Economic Systems xxx (xxxx) xxx

Contents lists available at ScienceDirect

Economic Systems

journal homepage: www.elsevier.com/locate/ecosys

Income inequality and systemic banking crises: A nonlinear nexus

Shengquan Wang

Institute of Advanced Studies in Humanities and Social Sciences, Beijing Normal University, Zhuhai 519087, China

ARTICLE INFO

JEL classification:
D3
E44
G20

Keywords:
Income inequality
Kuznets effect
Rajan effect
Systemic banking crises

ABSTRACT

Motivated by Rajan's work, we propose that income inequality and systemic banking crises have a nonlinear nexus. In addition to examining the linear "Rajan effect," we propose a "Kuznets effect" based on an assumption that income inequality has a nonlinear impact on growth, conditional on the stage of economic development, and thus plays a nonlinear role in modeling crises. We test the existence of this nexus using a sample covering 172 economies for a period of fifty years. We confirm that the relationship is U-shaped and identify the threshold level of income inequality that is beneficial for financial stability. Additionally, the U-test and area under the receiver operating characteristic (AUROC) statistics confirm that nonlinearity, and our model has good predictive performance of forecasting. Furthermore, the results of our panel regressions are consistent and are robust to several tests. We then identify the Rajan and Kuznets effects using a two-step test. We conclude that the impact of income inequality on the occurrence of systemic banking crises is U-shaped.

1. Introduction

Research on banking crises has a long history. Although such crises do not occur often, they are very costly and devastating (Angkinand and Willett, 2011). Many papers have probed the driving factors or early warning signals of systemic banking crises (Brunnermeier and Sannikov, 2014; Brunnermeier et al., 2020; Davis and Karim, 2008; Demirgüç-Kunt and Detragiache, 1998; Kim et al., 2013; Reinhart and Rogoff, 2009; Schularick and Taylor, 2012). During the financial crisis in 2008–2009, the United States experienced the largest rise in income inequality since World War II. Rajan (2010) was the first to propose that income polarization might contribute to the occurrence of financial crises, a view that has sparked intense debate (Coibion et al., 2020), though without reaching a consensus (Bordo and Meissner, 2012; Kirschenmann et al., 2016). We contribute to this debate by investigating the impacts of income inequality on banking fragility and, ultimately, systemic banking crises.

In addition to examining the "Rajan effect" (Rajan, 2010), in which income inequality monotonically causes crises by leading to credit booms, we innovative by formalizing a "Kuznets effect," which highlights the inequality-growth-crisis mechanism. In particular, following the literature, we argue that income inequality has a nonlinear effect on growth, conditional on the stage of economic development. Thus, income inequality could have a nonlinear effect on economic output (in a way that is similar to that of the Kuznets curve) and thus the occurrence of banking crises.

Furthermore, we contribute to the literature by empirically confirming whether income inequality could drive systemic banking crises, especially in a nonlinear manner. To our knowledge, our sample covers a larger number of countries than other related studies; our panel dataset includes 172 economies and covers the period 1970–2017. We employ a binary choice model and a linear probability model (LPM) to investigate the effect of income inequality on the incidence of systemic banking crises. We add a nonlinear term in the econometric model to identify the potential nonlinear effects.

E-mail address: wangshq@bnu.edu.cn.

https://doi.org/10.1016/j.ecosys.2023.101123
Received 11 October 2022; Received in revised form 9 July 2023; Accepted 10 July 2023
Available online xxxx

Please cite this article as: S. Wang, Income inequality and systemic banking crises: A nonlinear nexus, Economic Systems, https://doi.org/10.1016/j.ecosys.2023.101123

Econometrically, a significant quadratic term is insufficient for confirming a nonlinear relationship, so we do so with the U-test proposed by Lind and Mehlum (2010). Furthermore, we use an efficient criterion—the area under the receiver operating characteristic (AUROC) curve—to evaluate the model's predictive performance. Moreover, we conduct robustness checks—including adding control variables, considering the inertia of banking crises, filling in missing data, reducing noise in the annual data, distinguishing the "level" and "growth" hypotheses, applying the rare event strategy, and assessing the bias from unobservable quantities—and the results are in line with the baseline results. The Rajan and Kuznets effects are examined, and our proposed hypotheses are confirmed.

This paper is related to the strands of literature on the economic effects of income distribution and on the determinants or formation mechanism of systemic banking crises. Income distribution has been widely discussed, especially since the 2008–2009 financial crisis (Gertler and Klenow, 2019), with studies focusing on the economic effects of income distribution, such as economic growth (Acemoglu and Robinson, 2002; Grossman and Helpman, 2018; Kuznets, 1955), technological innovation (Foellmi and Zweimüller, 2006; Greenwood and Mukoyama, 2001), and current account imbalance (Behringer and van Treeck, 2018; Belabed et al., 2018; Kumhof et al., 2012). Many factors contribute to the development of systemic banking crises. In an influential work, Diamond and Dybvig (1983) model self-fulfilling bank runs. The 2008–2009 financial crisis showed that systemic financial instability is affected by the bursting of asset bubbles (Aoki and Nikolov, 2015; Brunnermeier et al., 2020), the "wind against" monetary policy (Gali, 2014), bank credit booms (Jordà et al., 2015; Schularick and Taylor, 2012), financial innovation (Beck et al., 2016; Kim et al., 2013), and the business cycle (Brunnermeier and Sannikov, 2014).

Only recently have studies addressed these two strands of the literature simultaneously. In a pioneering study, Rajan (2010) hypothesizes that income disparity is a structural driver of systemic financial crises and argues that income inequality creates credit booms, which destabilize the financial system. Like Rajan (2010), Reich (2010) finds that income inequality negatively affects financial stability. The concept of an inequality–credit nexus can be traced to Krueger and Perri (2006) and Iacoviello (2008), who argue that the expansion of income inequality increases the demand for insurance through the credit market, implying that households will increase their leverage. Recent studies have confirmed this relationship (Fischer et al., 2019; Malinen, 2016; Wang and Luo, 2023).

The literature on the nexus between income inequality and the occurrence of financial crises has not reached a consensus. In the first empirical investigation of this nexus, Bordo and Meissner (2012) find no evidence that income inequality begets credit booms and, then, financial crises. Furthermore, Morelli and Atkinson (2015) find no conclusive evidence to support Rajan's hypothesis. In contrast, other authors argue that income inequality is a powerful predictor of financial crises (Kirschenmann et al., 2016; Perugini et al., 2015; Rhee and Kim, 2018).

The first theoretical model of the nexus between income inequality and the occurrence of banking crises was built by Kumhof et al. (2015); their model formalizes the Rajan hypothesis and shows that higher household debt and banking crises are the endogenous results of widening income inequality. Moreover, Cairó and Sim (2018) construct a general equilibrium model in which income inequality increases credit growth by distributing income to the rich, who have a low marginal propensity to consume, causing an endogenous financial crisis.

Our hypothesis is correlated with the financial instability hypothesis by Hyman Minsky (1983) but with significant differences. The Minsky hypothesis states that a capitalist economy endogenously generates a financial structure that is susceptible to financial crises (Minsky, 1983), in which a financial structure is "the market interactions between borrowers and lenders and the balance sheets of non-financial firms, intermediaries and households that reflect these interactions" (Pollin, 1994). Minsky highlights that financial crises are the consequence of the systemic inability of firms in the financial sector to repay their debt. Our analysis is related to the Minsky hypothesis, which has been generalized to theories of endogenous financial instability and sometimes refers to a "Minsky moment" when the financial system collapses (Bartscher et al., 2020). Our analysis offers an explanation of financial collapse but differs from Minsky's hypothesis in two ways. First, rather than taking the firm perspective, we follow Rajan (2010) and explain financial instability from the perspective of households. Second, our analysis pays less attention to innate instability in a financial system and, instead, highlights income inequality as a driver of financial crises.

The remainder of the paper is structured as follows. In Section 2, we develop a theoretical hypothesis to formalize the nexus between income distribution and systemic banking crises, proposing the Rajan and the Kuznets effects. Section 3 presents the econometric approach and model evaluation methods. Section 4 describes the variables and tests the stationarity of the data and the dynamics of explanatory variables related to financial crises. The empirical results are presented, including cross-sectional evidence in Section 5 and panel evidence in Section 6. Section 7 concludes the paper.

2. Hypotheses development

We emphasize two direct factors, credit booms and economic growth, and a deep and structural factor, income inequality.

Credit booms are related to systemic banking crises (Kumhof et al., 2015; Reinhart and Rogoff, 2011; Schularick and Taylor, 2012), but the micro-foundation of this relationship has rarely been analyzed. In our framework, we argue that a bank's issuance of credit might increase the probability of borrower default, under the assumption that credit booms are likely a by-product of operational or regulatory failure in the financial system (Schularick and Taylor, 2012). Conditional on an unchanged economic environment and costly oversight, the oversupply of credit inevitably comes at the expense of relaxing credit requirements. Hence, borrowers' credit quality is not assured. When bank have limited liability and deposits are insured, bankers have more incentive for taking on excessive risk because they can obtain a risk premium from risk-taking (Aoki and Nikolov, 2015). Hence, credit booms increase default by borrowers, thus increasing the emergence of systemic banking crises.

Furthermore, we consider the reasons for credit booms, which might be significantly driven by income inequality. Rajan (2010) suggests that in the US the expansion of lending to low-income households (which are more likely to default) was a political response to rising inequality. Concomitantly, financial institutions (which represent the credit supply) were motivated to take on excessive risk in order to maximize profit. Meanwhile, consumers (representing the demand side) were happy to take out loans because the benefits (increased consumption) were immediate whereas the inevitable repayment could be postponed. According to Rajan (2010), as income inequality increases, the supply of credit and the demand for credit both expand, which essentially leads to credit booms. Many previous studies, such as Rhee and Kim (2018), Kumhof et al. (2015), and Bordo and Meissner (2012), are explicitly related to Rajan (2010), with the causal relation running from income inequality to credit booms.

Different from Rajan (2010), other mechanisms of income inequality and credit booms may exist. Iacoviello (2008) links income inequality to financial instability and argues that income inequality was the main driver of household debt booms in the United States during the 1980 s and 1990 s Stiglitz (2012) stresses that the rich may have incentives for loosening financial regulations and oversight to maximize their profit, thus enabling credit booms and financial vulnerability. As the rich are politically prioritized over the poor (Mitkov, 2020), they may influence financial regulations by pressuring (or lobbying) oversight agencies to increase their profit.

Although Rajan presents his ideas in the US context, many studies extend them to other developed and developing countries. For example, Rhee and Kim (2018) analyze a sample of 68 developing countries, whereas Bordo and Meissner (2012) use a sample comprising 14 advanced countries. They both find that income inequality causes credit booms.

Accordingly, we propose Hypothesis 1.

Hypothesis 1. (-Rajan Effect): Income distribution disparity can monotonically undermine banking stability by causing credit booms.

Do other mechanisms of income inequality affect financial stability? The Rajan effect assumes that the rate of economic growth remains unchanged as income inequality increases. But this assumption might not be reasonable if income inequality is correlated with financial stability.

First, economic growth is a fundamental determinant of financial stability. The underlying intuition is that the asset and liability sides of financial institutions are both largely dependent on economic growth. When the economy slows or trends downward, the liabilities of financial institutions will decline because of reductions in deposits, which can drive financial institutions into a liquidity crisis; returns shrink, partly because of the depreciation in asset prices and partly because of a rise in the risk of loan defaults. Demirgüç-Kunt and Detragiache (1998) argue that banking crises tend to erupt when growth of output is slow, and that low output growth is among the best early warning signals of such crises. Moreover, von Hagen and Ho (2007) show that a slowdown in real economic output precedes banking crises, and Davis and Karim (2008) arrive at a similar conclusion.

Second, many papers have investigated the effects of income inequality on economic growth, however, no consensus has been reached in the theoretical and empirical predictions (Lin et al., 2009, 2014). An increasing number of papers have tried to reconcile prior studies by modeling a nonlinear effect of income inequality on growth, conditional on the stage of economic development. Galor and Tsiddon (1997), Galor (2000), and Galor and Moav (2004) argue that, during the early stage of economic development driven by physical capital, income inequality could stimulate economic growth by shifting resources from the high-income because of their higher marginal propensity to save than the low-income; when economic development enters the stage driven by human capital, income inequality becomes harmful for growth because of credit constraints on human capital accumulation. Bandyopadhyay and Basu (2005) agree that the inequality-growth nexus is conditional on the nature of technology and the extent of redistribution. If knowledge spillover faces low barriers, technology is highly skill intensive, and the degree of redistribution is high, so income inequality will have a positive effect on growth; otherwise, the effect will be negative. If the nature of technology and the extent of redistribution change, the sign is reversed. Empirically, Lin et al. (2009, 2014) confirm a nonlinear nexus between income inequality and economic growth depending on the development level.

Therefore, a nonlinear nexus may exist between income inequality and financial stability because of the nonlinear effect of income inequality on economic growth.

Accordingly, we present the following hypotheses:

Hypothesis 2. (Kuznets effect): The nexus between income inequality and the occurrence of systemic banking crises is nonlinear because of the nonlinear effect of income inequality on economic growth.

Hypothesis 3. Synthetically, disparity in income distribution is nonlinearly correlated with the occurrence of systemic banking crises, considering both the Rajan and Kuznets effects.

H3 is illustrated in Fig. 1.

However, the combination of a linear effect and a nonlinear effect does not necessarily produce a nonlinear effect within a reasonable period. This implies that, in our analysis, a nonlinear Kuznets effect does not guarantee that the nexus between income inequality and the likelihood of systemic banking crises will be nonlinear in our sample. We use an empirical test to determine whether the nexus is nonlinear and whether the proposed Rajan and Kuznets effects exist.

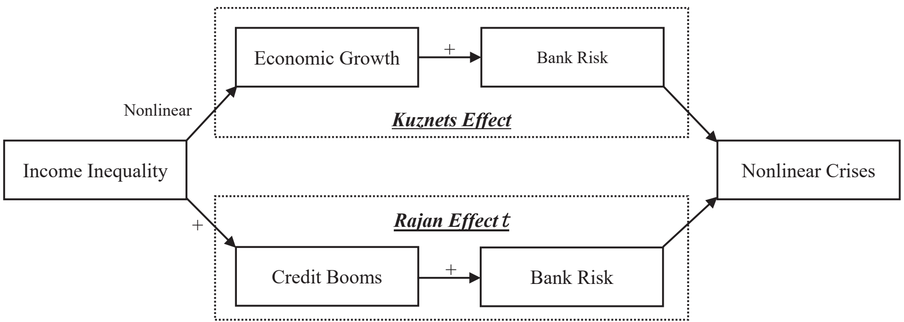

Fig. 1. Implicit mechanism of the nonlinear nexus between income inequality and the occurrence of systemic banking crises.

3. Methodology

3.1. Econometric model

The dependent variable—systemic banking crises—is binary. Thus, it is natural to use a discrete choice model, rather than a model designed for continuous dependent variables. In addition to constructing a logit model, we also report the results using the LPM. Our logit model is similar to the one employed by Schularick and Taylor (2012) and Kirschenmann et al. (2016).

To obtain multidimensional evidence on the effect of fluctuations in income inequality on the incidence of banking crises, we conduct cross-sectional regression (which can provide direct evidence) and panel two-way fixed-effects regression. The cross-sectional regression is a simplified version of the panel model, thus we introduce only the panel model used in this paper, as follows:

$$y_{it} = \begin{cases} 1, if \quad there \quad is \quad a \quad banking \quad crisis \quad in \quad economy \quad i \quad in \quad year \quad t \\ 0, \quad otherwise \end{cases} \tag{1}$$

$$y_{it} = \beta_I {}^{*}Inequality_{it-1} + \beta_{II} {}^{*}Inequality_{it-1}^2 + x_{it-1}'\beta_x + \mu_i + \delta_t + \eta_{it} \tag{2}$$

$$\text{logit}(y_{it}) = \beta_I {}^{*}Inequality_{it-1} + \beta_{II} {}^{*}Inequality_{it-1}^2 + x_{it-1}'\beta_x + \mu_i + \delta_t + \eta_{it} \tag{3}$$

where i is the economy, respectively, $i = 1, .., 172$, and t is the year $t = 1970, .., 2017$. y_{it} is a dummy variable that equals 1 if economy i suffers a banking crisis in year t. Eqs. (2) and (3) are the LPM and the logit model, respectively. $Inequality_{it-1}$ is the variable of interest and measures the degree of inequality of income distribution, and the corresponding coefficients β_I and β_{II} represent the effect of income inequality on the probability of a banking crisis. x_{it-1} is a matrix of predictors, and β_x is the corresponding coefficients vector. μ_i and δ_t are the economy fixed effect and the time fixed effect, respectively. η_{it} is the error term and is assumed to be well behaved.

3.2. Model evaluation

It is necessary to assess whether the proposed model can effectively predict banking crises. Following the approach by Anundsen et al. (2016), we evaluate two types of errors: those in which the model fails to predict a crisis (Type I error) and those in which the model erroneously signals a crisis (Type II error).

Using model m, when the estimated crisis incidence $\hat{p}_m$ exceeds a given threshold τ, the model sends a crisis signal. The true positive rate ($TPR_m(\tau)$) is defined as the ratio of correctly identified crises to all crises and equals one minus the probability of a Type I error. The false-positive rate ($FPR_m(\tau)$), the probability of a Type II error, is the ratio of incorrectly identified crises to all noncrises. The optimal threshold depends on the policy maker's preference regarding the trade-off between Type I and Type II errors. The linear loss function ($L_m(\bullet)$) is applied to determine the optimal threshold value:

$$L_m(\theta, \tau) = \theta p(1 - TPR_m(\tau)) + (1 - \theta)(1 - p)FPR_m(\tau) \tag{3}$$

where p is the sample probability of a banking crisis; θ is the relative weight assigned by the policy maker to undetected (i.e., missing) crises, and $\theta \in [0.5,1]$. We follow Anundsen et al. (2016) and set θ at 0.90 and 0.95, which implies that the policy maker is most interested in not missing any banking crises. Furthermore, we evaluate the usefulness of the proposed model using a statistic called *relative usefulness* ($U_m^r(\tau)$):

$$U_m^r(\tau) = \frac{\min\{\theta p, (1 - \theta)(1 - p)\} - L_m(\theta, \tau)}{\min\{\theta p, (1 - \theta)(1 - p)\}} \tag{4}$$

For a perfect model, $U_m^r(\tau) = 1$, whereas for a useless model, $U_m^r(\tau) = 0$. The model is useful if the numerator is positive, that is, if the loss related to the model is less than that without using a model.

We also evaluate the model using the AUROC (Eq. (5)), a widely used evaluation criterion for binary econometric models in other fields (e.g., survival analysis in medicine, the signal detection in military affairs and psychology), and, of course, in forecasting financial crises (Herrera et al., 2020; Schularick and Taylor, 2012). The ROC curve plots all possible combinations of $TPR_m(\tau)$ and $FPR_m(\tau)$ across different values of τ.

$$AUROC_m = \int_{\tau=0}^{1} TPR_m(FPR_m(\tau))FPR_m'(\tau)d\tau \tag{5}$$

For a perfect model, AUROC = 1, and, for a useless model, AUROC = 0.5.

We hypothesize that the impact of income inequality on the occurrence of systemic banking crises is nonlinear. Most studies in which a significant quadratic term is added to an otherwise classical linear regression model find a U-shaped relationship between the variables of interest. However, Lind and Mehlum (2010) argue that this approach is too weak when the true nexus is convex but monotone and develop a U-test statistic to check the nonlinearity of the relationship at a reasonable interval. A U-shaped relationship can be confirmed by the following condition:

$$\beta_I + \beta_{II} {}^{*}Inequality_{it-1}^l < 0 < \beta_I + \beta_{II} {}^{*}Inequality_{it-1}^h \tag{6}$$

Here, $[Inequality_{it-1}^l, Inequality_{it-1}^h]$ is the interval of interest. In our case, this interval is the observed data range. If these inequalities are violated, the curve will take an inverted U shape or be monotonic. Thus, the U-test essentially examines whether the null hypothesis (H_0) can be rejected by the following alternative hypothesis (H_1):

$$H_0\text{: } \beta_I + \beta_{II} {}^{*}Inequality_{it-1}^l \geq 0 \; and/or \quad \beta_I + \beta_{II} {}^{*}Inequality_{it-1}^h \leq 0 \tag{7}$$

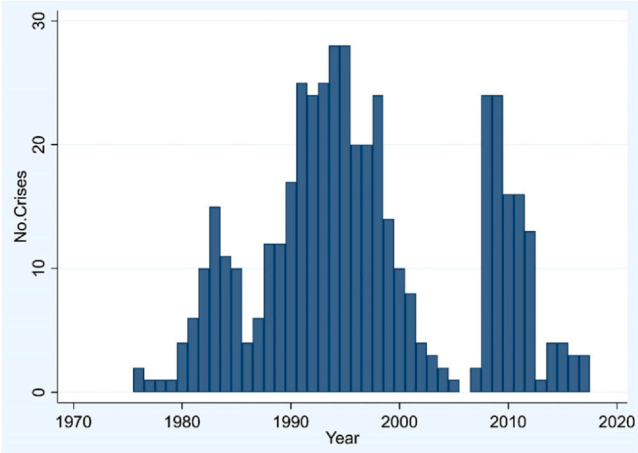

Fig. 2. Temporal distribution of the systemic banking crises in the sample.

$$H_1 : \beta_I + \beta_{II} * Inequality_{it-1}^{l} < 0 \text{ and } \quad \beta_I + \beta_{II} * Inequality_{it-1}^{h} > 0 \tag{8}$$

The U-test offers a more robust approach for checking the nonlinearity of a relationship compared with simply focusing on the significance of the coefficient of the quadratic term added.

4. Data description and temporal properties

We collect annual panel data from diverse sources to cover as many economies as possible. Our final sample covers 172 economies and spans the period 1970–2017 (for a complete list of the economies in the sample, see Appendix Table A1).

4.1. Systemic banking crises

We define a systemic banking crisis with the criteria by Laeven and Valencia (2013) as follows:

1. The banking system shows signs of financial distress (as indicated by significant bank runs, losses in the banking system, and bank liquidations).
2. Significant banking policy intervention measures are taken in response to significant losses in the banking system.

Fig. 2 plots the temporal distribution of the systemic banking crises in our sample. More intensive banking crises occurred in the last decade of the twentieth century than in other periods—for example, the European monetary system crisis (1992–1993), the Mexican financial crisis (1994–1995), and the Asian financial crisis (1997–1998). Furthermore, 24 economies suffered a systemic banking crisis during the global financial crisis in 2008–2009.

4.2. Explanatory variables

As shown by Kauko (2014), the literature discusses many factors that could drive banking crises. For simplicity, we include only income inequality and its quadratic term in the baseline model; however, in robustness checks, we investigate the effects of other factors chosen based on their importance and data availability. In the following discussion, we present the sources and measurements of the explanatory variables used in the model.

4.2.1. Income inequality

In light of the hypothesis development and the subsequent nonparametric evidence presented in Section 5.1, we add income inequality and its quadratic term in the regression to prove the existence of a nonlinear nexus between income distribution and the occurrence of banking crises. We measure income inequality in two ways: the Gini index (a widely used measure of income distribution in a population across income percentiles) and the share of the income contributed by the 10% of the population with the highest income. The data are collected from the World Bank's World Development Indicator (WDI) database and the Standardized World Income Inequality Database (SWIID).[1] The WDI offers data on the Gini index and the top 10% income share, and SWIID offers another dataset on the Gini index. Although both WDI and SWIID offer the Gini index, they are different. The Gini index from the WDI is collected or calculated using income/tax files, so many observations are missing. The

[1] https://fsolt.org/swiid/.

Gini index from SWIID is re-estimated based on the Luxembourg Income Study (LIS), which fills in some missing data and maximizes the comparability of income inequality across countries and years.[2]

4.2.2. Credit

Credit is widely considered as one of the most important drivers of economic crises (Reinhart and Rogoff, 2008; Schularick and Taylor, 2012). However, the multicollinearity between income inequality and credit may lead to misleading results; thus, we explicitly analyze only credit to identify the mechanism underlying the relationship between income inequality and credit. The ratio of credit to nonfinancial sectors in the gross domestic product (GDP) is used in the regressions. Credit data are obtained from the website of the Bank for International Settlements.[3]

4.2.3. Economic growth

Economic growth drives banks' fundamental risk. The effect of economic growth on the occurrence of systemic banking crises is linear (Davis and Karim, 2008; Demirgüç-Kunt and Detragiache, 1998; von Hagen and Ho, 2007), whereas the nonlinear Kuznets effect might shape the interaction between income inequality and economic growth. Thus, economic growth is added to the baseline model, whereas economic growth and its quadratic term are applied to identify the underlying mechanism. Economic growth data are obtained from the WDI database.

4.2.4. Current account

Many studies find that most economies run current account deficits in the leadup to banking crises (Lo Duca and Peltonen (2013); Reinhart and Rogoff, 2008). Accordingly, we add the current account balance as a ratio of GDP to the econometric model. We obtain these data from the WDI database.

However, Belabed et al. (2018) suggest that income distribution can linearly explain current account developments, which implies that a multicollinearity issue will arise. Thus, below, we report the results both with and without consideration of current account data.

4.2.5. Others

Price instability may increase the risk of financial instability, especially in emerging economies (Angkinand and Willett, 2011; Lo Duca and Peltonen (2013)). At high inflation levels, banks worry that their main sources of revenue will disappear after inflation decreases sharply, and systemic bank default might follow. Thus, the GDP deflator, as a proxy for inflation, is included in the regressions. This deflator is obtained from the WDI database.

Fiscal deficits can predict the occurrence of banking crises (von Hagen and Ho, 2007). In this study, a fiscal deficit is operationalized using the government debt balance as a ratio of GDP. These data are obtained from the WDI database.

Deposit insurance can lead to excessive risk-taking by bank managers and thus affect the probability of banking crises (Beck et al., (2010)). Data on deposit insurance are obtained from the World Bank Group's Deposit Insurance Dataset.[4]

4.3. Descriptive statistics

Table 1 shows the descriptive statistics of the variables examined in this study. We winsorized the sample at the 1% quantile, for use in our analysis. The probability of a banking crisis in our sample is 5.77%. The three measures of income inequality—GINI (WDI), 10%Top, and GINI (SWIID)—average 38.931, 29.8388, and 37.4232, respectively; these values are used as the benchmarks in the subsequent analysis in which we compare our results with those in the literature. Average growth of credit is 38.6093%, and average economic growth is 3.7632%; these two variables are used as mediating variables in the subsequent analysis. The other variables follow patterns comparable to those reported in previous studies.

4.4. Data stationarity

Data stationarity is important for ensuring the validity of econometric inferences, even for binary choice models (Park and Phillips, 2000). Thus, to ensure the reliability of our inferences, we perform unit-root testing of the panel data. Because our sample is unbalanced and contains missing data, we can only perform a Fisher-type test (Choi, 2001) for the five control variables. The results are shown in Table 2. The time series of all five control variables reject the null hypothesis, which implies that our sample is stationary and reliable.

4.5. Dynamics around crisis episodes

To further analyze the determinants of systemic banking crises, we investigate the behavior of the variables around a crisis. Following Gourinchas and Obstfeld (2012), we regress the following linear model to document the expectation of the variable related to the distance from a crisis.

$$x_{ijt} = \alpha_{ij} + \beta_{js}\delta_{ijs} + \varepsilon_{ijt} \tag{9}$$

[2] For a detailed explanation of SWIID, see Solt (2020).

[3] https://www.bis.org/statistics/totcredit.htm?m= 6%7C380%7C669/.

[4] https://datacatalog.worldbank.org/dataset/deposit-insurance-dataset/.

Table 1
Descriptive statistics of the variables.

Variable	Mean	Std. Dev.	Min.	Max.	Obs.
Crises	0.0577	0.2332	0	1	7938
GINI(WDI)	38.9310	9.0783	24.8000	59.6000	1238
10%Top	29.8388	8.0231	0	47.5000	1263
GINI(SWIID)	37.4232	7.8570	21.9000	53.3000	4592
Creditgdp	38.6093	36.0063	1.4823	170.7510	6614
Curaccount	-3.3173	9.1015	-33.2802	28.7125	5595
M2growth	19.8134	24.5707	-14.1329	165.2856	5983
Govdebt	54.9221	34.0554	5.0390	187.7810	1256
Capitalgrowth	5.6366	17.1669	-41.5786	73.4901	4884
ExplicitDGS	0.2915	0.4545	0	1	7824
GDPgrowth	3.7632	4.9520	-12.9121	20.1554	6935
GDPdeflator	13.3685	30.8371	-10.9760	233.0309	6933

Table 2
Results of the Fisher-type unit-root test based on the augmented Dickey–Fuller (ADF) test.

Variables	Inverse χ^2	Inverse Normal	Inverse Logit t	Modified Inverse χ^2
GDP growth	1989.5214 ***	-34.6892 ***	-41.8501 ***	63.2563 ***
	(0.000)	(0.0000)	(0.0000)	(0.0000)
Creditgdp	798.7787 ***	-13.3740 ***	-13.7625 ***	17.8521 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Currentgdp	1175.2035 ***	-22.9687 ***	-24.5652 ***	33.0777 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Govdebtgdp	262.5870 ***	-8.6293 ***	-8.9175 ***	10.9960 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Gdpdeflator	1701.2775 ***	-29.7996 ***	-35.5801 ***	52.2026 ***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Notes: *p*-values are in parentheses. * ** significance at 1%.

where j is the j^{th} variable, and s indicates the number of years between the current year and the year of a systemic banking crisis. In our analysis, s is from –5–5, that is, we evaluate the behavior of key variables in the five years preceding and after a crisis. x_{ijt} represents the variables of interest, which include income inequality, the ratio of credit to GDP, GDP growth, the ratio of current account balance to GDP, the ratio of government debt to GDP, and the GDP deflator. δ_{ijs} is a dummy variable that equals one when a crisis occurs and, otherwise, zero. ε_{ijt} is the disturbance term, and $\varepsilon_{ijt} \sim i.\ i.\ d.\ N(0, \sigma_{xi}^2)$. α_{ij} is the fixed effect. β_{js} is the parameter of interest and measures the conditional impact of being s years away from a banking crisis on the expectation of x_{ijt} relative to "normal times"—that is, all country-year observations outside the event window.

Fig. 3 plots the behavior of income inequality, the ratio of credit to GDP, GDP growth, the ratio of the current account balance to GDP, the ratio of government debt to GDP, and the GDP deflator around crisis episodes. The three measures of income inequality (the first row in the figure) tend to be significantly higher in the two years before a crisis than in normal times, which is consistent with the view that income inequality seeds systemic banking crises (Cairó and Sim, 2018; Kirschenman et al., 2016; Kumhof et al., 2015; Rajan, 2010). The income inequality measures (the first two figures) from the WDI database (the Gini index and 10%Top) are approximately twenty percentage points higher at the peak value of income inequality (two years preceding the crisis) than in normal times; the Gini index from the SWIID database shows a persistent increase during the same period and is significantly higher than in normal times. Furthermore, the income distribution remains significantly worse in the five years after a crisis than in normal times, which implies that banking crises erode the power of labor and increase income inequality (Dufour and Orhangazi (2016)).

The ratio of credit to GDP (the first figure in the second row) continuously increases from five years before a crisis until two years after the crisis, and the ratio of credit to GDP is ten percentage points higher around crisis episodes than in normal times. It is widely accepted that credit booms can contribute to financial instability, and the mechanism underlying this effect is thought to be one in which credit booms are typically accompanied by excessive risk-taking by financial institutions. For example, Jordà et al. (2011) find that credit is the strongest driver of financial instability.

As shown in Fig. 3, a downturn in economic growth (the second figure in the second row) can predict the risk exposure to banking crises (Davis and Karim, 2008; Demirgüç-Kunt and Detragiache, 1998; von Hagen and Ho, 2007). When economic growth is at its peak (one year before the crisis), economic growth is about seven percentage points lower than in normal times. However, economic growth is seventeen percentage points lower in the two years after a crisis than in normal times.

A current account imbalance (or external imbalance) can increase financial instability (Jordà et al., 2011) and is widely considered to have been a driving factor in the 2008–2009 financial crisis (Belabed et al., 2018; Obstfeld and Rogoff, 2009). Consistent with prior studies, in our sample, there is a current account deficit, and the ratio of the current account balance to GDP (the third figure in the second row) decreases when a crisis is approaching. The ratio of the current account balance to GDP is about ten percentage points lower in the year before a crisis than in normal times but insignificantly, at the 95% confidence interval.

The third row of Fig. 3 plots the dynamics of the ratio of government debt to GDP, the GDP deflator, and deposit insurance around crisis episodes. As shown in the first graph, the ratio of government debt to GDP tends to decrease significantly before crises, which implies that

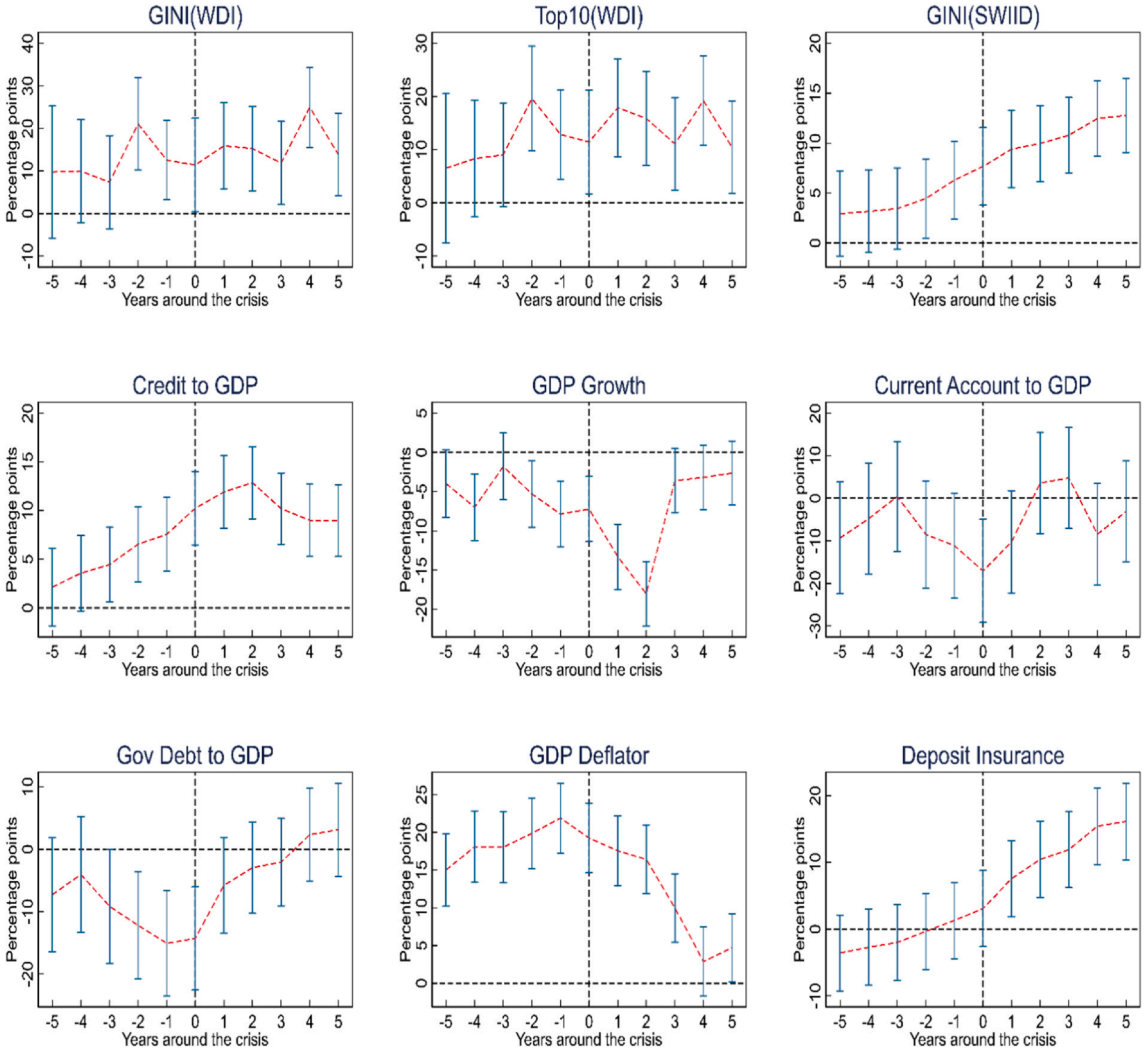

Fig. 3. Behavior of the explanatory variables around banking crisis episodes. *Notes*: The dashed lines are the conditional effects of being 5 to − 5 years away from a crisis, while the vertical bars show the corresponding 95% confidence intervals. A value different from zero means that the variable takes values that deviate from those in "normal times," defined as all country–year observations that do not fall within the event window.

government debt could predict banking crises (Lo Duca and Peltonen (2013); Reinhart and Rogoff, 2008). The GDP deflator is about twenty percentage points higher in the year before a crisis than in normal times, indicating that this variable may be useful in predicting crises (Demirgüç-Kunt and Detragiache, 1998). As shown in the third graph, the number of economies offering deposit insurance increases over time. Specifically, this number exceeds the usual number one year before a crisis, which suggests that deposit insurance increases financial risk (Beck et al., (2010)).

5. Empirical analysis: cross-sectional evidence

We first report the empirical evidence obtained using pooled data. The results have some referential value by providing direct evidence on the nexus between income inequality and the incidence of banking crises. To confirm the existence of a nonlinear relationship between these two parameters, we first plot the nonparametric relationship between income inequality and banking crises.

5.1. Nonparametric analysis

Unlike a regression analysis in which the shape of the relation between the variables of interest is specified, we assume that no prior information is available about the shape of the relationship between income inequality and banking crises. Thus, we perform the following nonparametric regression:

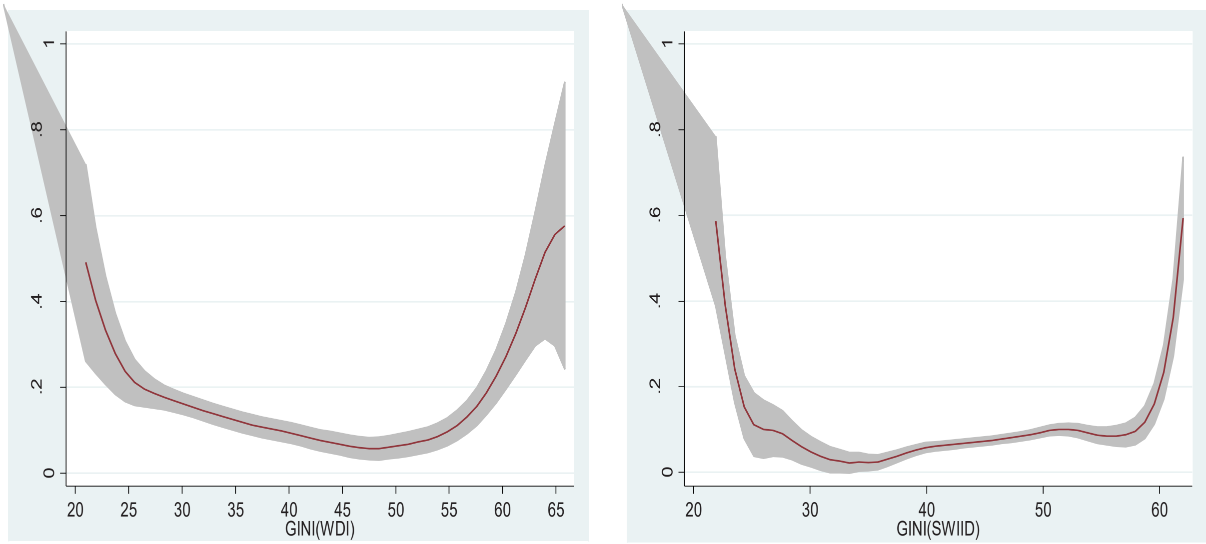

Fig. 4. A nonparametric plot between income inequality and the occurrence of banking crises. Notes: The shaded area is the 95% confidential band. The kernel function is Epanechnikov. The bandwidths are 6.8 and 3.56, respectively.

$$y_{it} = f(Inequality_{it-1}) + v_{it} \tag{10}$$

In this regression, the Epanechnikov kernel function is applied and the maximum degree of the polynomial is set at 2. The nonparametric plot is shown in Fig. 4. To cross-validate the nonlinear relationship, we use the Gini index from two sources, and present two plots. As shown in Fig. 4, the U-shaped nexus between income inequality and banking crises is significant within a 95% confidence interval. This primarily offers evidence of our nonlinear hypothesis. Thus, the subsequent model specification with a quadratic term is reasonable. The average optimal Gini index is approximately 0.4. When the income distribution is skewed to the left of the threshold, the probability of banking crises decreases from 0.5 or 0.6 to nearly 0 as inequality increases, and the opposite is true when the income distribution is skewed to the right. To summarize, the nonparametric analysis supports the existence of a nonlinear nexus between income distribution and the occurrence of banking crises, which mirrors the hypotheses in Section 2, and the model specification in the following regression analysis.

Furthermore, the plot in Fig. 4 might illustrate why prior empirical studies obtained different results. The figure shows a transitional period in which the probability of banking crises is generally stable as income inequality changes. Hence, the heterogeneity in the findings of prior studies may be a result of differences in their samples. Specifically, researchers who use samples in which the variables in Fig. 3 skew rightward might find that income inequality can drive financial instability (Kirschenmann et al., 2016), whereas others who use left-skewing samples might find no correlations between inequality and banking crises (Bordo and Meissner, 2012). Our large sample avoids sample discrimination issues and can thus indicate the implicit and real impacts of income distribution on the occurrence of financial crises. A rigorous comparison is presented in Section 6.3.

5.2. Regression analysis

For a more accurate analysis, we regress income inequality on the occurrence of systemic banking crises. The results based on the logit model and the LPM are shown in Table 3. Income inequality is measured by the Gini index collected from the WDI database in Columns (1) and (2) and scaled by the income share of the 10% of the population with the highest income (obtained from the same database, referred to hereafter as the "top 10% income share") in Columns (3) and (4). Columns (5) and (6) use the standardized Gini index from the SWIID database as a proxy for income inequality.

The coefficients of lagged Gini and its quadratic term are all significant at the 99% confidence interval. Additionally, the former is negative while the latter is positive, which implies that the relationship between Gini and systemic banking crises follows a quadratic function with a positive coefficient of the quadratic term. The threshold Gini for financial stability in the logit model is calculated at 0.39 (=[–(–0.2211)/(2 × 0.0028)]/100 in Column (5)) and 0.45 (=[–(–0.3164)/(2 × 0.0035)]/100 in Column (1)), which is consistent with the nonparametric analysis in Section 5.1. The relationship between the Gini and banking crises also holds with the LPM, and the threshold Gini is 0.46 (=[–(–0.0368)/(2 × 0.0004)]/100 in Column (2)) and 0.48 (=[–(–0.0190)/(2 × 0.0002)]/100 for Column (6)). These results suggest that income inequality, as measured by Gini, plays a significant role in driving systemic banking crises, and their relationship takes a U shape.

For a deeper analysis, we determine the marginal effect of Gini on the incidence of systemic banking crises. Because of the quadratic term, the marginal effect varies with Gini. We calculate the marginal effect and its corresponding standard error as follows:

$$Marginal \quad Effect = \frac{\partial(\Pr(y_{it} = 1))}{\partial(L.\ GINI)} = \beta_1 + 2\beta_2(L.\ GINI) \tag{11}$$

Table 3
Impact of income inequality on the occurrence of systemic banking crises: Cross-sectional evidence.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
L.GINI(WDI)	-0.3164 * **	-0.0368 * **				
	(0.0849)	(0.0104)				
L.GINI2(WDI)	0.0035 * **	0.0004 * **				
	(0.0010)	(0.0001)				
L.10%Top(WDI)			-0.0364	-0.0064		
			(0.0317)	(0.0049)		
L.10%Top2(WDI)			0.0001	0.0001		
			(0.0007)	(0.0001)		
L.GINI(SWIID)					-0.2211 * *	-0.0190 * *
					(0.0911)	(0.0086)
L.GINI2(SWIID)					0.0028 * **	0.0002 * *
					(0.0010)	(0.0001)
Cons.	4.6367 * **	0.8988 * **	-1.0332 * *	0.2606 * **	1.6532	0.4409 * *
	(1.6265)	(0.2069)	(0.4069)	(0.0736)	(2.0752)	(0.1932)
Obs.	1211	1211	1235	1235	4659	4659
Pseudo/Adj. R^2	0.0258	0.0203	0.0095	0.0076	0.0082	0.0049
Correctly Classified	88.11%		88.02%		92.68%	
AUROC	0.6205		0.5996		0.5711	
U-Test						
Slope at $Inequality^l$	-0.1689 * **	-0.0197 * **		-0.0064 *	-0.0994 * *	-0.0087 * *
	[− 4.0850]	[− 3.8232]		[− 1.3248]	[− 2.0645]	[− 1.9358]
Slope at $Inequality^h$	0.1457 * **	0.0170 * **		0.0002	0.1236 * **	0.0103 * **
	[2.6790]	[2.6773]		[0.0360]	[3.7346]	[3.1152]
SLM test for U shape	2.68	2.68		0.04	2.06	1.94
	{0.0037}	{0.0038}		{0.4860}	{0.0195}	{0.0265}

Notes: Robust standard errors are in parentheses, *t*-values are in square brackets, and *p*-values are in curly brackets. * , * *, and * ** indicate significance at the 10%, 5%, and 1% levels, respectively.

$$S.\ D.\ (Marginal \quad Effect) = \sqrt{Var(\beta_1) + (L.\ GINI)^2 Var(\beta_2) + 2(L.\ GINI)Cov(\beta_1, \beta_2)} \tag{12}$$

Thus, we construct the 95% confidence interval based on Eq. (12): ($Marginal\ Effect - 1.96 \times S.\ D.$, $Marginal\ Effect + 1.96 \times S.\ D.$). For example, in Column (1), we plot the marginal effect of Gini on the occurrence of systemic banking crises in Fig. 5, as well as the 95% confidence interval and the frequency of Gini. It also shows that the marginal effect is conditional on the value of Gini and turns from negative to positive with an increase in Gini. The marginal effect is negative and statistically significant if income distribution is relatively even, whereas the marginal effect tends to be positive when Gini is higher than a threshold (0.45). These results indicate that the impact of increased income inequality on financial risk depends on the initial level of income inequality.

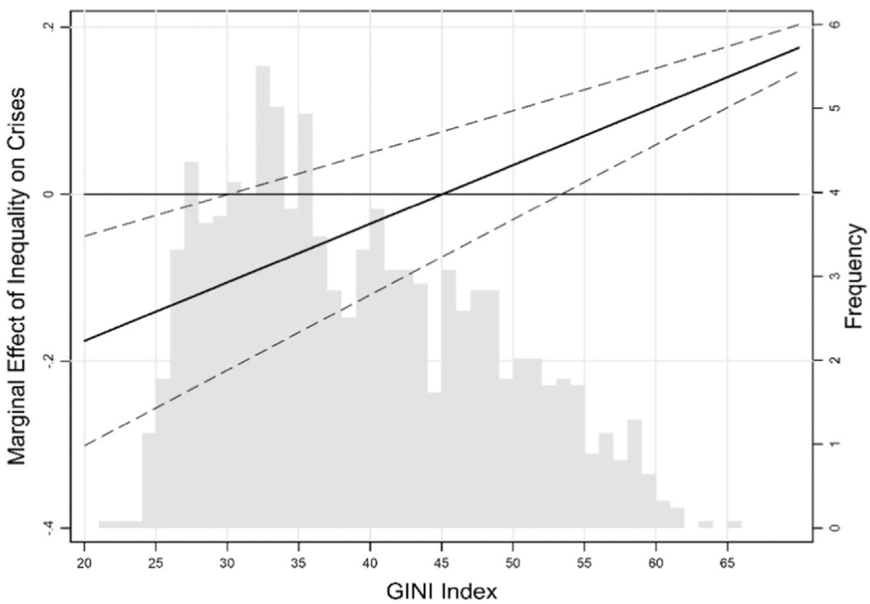

Fig. 5. Marginal effect of income inequality on the probability of systemic banking crises conditional on income inequality. Notes: The shaded area is the histogram of Gini. The dashed lines are 95% confidence bands.

Table 4
Results of the prediction test between the "credit" model and the "inequality" model.

Models	Obs.	AUROC
Credit model	1195	0.5815 *** (0.0252)
GINI model	1211	0.6205 *** (0.0252)
Equality test: χ^2	4.70 * *{0.0301}	

Notes: Robust standard errors are in parentheses, and *p*-values are in curly brackets. * * and * ** indicate significance at the 5% and 1% levels, respectively.

Fig. 5 displays a histogram of Gini, showing that our sample covers a large range of income inequality, from 0.21 to 0.658, and has a distribution with a slightly positive skew.

As the nonlinear relationship judged by the significant quadratic specification may be misleading (Lind and Mehlum, 2010), we apply the appropriate U-test to check for the existence of a U-shaped nexus. The results are shown in Table 3. In Column (1), the test yields a negative slope at $Inequality^l$ with a *t*-value of –4.0850 and a positive slope at $Inequality^h$ with a *t*-value of 2.679; thus, an inverted-U-shaped relationship exists. These results reject the null hypothesis—namely, that the relationship is monotonic—and show that the U-shaped nexus between income inequality and banking crises holds within at least the 95% confidence interval.

5.3. Predictive performance

Table 3 reports the percentage of crises that are correctly identified. The results suggest that our banking crisis prediction models have adequate predictive performance; the correctly classified rates of the proposed models are 88.11% (model 1), 88.02% (model 3), and 92.68% (model 5), showing that our models have an extremely low probability of making Type II errors.

Additionally, we use the AUROC statistics to estimate the predictive power of the models. The three models all have higher AUROC than the critical value of 0.5, which suggests that the proposed models can predict crises with an acceptable probability of Type I errors. Furthermore, in Table 4, we compare the predictive power of the widely used "credit" model with that of the "inequality" model (Column (1) in Table 3). The AUROC of the former is 0.5815 and that of the latter is 0.6205; the χ^2 equality test rejects the null hypothesis that the two models have no systemic differences in predictive power, which means that our proposed inequality model outperforms the credit model.

6. Empirical analysis: panel evidence

6.1. Baseline results

To rationalize our fixed-effect panel models (2) and (3), we first conduct the Hausman test to determine whether to include the fixed effect. The results are reported in Table 5. The null hypotheses are rejected in all cases, which implies that our model with the country-fixed effect is reasonable. The country-fixed effect includes some time-invariant but country-variant variables, such as the legal system and the political system. Laeven (2002) concludes that a legal system with proper enforcement of rules reduces the negative effects of deposit insurance on bank risk-taking. Additionally, Ashraf (2017) finds that stable political systems can strengthen credit market competition and lead banks to increase risk-taking. Therefore, our fixed-effect models can effectively control for these time-invariant but country-variant variables.

Table 6 reports the baseline results. The time- and country-fixed effects are controlled for in the regressions. The results provide stronger support for the U-shaped nexus confirmed in Section 5. The coefficients for the quadratic terms of income inequality for the three measures of income inequality are both positive and significant at least at the 5% confidence level, whereas the coefficients for income inequality are negative with different significance levels, suggesting that the effect of income inequality on the probability of systemic banking crises is U-shaped.

The threshold values of the income inequality measures in the logit model are as follows: Gini: 0.398 (=[–(–0.7960)/(2 × 0.0100)]/100, Column (1)); the top 10% income share: 32.35% (=[–(–0.8411)/(2 × 0.0130)]/100; Column (3)); and the standardized Gini index: 0.338 (=[–(–0.5275)/(2 ×0.0078)]/100, Column (5)). The threshold values of the income inequality measures for the LPM are as follows: the Gini index: 0.398 (=[–(–0.0477)/(2 × 0.0006)]/100, Column (2)); the top 10% income share: 32.33% (=[–(–0.0194)/(2 × 0.0003)]/100, Column (4)); and the standardized Gini index: 0.39 (=[–(–0.0236)/(2 × 0.0003)]/100, Column (6)). The coefficients of inequality squared are positive and significant, which implies that when income inequality is lower than the threshold, its marginal impact on the likelihood of a banking crisis is negative, whereas when it is higher than the threshold, its marginal impact is positive. Thus, we conclude that a threshold exists at which income inequality can be beneficial to financial stability: when income inequality skews to the left (right) of this threshold, the banking system tends to stabilize

Table 5
Results of the Hausman test for country-fixed effects.

Statistics	(1) LogitGINI (WDI)	(2) LPMGINI (WDI)	(3) Logit10%Top	(4) LPM10%Top	(5) LogitGINI (SWIID)	(6) LPMGINI (SWIID)
χ^2	5.86 * *	6.27 * *	7.84 * *	4.53 *	21.29 * **	21.56 * **

Notes: The null hypothesis of the Hausman test is "difference in coefficients is not systemic." * , * *, and * ** significance at the 10%, 5%, and 1% levels, respectively.

Table 6
Impact of income inequality on the occurrence of systemic banking crises: Baseline panel evidence.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
L.GINI(WDI)	-0.7960 * **	-0.0477 * **				
	(0.2299)	(0.0155)				
L.GINI2(WDI)	0.0100 * **	0.0006 * **				
	(0.0027)	(0.0002)				
L.10%Top(WDI)			-0.8411 * **	-0.0194 *		
			(0.2995)	(0.0116)		
L.10%Top2(WDI)			0.0130 * **	0.0003 * *		
			(0.0043)	(0.0002)		
L.GINI(SWIID)					-0.5678 *	-0.0236 * *
					(0.2704)	(0.0060)
L.GINI2(SWIID)					0.0078 * *	0.0003 * *
					(0.0036)	(0.0001)
Obs.	676	1211	676	1211	3170	4412
R^2		0.1699		0.1664		0.0562
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses. * , * *, and * ** indicate significance at the 10%, 5%, and 1% levels, respectively.

(destabilize) as inequality worsens. Furthermore, we can calculate the average income inequality for each country to evaluate its relative position to this threshold value of income inequality. We find that, in most Organization for Economic Cooperation and Development (OECD) countries (e.g., Germany, France, Japan, Australia, and Italy) and northern European countries (e.g., Denmark, Finland, Sweden, and Norway), income inequality is less than the threshold value, which means that in these countries, the probability of banking crises declines as income inequality worsens. In particular, the top 10% income share in the United States is around the threshold, whereas the other two measures are to the left of the threshold. In contrast, in the most developing countries (e.g., Brazil, China, India, Mexico, Peru, and Thailand), income inequality is substantially larger than the threshold, which alerts these countries that, to achieve financial stability, they should decrease income inequality.

These results confirm H3. The hypotheses by Rajan (2010) and in subsequent studies (Bordo and Meissner, 2012; Kirschenmann et al., 2016; Kumhof et al., 2015) show that income inequality has a linear causal effect on banking crises. These studies either support Rajan's hypothesis (Kirschenmann et al., 2016) or find no significant evidence for it (Bordo and Meissner, 2012). We find evidence of a nonlinear effect, and the results in Table 6 support H3.

6.2. *Robustness checks*

We check the stability of the baseline results by conducting the following robustness checks. First, we add the lagged banking crises to the regression model in order to capture the persistence of crises. Second, we introduce several widely used control variables in the banking crisis literature to the model. Third, as the sample contains many missing observations, data completion techniques are applied, and the results are produced with the new dataset. Fourth, we reduce the noise in the annual macroeconomic dataset using an averaging approach. Finally, we check the growth hypothesis. These robustness checks confirm the baseline results, indicating that income inequality is a significant and U-shaped driver of systemic banking crises.

6.2.1. *Persistence of systemic banking crises*

Chen (2015) shows that a year with (without) a banking crisis is more likely to be followed by a year with (without) a banking crisis. Therefore, we estimate the model by adding one-lagged term of the dependent variable (*L.Crisis*) to the econometric model. Thus, a dynamic panel model is established. Methods such as difference Generalized Method of Moments (GMM) and system GMM are usually applied to address any endogeneity issues arising from the introduction of a lagged dependent variable. However, applying these methods in models with a binary dependent variable is difficult (Chen, 2015). Therefore, following Chen (2015), we use a fixed-effect panel logit model and fixed-effect PLM as in the previous section. The results are presented in Table 7.

As reported in Table 4, the coefficients of the lagged crises are positive and significant at the 1% confidence level irrespective of the specification, indicating that systemic banking crises are persistent. Furthermore, the coefficients of income inequality and its quadratic term in Columns (1), (2), and (3) are significant at the 95% or 99% confidence intervals. Compared with the baseline results in Table 6, the significance of the variables of interest declines because of the absorptive effect of the lagged dependent variable. This evidence again confirms that income inequality plays a U-shaped role in shaping the probability of systemic banking crises.

6.2.2. *Control variables*

The analyses presented thus far do not account for the control variables. Table 8 presents the results with control variables, which suggest that income inequality significantly and nonlinearly affects the odds of systemic banking crises. All specifications show that the coefficients of income inequality and its quadratic term are significant, and their corresponding signs indicate that the effect of income inequality on the occurrence of banking crises is U-shaped. The threshold levels of income inequality are 0.421 in Column (1), 31.42% in Column (3), and 0.427% in Column (5), which are consistent with the baseline results.

Table 7
Impact of income inequality on the occurrence of systemic banking crises: Crisis persistence.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
L.Crisis	5.1987 * **	0.5929 * **	5.2621 * **	0.5943 * **	3.8117 * **	0.6182 * **
	(0.7010)	(0.0238)	(0.7082)	(0.0315)	(0.1839)	(0.0192)
L.GINI(WDI)	-0.5143 *	-0.0243 * *				
	(0.3017)	(0.0123)				
L.GINI2(WDI)	0.0070 *	0.0003 *				
	(0.0036)	(0.0001)				
L.10%Top(WDI)			-0.6352 * *	-0.0229 * *		
			(0.2834)	(0.0098)		
L.10%Top2(WDI)			0.0104 * *	0.0003 * *		
			(0.0044)	(0.0001)		
L.GINI(SWIID)					-0.4345 * *	-0.0211 * *
					(0.1871)	(0.0086)
L.GINI2(SWIID)					0.0071 * **	0.0001 * *
					(0.0009)	(0.0000)
Obs.	676	1211	676	1211	3170	4412
R^2		0.5351		0.5377		0.4494
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses. * , * *, and * ** significance at the 10%, 5%, and 1% levels, respectively.

Table 8
Impact of income inequality on systemic banking crises: Control variables.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
L.GINI(WDI)	-18.1971 * **	-0.1918 * *				
	(4.8861)	(0.0839)				
L.GINI2(WDI)	0.2162 * **	0.0021 * **				
	(0.0560)	(0.0009)				
L.10%Top(WDI)			-18.8521 * **	-0.2373 *		
			(6.3280)	(0.1204)		
L.10%Top2(WDI)			0.3000 * **	0.0035 *		
			(0.0890)	(0.0017)		
L.GINI(SWIID)					-7.6330 * ** (1.6379)	-0.2950 * **
						(0.0887)
L.GINI2(SWIID)					0.0894 * **	0.0034 * **
					(0.0200)	(0.0010)
L.GDPgrowth	-0.4657 * *	-0.0146 * **	-0.5127 * **	-0.0215 * **	-0.1128 * **	-0.0191 * **
	(0.1958)	(0.0046)	(0.1857)	(0.0054)	(0.0281)	(0.0037)
L.Credirgdp	0.0581 * **	0.0059 * **	0.0619 * **	0.0059 * **	0.1450 * **	0.0073 * **
	(0.0152)	(0.0020)	(0.0157)	(0.0021)	(0.0246)	(0.0017)
L.Curaccount	0.0476 (0.0498)	0.0077	0.0489	0.0085	0.1137 (0.0734)	0.0038
		(0.0059)	(0.0493)	(0.0068)		(0.0037)
L.GDPdeflator	0.0077 (0.0099)	0.0007	0.0045 (0.0095)	0.0003	0.0384	0.0000
		(0.0007)		(0.0006)	(0.0310)	(0.0003)
L.Govdebt	-0.0201	-0.0001	-0.0113	-0.0005	-0.0273	-0.0004
	(0.0169)	(0.0029)	(0.0278)	(0.0029)	(0.0177)	(0.0014)
L.ExplicitDGS	0.7173 (0.8564)	-0.0737	0.9490 (0.8175)	-0.0581	2.1312 * * (1.1054)	0.0180
		(0.1122)		(0.1105)		(0.0763)
L.Capitalgrowth	-0.0007	-0.0024	-0.0024	-0.0023	-0.0025	-0.0006 *
	(0.0204)	(0.0031)	(0.0197)	(0.0030)	(0.0114)	(0.0003)
L.M2growth	-0.0004	-0.0005	-0.0017	-0.0007	-0.0779 * **	-0.0001
	(0.0056)	(0.0009)	(0.0054)	(0.0009)	(0.0255)	(0.0003)
Obs.	441	284	441	284	405	717
R^2		0.2583		0.1840		0.0729
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses. * , * *, and * ** significance at the 10%, 5%, and 1% levels, respectively.

Table 9
Impact of income inequality on the occurrence of systemic banking crises: Controlling for current GDP growth.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
L.GINI(WDI)	-15.0370 * *	-0.1230 * **				
	(7.5664)	(0.0396)				
L.GINI2(WDI)	0.1985	0.0015				
	(0.1999)	(0.0009)				
L.10%Top(WDI)			-11.4365 * *	-0.1032 * *		
			(5.2877)	(0.0451)		
L.10%Top2(WDI)			0.1938	0.0022		
			(0.1424)	(0.0014)		
L.GINI(SWIID)					-2.2901 * **	-0.1625 * **
					(0.8770)	(0.0654)
L.GINI2(SWIID)					0.0281	0.0019
					(0.1117)	(0.0050)
GDPgrowth	-1.2715 * *	-0.0275 * **	-0.7418 * **	-0.0286 * **	-0.3446 * **	-0.0212 * **
	(0.6416)	(0.0057)	(0.2729)	(0.0058)	(0.0627)	(0.0028)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	441	284	441	284	405	717
R^2		0.2578		0.1247		0.1084
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses. * , * *, and * ** significance at the 10%, 5%, and 1% levels, respectively. Other controls include those in Table 8, except L.GDPgrowth.

Next, we add eight control variables, which have been extensively discussed in the literature, to the regressions. As shown in Table 8, lagged economic growth and the ratio of credit to GDP are the driving forces of systemic banking crises, as predicted by the Rajan and Kuznets effects. The coefficient of lagged GDP growth is negative and significant at the 1% or 5% confidence level, indicating that economic prosperity is beneficial to financial stability, and the risk of systemic banking crises is countercyclical, which is consistent with Demirgüç-Kunt and Detragiache (1998), von Hagen and Ho (2007), and Davis and Karim (2008). Accordingly, the ratio of lagged credit to GDP could significantly predict the occurrence of banking crises under any specifications, as found by Reinhart and Rogoff (2008) and Schularick and Taylor (2012).

In Table 8, we control for lagged GDP growth and find that it is a key factor that influences the occurrence of crises. H2 implies that if we control for current economic growth in the model, the coefficients of income inequality squared will become nonsignificant because the nonlinear effect of income inequality on crises will be absorbed by the economic growth. We test this implication in Table 9, in which the coefficients of current economic growth are positive and significant, whereas those of inequality squared are insignificant.

As shown in Table 8, after the control variables are added to our regressions, the number of observations declines significantly. A problem might arise if the missing observations are randomly distributed across countries and years. If the distribution of missing observations is not random, we cannot exclude the contingency of the results in Table 8. To address this problem, we adopt a counterfactual strategy by including the

Table 10
Results of the test of sample randomness.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
L.GINI(WDI)	-1.8587 * **	-0.1821 * *				
	(0.3744)	(0.0758)				
L.GINI2(WDI)	0.0210 * **	0.0020 * *				
	(0.0043)	(0.0008)				
L.10%Top(WDI)			-2.0117 * **	-0.0546 * *		
			(0.4495)	(0.0220)		
L.10%Top2(WDI)			0.0282 * **	0.0011 * *		
			(0.0064)	(0.0005)		
L.GINI(SWIID)					-4.4675 * **	-0.3792 * *
					(1.1209)	(0.1525)
L.GINI2(SWIID)					0.0528 * **	0.0045 * *
					(0.0138)	(0.0018)
Obs.	441	284	441	284	405	717
R^2		0.2222		0.3476		0.1690
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses. * , * *, and * ** significance at the 10%, 5%, and 1% levels, respectively.

Table 11
Impact of income inequality on the occurrence of systemic banking crises: Filling in missing data.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
L.GINI(WDI)	-0.5315 * **	-0.0348 * **				
	(0.1365)	(0.0090)				
L.GINI2(WDI)	0.0061 * **	0.0004 * **				
	(0.0016)	(0.0001)				
L.10%Top(WDI)			-0.8631 * **	-0.0091 *		
			(0.1806)	(0.0054)		
L.10%Top2(WDI)			0.0126 * **	0.0001 *		
			(0.0026)	(0.0001)		
L.GINI(SWIID)					-0.6549 * *	-0.0304 *
					(0.3280)	(0.0177)
L.GINI2(SWIID)					0.0093 * **	0.0004 * *
					(0.0036)	(0.0002)
Obs.	1727	2807	1794	2864	3480	4954
R^2		0.0788		0.0586		0.0401
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses. * , * *, and * ** significance at the 10%, 5%, and 1% levels, respectively.

observations employed in Table 8 but without the controls. The observations in Table 8 are largely different from those in Table 6. If the missing observations are not randomly distributed, the expected results should hold only in Table 8. If the expected results still hold, this provides a high degree of confidence that the missing observations are randomly distributed. The results in Table 10 are consistent with earlier results and support our expectations, suggesting that income inequality can robustly predict banking crises in a random sample.

6.2.3. *Filling in missing values*

Because of the asymmetric nature of the missing data and considering data availability, thus far the analyses have used only limited observations from the available dataset. The credibility of our baseline results may be low if too much information is missing. Thus, we use the Carryforward command in Stata to fill in the missing data by carrying observations forward from one observation to the next. This approach is especially applicable to panel data.

The results obtained using these carryforward data are presented in Table 11. The outcomes are consistent with the baseline results. The coefficients of income inequality and its quadratic term are both significant, and the corresponding signs indicate that the relation between income inequality and the occurrence of systemic banking crises is U-shaped. The thresholds of income inequality are 0.436, 34.25%, and 0.352 in Columns (1), (3), and (5), respectively. Thus, H3 is further confirmed.

6.2.4. *Removing noise from annual data*

Our analysis uses annual macroeconomic data, which are potentially noisy (Roine et al., 2009). Thus, we average the variables across each of the five non-overlapping years to reduce noise. The dependent variable—banking crises—equals one when the economy experiences at least one

Table 12
Impact of income inequality on the occurrence of systemic banking crises: Non-overlapping five-year averages, 1973–2017.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
GINI(WDI)	-0.3894 *	-0.0362				
	(0.2097)	(0.0326)				
GINI2(WDI)	0.0048 * *	0.0004				
	(0.0025)	(0.0004)				
10%Top(WDI)			-0.4022 * *	-0.0115		
			(0.2222)	(0.0438)		
10%Top2(WDI)			0.0063 * **	0.0002		
			(0.0012)	(0.0006)		
GINI(SWIID)					-0.3731 * *	-0.0614 * *
					(0.1735)	(0.0283)
GINI2(SWIID)					0.0055 * *	0.0009 * *
					(0.0023)	(0.0004)
Obs.	417	557	417	557	872	990
R^2		0.0749		0.0724		0.0880
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses. * and * * significance at the 10% and 5% levels, respectively.

Table 13
Impact of income inequality on systemic banking crises: Growth hypothesis.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
GINI(WDI)	-0.3894 *	-0.0362				
	(0.2097)	(0.0326)				
GINI2(WDI)	0.0048 * *	0.0004				
	(0.0025)	(0.0004)				
10%Top(WDI)			-0.4022 * *	-0.0115		
			(0.2222)	(0.0438)		
10%Top2(WDI)			0.0063 * **	0.0002		
			(0.0012)	(0.0006)		
GINI(SWIID)					-0.3731 * *	-0.0614 * *
					(0.1735)	(0.0283)
GINI2(SWIID)					0.0055 * *	0.0009 * *
					(0.0023)	(0.0004)
Obs.	417	557	417	557	872	990
R^2		0.0749		0.0724		0.0880
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses. * and * * significance at the 10% and 5% levels, respectively.

crisis during the corresponding five years and equals zero otherwise. The regression results using the post-treatment data are reported in Table 12. The baseline conclusion is robust in Columns (1), (3), (5) and (6), which indicates that income inequality is nonlinearly related to the occurrence of systemic banking crises. Although the coefficients of income inequality under specifications (2) and (4) are insignificant, their signs suggest that the relationship between income inequality and the occurrence of banking crises is nonmonotonic.

6.2.5. Distinguishing the level hypothesis from the growth hypothesis

Morelli and Atkinson (2015) distinguish the level and growth hypotheses: the former states that the level of income inequality contributes to banking instability, whereas the latter states that the change in income inequality contributes to banking instability. Intuitively, it is not possible to credibly judge whether economic actors care more about the absolute level of inequality or the change in inequality. Studies on the inequality–growth nexus have not reached a consensus on whether the level of inequality or the change in inequality should be used when the connection between income inequality and banking crises is investigated. For example, Acemoglu and Robinson (2002) use the level of inequality, whereas Banerjee and Duflo (2003) use the change in inequality. We follow Morelli and Atkinson (2015) and use both.

The foregoing analyses focus mainly on the level hypothesis. The results in Table 13 support the growth hypothesis. in the growth hypothesis, unlike the level hypothesis, the change in income inequality linearly causes the occurrence of banking crises during the sample interval. Although the coefficients of the quadratic term are significant in Column (1), the optimal income inequality calculated is outside the rational range.

Table 14
Impact of income inequality on systemic banking crises: Rare events.

Variables	(1) relogit	(2) cloglog	(3) relogit	(4) cloglog	(5) relogit	(6) cloglog
L.GINI(WDI)	-0.3197 * **	-0.6269				
	(0.0846)	(0.3914)				
L.GINI2(WDI)	0.0036 * **	0.0083 * *				
	(0.0010)	(0.0041)				
L.10%Top(WDI)			-0.4123 * ** (0.1280)	-− 0.7034		
				(0.4790)		
L.10%Top2(WDI)			0.0057 * **	0.0113 * *		
			(0.0020)	(0.0053)		
L.GINI(SWIID)					-0.2198 * *	-0.1629 * *
					(0.0507)	(0.0715)
L.GINI2(SWIID)					0.0028 * **	0.0025 * *
					(0.0007)	(0.0012)
Obs.	1211	588	1211	588	4412	2927
LL		-200.2253		-201.7154		-870.3794
Country-fixed effects	No	Yes	No	Yes	No	Yes
Time-fixed effects	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors are in parentheses. * , * *, and * ** significance at the 10%, 5% and 1% levels, respectively.

Table 15
Results of the use of selection of observables to assess bias due to unobservable variables.

Controls in the restricted set	Controls in the full set		(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
None	All controls in Table 8	inequality	1.0457	1.3310	1.0467	1.0890	1.0804	1.0870
		quadratic term of inequality	1.0485	1.4000	1.0453	1.0938	1.0956	1.0968

Notes: All regressions include the country- and time-fixed effects. The reported ratio is calculated as $\mid \beta^F/\beta^F - \beta^R \mid$.

6.2.6. *Rare event analysis*

As shown in Fig. 2, systemic banking crises are rare events (King and Zeng, 1999). Studies show that the traditional logistic procedure might underestimate the probability of rare events and that the commonly used data collection approach is inefficient. Accordingly, King and Zeng (KZ, 1999) propose an estimation strategy for producing less-biased and lower-variance estimates of logit coefficients and the corresponding variance-covariance matrix by correcting for small samples and rare events. Moreover, a complementary log–log model is typically used when one of the outcomes is rarer (systemic banking crises) than other outcomes (no crises).

Table 14 reports the results of the KZ estimates (relogit) and complementary log–log estimates (cloglog), which are essentially unchanged from our baseline results, showing that income distribution affects the probability of banking crises in a U-shaped manner.

6.2.7. *Assessment of bias from unobservable variables*

Despite our attempts to control for the observable variables, these estimation results may be biased by unobservable factors that moderate the relationship between banking crises and income distribution. Thus, in this section, we evaluate the likelihood that the results are biased by unobservable variables. We conduct this assessment following the approach by Nunn and Wantchekon (2011). The basic idea is that the selection of observables can be used to assess potential bias caused by unobservable variables. We calculate a ratio to measure the strength of the likely bias arising from the unobservable variables: the ratio is calculated as $\mid \beta^F/\beta^F - \beta^R \mid$, where β^F is the coefficient of income inequality in the full model, which includes the full set of controls as in Table 8, and β^R is the coefficient of income inequality in the restricted model, which includes the restricted set of control variables. The restricted model is the same as the model with no controls in Table 6. The rule is that when the ratio is larger, the bias from the unobservable is smaller. The underlying intuition is straightforward: a large numerator indicates that the effects of the unobservables on systemic financial crises are limited, and a small denominator implies that the addition of observable variables offers little information to explain systemic banking crises.

The ratios calculated are reported in Table 15, and none of the twelve ratios in it is less than one. The ratios range from 1.0453 to 1.4000, with a mean of 1.1216. Therefore, if the unobservable variables are key factors in systemic banking crises, their effects on the banking crises on average should be at least 1.1216 times greater than those of the observable variables. Generally, these results make it nearly impossible for the estimated effect of income inequality on the likelihood of banking crises to be fully driven by unobservables.

6.3. *A comparison with previous studies*

In Section 5.1, we assume that the differences in the conclusions in the literature regarding the relationship between income inequality and the occurrence of banking crises originates in sample selection. We formally test this assumption. We compare our study to five representative studies: Bordo and Meissner (BM, 2012) and Kirschenmann et al. (KMN, 2016), who use a sample of 14 developed countries for the period 1920–2007; Morelli and Atkinson (MA, 2015), who use a sample of 25 countries for the period 1900–2012; Perugini et al. (PHC, 2015), who use data for 18 OECD countries for the period 1970–2007; and Rhee and Kim (RK, 2018), who study 68 countries using data for the period 1973–2010. To enable fair comparisons between these five studies and our study, we use the same but modify their time period to match that of our sample.

First, we investigate the differences in the descriptive statistics between the five studies and our sample in Table 16. Bordo and Meissner (2012), Kirschenmann et al. (2016), and Morelli and Atkinson (2015) find little evidence of a link between income inequality and the occurrence of banking crises, whereas Perugini et al. (2015) and Rhee and Kim (2018) find a positive and significant link. The means of the three measures of inequality in Perugini et al. (2015) and Rhee and Kim (2018) are typically larger than those in Bordo and Meissner (2012), Kirschenmann et al. (2016), and Morelli and Atkinson (2015) and are close to those of our sample statistics in Table 1. This indicates that samples with average inequality higher than the threshold determined in this paper yields a positive link between income inequality and banking crises in our nonlinear framework.

Furthermore, we run the regression analysis as in our baseline specification but use the samples from each of the five studies. The results, shown in Table 17, support our assumption. We find a significant and positive nexus between income inequality and the occurrence of banking crises using the samples of PHC (2015) and RK (2018), whereas we find no significant link using the samples of the three other studies.

Table 16
Descriptive Statistics in Previous Studies.

Statistics	**BM (2012)/KMN (2016)**			**MA (2015)**			**PHC (2015)**			**RK (2018)**		
	GINI (WDI)	10%Top	GINI (SWIID)	GINI (WDI)	10%Top	GINI (SWIID)	GINI (WDI)	10%Top	GINI (SWIID)	GINI (WDI)	10%Top	GINI (SWIID)
Mean	31.6896	24.6500	28.8373	31.8150	24.8560	32.9427	37.2874	28.9677	38.9868	40.7813	31.6718	38.7180
Std. Dev.	3.7940	2.2637	3.8329	3.8169	2.3664	7.3458	9.7391	7.4216	3.9743	9.6482	7.2021	3.9599

Notes: BM (2012) is Bordo and Meissner (2012), KMN (2016) is Kirschenmann et al. (2016), MA (2015) is Morelli and Atkinson (2015), PHC (2015) is Perugini et al. (2015), and RK (2018) is Rhee and Kim (2018).

Table 17
Comparison of Our Results with Those in Previous Studies.

	BM (2012)/KMN (2016)			MA (2015)			PHC (2015)			RK (2018)		
Variables	GINI (WDI)	10%Top	GINI (SWIID)	GINI (WDI)	10%Top	GINI (SWIID)	GINI (WDI)	10%Top	GINI (SWIID)	GINI (WDI)	10%Top	GINI (SWIID)
L.GINI(WDI)	-0.6685			-1.0643			0.6034 * **			0.7542 * *		
	(1.9522)			(1.5290)			(0.1537)			(0.2972)		
L.GINI2(WDI)	0.0154			0.0231			-0.0079			0.0102		
	(0.0302)			(0.0178)			(0.0226)			(0.3623)		
L.10%Top(WDI)		-0.9657			-0.0494			2.9429 * **			0.7623 * *	
		(3.9734)			(1.4909)			(0.2553)			(0.3632)	
L.10%Top2(WDI)		0.0336			0.0144			-0.0508			0.0127	
		(0.0801)			(0.0214)			(0.0481)			(0.6035)	
L.GINI(SWIID)			-5.1706			-0.9916			5.2618 * **			4.1939 * **
			(6.7503)			(3.6879)			(1.2710)			(1.1105)
L.GINI2(SWIID)			0.1029			0.0154			0.0911			0.0695
			(0.3182)			(0.0488)			(0.2254)			(0.1902)
Obs.	121	121	535	204	204	840	147	147	673	305	305	765
LL	-45.8115	-44.0453	-55.3076	-13.9002	-16.3537	-154.5430	-61.4346	-60.2099	-77.3505	-66.2625	-67.8972	-88.2939
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: BM (2012) is Bordo and Meissner (2012), KMN (2016) is Kirschenmann et al. (2016), MA (2015) is Morelli and Atkinson (2015), PHC (2015) is Perugini et al. (2015), and RK (2018) is Rhee and Kim (2018). Robust standard errors are in parentheses, following Schularick and Taylor (2012). * * and * ** significance at the 5% and 1% levels, respectively.

Table 18

Impacts of credit and economic growth on the occurrence of systemic banking crises.

Variables	(1) Logit	(2) LPM	(3) Logit	(4) LPM	(5) Logit	(6) LPM
L.GINI(WDI)	-13.1037	-0.1123				
	(8.0391)	(0.0788)				
L.GINI2(WDI)	0.1544	0.0012				
	(0.0971)	(0.0009)				
L.10%Top(WDI)			-13.5331	-0.1439		
			(8.1037)	(0.1036)		
L.10%Top2(WDI)			0.1986	0.0023		
			(0.1241)	(0.0015)		
L.GINI(SWIID)					-5.6084	-0.2451
					(4.0937)	(0.1885)
L.GINI2(SWIID)					0.0673	0.0029
					(0.0530)	(0.0017)
GDPgrowth	-0.3284 * **	-0.0248 * **	-0.3252 * **	-0.0256 * **	-0.2890 * **	-0.0147 * **
	(0.0628)	(0.0089)	(0.0616)	(0.0089)	(0.0669)	(0.0045)
Creditgdp	0.0745 * **	0.0042 * *	0.0749 * **	0.0044 *	0.1037 * **	0.0067 * **
	(0.0136)	(0.0019)	(0.0136)	(0.0024)	(0.0169)	(0.0016)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	181	332	197	332	504	892
R^2		0.3045		0.2740		0.0870
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Other controls are same as in Table 8, except L.GDPgrowth and L.Creditgdp. * ** indicates significance at the 1% level.

Table 19

Impact of income inequality on credit and economic growth.

Variables	(1) Creditgdp	(2) GDPgrowth	(3) Creditgdp	(4) GDPgrowth	(5) Creditgdp	(6) GDPgrowth
L.GINI(WDI)	0.4805 * **	0.4115 * *				
	(0.1848)	(0.1941)				
L.GINI2(WDI)		-0.0058 * *				
		(0.0027)				
L.10%Top(WDI)			0.5151 * *	0.3194 * *		
			(0.1997)	(0.1472)		
L.10%Top2(WDI)				-0.0052 *		
				(0.0029)		
L.GINI(SWIID)					1.0061 * **	0.8857 *
					(0.3422)	(0.4613)
L.GINI2(SWIID)						-0.0085 * *
						(0.0041)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	276	287	276	287	718	741
R^2	0.0071	0.2654	0.0058	0.2840	0.0128	0.1623
Country-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors are in parentheses. * * and * ** significance at the 5% and 1% levels, respectively. The control variables are same as in Table 8.

We conclude that the five other studies use a subset of our sample and do not adequately consider the potential nonlinear relationship between income inequality and the occurrence of banking crises. Our results may help to reconcile the differences in the literature.

6.4. Mechanism identification

Following the methodology used to check the mediating effect, we perform two additional steps with the baseline results to identify the implicit mechanisms hypothesized in Section 2. First, we examine the direct effects of the ratio of current credit to GDP and current economic growth on the occurrence of banking crises. Second, we regress income inequality on the ratio of credit to GDP and economic growth, respectively. If the coefficients of the variables of interest are both significant with this procedure, the Rajan effect (H1) and the Kuznets effect (H2) are confirmed.

The results of the first step are reported in Table 18. After controlling for the ratio of current credit to GDP and current economic growth, we find that the effects of income inequality on the occurrence of crises are completely absorbed. The coefficients of lagged inequality and inequality squared are both insignificant. In Table 9, after controlling for current economic growth, the coefficients of inequality squared are insignificant, whereas the coefficients of lagged inequality remain significant. The implicit logic is that current economic growth and the ratio of current credit to GDP are direct drivers of crises, as discussed in Section 2. Thus, after we control for these drivers in regressions, the impacts of income inequality are completely absorbed. Thus, the proposed implicit mechanisms are partially confirmed.

Furthermore, we test the effects of lagged income inequality on the mediating variables, that is, the ratio of credit to GDP and economic growth. Table 19 presents the empirical results. In Columns (1), (3), and (5), the dependent variable is the ratio of credit to GDP. In Columns (2), (4), and (6), the dependent variable is economic growth. We find that income inequality significantly and positively drives the dynamics of the ratio of credit to GDP. These results, combined with those in Table 18, show that income inequality affects the risk of banking crises via an increase in the extension of credit, supporting H1 (Rajan effect). Additionally, the results demonstrate that income inequality is a significant and nonlinear determinant of economic growth. These results, combined with those in Table 18, support H2 (Kuznets effect), that is, the inverted-U-shaped nexus between income inequality and economic growth leads to a U-shaped relationship between income inequality and the occurrence of systemic banking crises. Hence, Table 19 confirms H1 and H2.

7. Conclusions

We study the role of income inequality in the occurrence of systemic banking crises. In a pioneering study, Rajan (2010) argues that income inequality seeded the 2008–2009 financial crisis. The central idea behind Rajan's argument is that income polarization can drive credit booms for many reasons, meaning that prosperity in credit markets is detrimental to financial stability. Some studies theoretically and empirically support the Rajan effect, whereas others find a negligible impact of income inequality on the occurrence of financial crises. Thus, the literature has not achieved a consensus.

In addition to examining the Rajan effect, we investigate another mechanism through which income inequality might affect the occurrence of systemic banking crises nonlinearly. We refly on theories in which the effect of income inequality on growth is nonlinear, conditional on the stage of economic development, combined with the linear effect of growth on financial stability. Therefore, income inequality could affect economic growth and then financial stability in a nonlinear way. We call this the Kuznets effect.

Furthermore, we use panel logit and linear probability models to investigate the effects of income distribution on the likelihood of systemic banking crises. We use a wide sample, covering 172 economies for the period 1970–2017. We show that, because of potential nonlinearity, the smaller subsamples used in previous studies may lead to misleading results, which might explain the variety of results in the literature regarding the link between income inequality and the occurrence of banking crises. We further evaluate the model's predictive efficiency using AUROC statistics and test the nonlinear relationship using the U-test.

First, we report the empirical results based on pooled data. The cross-sectional analysis supports our hypotheses. The nonparametric plot indicates that the nexus between income inequality and the occurrence of systemic banking crises is U-shaped. Then, the pooled regressions are used to identify the optimal level of income equality for financial stability, the nonlinear nexus is confirmed with the U-test, and the AUROC statistics reveal that our model outperforms a credit model that is widely used for forecasting banking crises.

Furthermore, the panel regression confirms our hypotheses. The baseline results show that income inequality plays a U-shaped role in driving systemic banking crises. Robustness checks are conducted—including considering the persistence of banking crises, adding control variables, filling in missing values, removing noise in annual data, disentangling the level and growth hypotheses, applying rare event strategy, and assessing bias from unobservable variables—and the baseline findings remain unchanged. Then we compare our results to those of previous studies. Finally, we identify the proposed mechanisms through which income inequality affects the occurrence of banking crises—that is, the Rajan and Kuznets effects—by applying a two-step mediating effect test.

This study has several implications for both policy and research. First, financial regulators and income distribution managers should coordinate with each other. Traditional financial regulation policies ignore the role of income distribution in shaping financial stability; at the same time, traditional income distribution policies pay limited attention to financial stability. Second, the robust U-shaped nexus indicates that, as much as possible, an optimal level of income inequality should be maintained for the purpose of financial stability, rather than aiming for absolute equality of income. Third, this study smooths over the differences in prior literature, which can be attributed to their overlooking nonlinear nexus.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (72003205), the General Project of Social Science Planning in Guangdong Province (GD22CYJ12), and the Humanities and Social Sciences Foundation of Chinese Ministry of Education (20YJC790142).

Appendix A

See Appendix Table A1 here.

Table A1
A list of the countries/regions in the sample.

Countries/Regions	Countries/Regions	Countries/Regions	Countries/Regions
Afghanistan	Ecuador	Libya	Samoa
Albania	Egypt	Lithuania	San Marino
Algeria	El Salvador	Luxembourg	São Tomé and Principé
Angola	Equatorial Guinea	Macao, China	Saudi Arabia
Antigua and Barbuda	Estonia	Madagascar	Senegal
Argentina	Eswatini	Malawi	Serbia
Armenia	Ethiopia	Malaysia	Seychelles
Aruba	Fiji	Maldives	Sierra Leone
Australia	Finland	Mali	Singapore
Austria	France	Malta	Slovakia
Bahamas	French Polynesia	Mauritania	Slovenia
Bahrain	Gabon	Mauritius	Solomon Islands
Bangladesh	The Gambia	Mexico	Spain
Barbados	Georgia	Micronesia	Sri Lanka
Belarus	Germany	Moldova	St. Kitts and Nevis
Belgium	Ghana	Mongolia	St. Lucia
Benin	Greece	Morocco	St. Vincent/Grenadines
Bhutan	Grenada	Mozambique	Sudan
Bolivia	Guinea	Myanmar	Sweden
Bosnia and Herzegovina	Guinea-Bissau	Nauru	Switzerland
Brazil	Guyana	Nepal	Syrian Arab Republic
Brunei	Haiti	Netherlands	Tajikistan
Bulgaria	Hong Kong, China	New Caledonia	Tanzania
Burkina Faso	Hungary	New Zealand	Thailand
Burundi	Iceland	Nicaragua	Togo
Cabo Verde	India	Niger	Tonga
Cambodia	Indonesia	Nigeria	Trinidad and Tobago
Cameroon	Iran	North Macedonia	Tunisia
Canada	Iraq	Norway	Turkey
Central African Republic	Ireland	Pakistan	Uganda
Chad	Israel	Palau	Ukraine
Chile	Italy	Panama	United Arab Emirates
China	Jamaica	Papua New Guinea	United Kingdom
Colombia	Japan	Paraguay	United States
Comoros	Jordan	Peru	Uruguay
Congo	Kazakhstan	Philippines	Vanuatu
Costa Rica	Kenya	Poland	Venezuela
Côte d′Ivoire	Korea	Portugal	Vietnam
Croatia	Korea	Qatar	Yemen
Cyprus	Kuwait	Russia	Zambia
Czechia	Kyrgyz Republic	Rwanda	Zimbabwe
Denmark	Lao PDR		
Djibouti	Latvia		
Dominica	Lebanon		
Dominican Republic	Liberia		

References

Acemoglu, D., Robinson, J.A., 2002. The political economy of the Kuznets curve. Rev. Dev. Econ. 6 (2), 183–203.
Angkinand, A.P., Willett, T.D., 2011. Exchange rate regimes and banking crises: The channels of influence investigated. Int. J. Financ. Econ. 16 (3), 256–274.
Anundsen, A.K., Gerdrup, K., Hansen, F., Kragh-Sørensen, K., 2016. Bubbles and crises: The role of house prices and credit. J. Appl. Econ. 31 (7), 1291–1311.
Aoki, K., Nikolov, K., 2015. Bubbles, banks and financial stability. J. Monet. Econ. 74, 33–51.
Ashraf, B.N., 2017. Political institutions and bank risk-taking behavior. J. Financ. Stab. 29, 13–35.
Bandyopadhyay, D., Basu, P., 2005. What drives the cross-country growth and inequality correlation? Can. J. Econ. 38 (4), 1272–1297.
Banerjee, A.V., Duflo, E., 2003. Inequality and growth: What can the data say? J. Econ. Growth 8 (3), 267–299.
Bartscher, A., Kuhn, M., Schularick, M., Steins, U., 2020. Modigliani meets Minsky: Inequality, Debt, and Financial Fragility in America, 1950-2016. CEPR Press Discussion Paper No. 14667,.
Beck, T., Levine, R., Levkov, A., 2010. Big bad banks? The winners and losers from bank deregulation in the United States. J. Financ. 65 (5), 1637–1667.
Beck, T., Chen, T., Lin, C., 2016. Financial innovation: The bright and the dark sides. J. Bank. Financ. 72, 28–51.
Behringer, J., van Treeck, T., 2018. Income distribution and the current account. J. Int. Econ. 114, 238–254.
Belabed, C.A., Theobald, T., van Treeck, T., 2018. Income distribution and current account imbalances. Camb. J. Econ. 42, 47–94.
Bordo, M.D., Meissner, C.M., 2012. Does inequality lead to a financial crisis? J. Int. Money Financ. 31 (8), 2147–2161.
Brunnermeier, M., Rother, S., Schnabel, I., 2020. Asset price bubbles and systemic risk. Rev. Financ. Stud. 33 (9), 4272–4317.
Brunnermeier, M.K., Sannikov, Y., 2014. A macroeconomic model with a financial sector. Am. Econ. Rev. 104, 379–421.
Cairó, I., Sim, J.W. Income inequality, financial crises, and monetary policy. FEDS Working Paper No. 2018–048, 2018.
Chen, Q., 2015. Climate shocks, state capacity and peasant uprisings in North China during 25–1911 CE. Economica 82, 295–318.
Choi, I., 2001. Unit root tests for panel data. J. Int. Money Financ. 20, 249–272.
Coibion, O., Gorodnichenko, Y., Kudlyak, M., Mondragon, J., 2020. Greater inequality and household borrowing: new evidence from household data. J. Eur. Econ. Assoc. 18 (6), 2922–2971.

Davis, E.P., Karim, D., 2008. Could early warning systems have helped to predict the sub-prime crisis? Natl. Inst. Econ. Rev. 206 (1), 35–47.
Demirgüç-Kunt, A., Detragiache, E., 1998. The determinants of banking crises in developing and developed countries. IMF Econ. Rev. 45 (1), 81–109.
Diamond, D.W., Dybvig, P.H., 1983. Bank runs, deposit insurance, and liquidity. J. Political Econ. 91 (3), 401–419.
Dufour, M., Orhangazi, Ö., 2016. Growth and distribution after the 2007–2008 US financial crisis: Who shouldered the burden of the crisis? Rev. Keynes. Econ. 4, 151–174.
Foellmi, R., Zweimüller, J., 2006. Income distribution and demand-induced innovations. Rev. Econ. Stud. 73 (4), 941–960.
Gali, J., 2014. Monetary policy and rational asset price bubbles. Am. Econ. Rev. 104, 721–752.
Galor, O., 2000. Income distribution and the process of development. Eur. Econ. Rev. 44 (4–6), 706–712.
Galor, O., Moav, O., 2004. From physical to human capital accumulation: Inequality and the process of development. Rev. Econ. Stud. 71 (4), 1001–1026.
Galor, O., Tsiddon, D., 1997. Technical progress, mobility, and economic growth. Am. Econ. Rev. 87 (3), 363–382.
Gertler, M., Klenow, P., 2019. Economic fluctuations and growth. NBER Report. 2, 1–13.
Gourinchas, P.O., Obstfeld, M., 2012. Stories of the twentieth century for the Twenty-First. Am. Econ. J.: Macroecon. 3 (4), 226–265.
Grossman, G.M., Helpman, E., 2018. Growth, trade, and inequality. Econometrica 86 (1), 37–83.
von Hagen, J., Ho, T., 2007. Money market pressure and the determinants of banking crises. J. Money, Credit, Bank. 39 (5), 1037–1066.
Herrera, H., Ordonez, G., Trebesch, C., 2020. Political booms, financial crises. J. Political Econ. 128 (2), 507–543.
Iacoviello, M., 2008. Household debt and income inequality, 1963–2003. J. Money, Credit, Bank. 40 (5), 929–965.
Jordà, Ò., Schularick, M., Taylor, A.M., 2011. Financial crises, credit booms, and external imbalances: 140 years of lessons. IMF Econ. Rev. 59, 340–378.
Jordà, Ò., Schularick, M., Taylor, A.M., 2015. Leveraged bubbles. J. Monet. Econ. 76, S1–S20.
Kauko, K., 2014. How to foresee banking crises? A survey of the empirical literature. Econ. Syst. 38 (3), 289–308.
Kim, T., Koo, B., Park, M., 2013. Role of financial regulation and innovation in the financial crisis. J. Financ. Stab. 9 (4), 662–672.
King, G., Zeng, L., 1999. Logistic regression in rare events data. Department of Government, Harvard University, (available from ⟨http://GKing.Harvard.Edu⟩).
Kirschenmann, K., Malinen, T., Nyberg, H., 2016. The risk of financial crises: Is there a role for income inequality? J. Int. Money Financ. 68, 161–180.
Krueger, D., Perri, F., 2006. Does income inequality lead to consumption inequality? Evidence and theory. Rev. Econ. Stud. 73 (1), 163–193.
Kumhof, M., Rancière, R., Winant, P., 2015. Inequality, leverage, and crises. Am. Econ. Rev. 105 (3), 1217–1245.
Kuznets, S., 1955. Economic growth and income inequality. Am. Econ. Rev. 45 (1), 1–28.
Laeven, L., 2002. Bank risk and deposit insurance. World Bank Econ. Rev. 16 (1), 109–137.
Laeven, L., Valencia, F., 2013. Systemic banking crises database. IMF Econ. Rev. 61 (2), 225–270.
Lin, S.C., Huang, H.C., Kim, D.H., Yeh, C.C., 2009. Nonlinearity between inequality and growth. Stud. Nonlinear Dyn. Econ. 13 (2) Article 3.
Lin, Y.C., Huang, H.C., Yeh, C.C., 2014. Inequality-growth nexus along the development process. Stud. Nonlinear Dyn. Econ. 18 (3), 237–252.
Lind, J.T., Mehlum, H., 2010. With or without U? The appropriate test for a U-shaped relationship. Oxf. Bull. Econ. Stat. 72 (1), 109–118.
Lo Duca, M., Peltonen, T.A., 2013. Assessing systemic risks and predicting systemic events. J. Bank. Financ. 37 (7), 2183–2195.
Malinen, T., 2016. Does income inequality contribute to credit cycles? J. Econ. Inequal. 14 (3), 309–325.
Minsky, H.P., 1983. The financial instability hypothesis: An interpretation of Keynes and an alternative to “Standard” theory. In: Wood, J.C., John Maynard Keynes (Eds.), Critical assessments. Macmillan, London.
Mitkov, Y., 2020. Inequality and financial fragility. J. Monet. Econ. 115, 233–248.
Morelli, S., Atkinson, A.B., 2015. Inequality and crises revisited. Econ. Polit. 32 (1), 31–51.
Nunn, N., Wantchekon, L., 2011. The slave trade and the origins of mistrust in Africa. Am. Econ. Rev. 101 (7), 3221–3252.
Obstfeld, M., Rogoff, K. Global imbalances and the financial crisis: Products of common causes. CEPR Discussion Papers, 2009.
Park, J.Y., Phillips, P.C.B., 2000. Nonstationary binary choice. Econometrica 68 (5), 1249–1280.
Perugini, C., Hölscher, J., Collie, S., 2015. Inequality, credit and financial crises. Camb. J. Econ. 40 (1), 227–257.
Pollin, R., 1994. Marxian and Post-Keynesian developments in the sphere of money, credit and finance: Building alternative perspectives in monetary macroeconomics. In: Glick, M.A. (Ed.), Competition, technology and money. Classical and Post-Keynesian perspectives. Edward Elgar, Aldershot.
Rajan, R.G., 2010. Fault Lines: How Hidden Fractures Still Threaten the World Economy. Princeton University Press, Princeton.
Reich, R.B., 2010. Aftershock: The Next Economy and America’s Future. Random House, New York.
Reinhart, C.M., Rogoff, K.S., 2009. The aftermath of financial crises. Am. Econ. Rev. 99 (2), 466–472.
Reinhart, C.M., Rogoff, K.S., 2011. From financial crash to debt crisis. Am. Econ. Rev. 101 (5), 1676–1706.
Rhee, D.E., Kim, H., 2018. Does income inequality lead to banking crises in developing countries? Empirical evidence from cross-country panel data. Econ. Syst. 48 (2), 206–218.
Roine, J., Vlachos, J., Waldenström, D., 2009. The long-run determinants of inequality: What can we learn from top income data? J. Public Econ. 93, 974–988.
Schularick, M., Taylor, A.M., 2012. Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870–2008. Am. Econ. Rev. 102 (2), 1029–1061.
Solt, F., 2020. Measuring income inequality across countries and over time: The standardized world income inequality database. Soc. Sci. Q. 101 (3), 1183–1199.
Stiglitz, J.E., 2012. The Price of Inequality: How Today's Divided Society Endangers Our Future. W.W. Norton,.
Wang, S., Luo, R., 2023. Income distribution, financial liberalisations and banking stability: Theory and international evidence. Int. J. Financ. Econ (Forthcoming).