tricks for econometrics', *Journal of economic perspectives* **28**(2), 3–28.